Location Choices over the Life Cycle: The Role of Relocation for Retirement

Kristina Komissarova*

August 18, 2022

Abstract

This paper develops an overlapping-generations spatial equilibrium model to study the allocation of workers and retirees across cities. I show that differences in locations' wages and amenities for the young and the old lead to geographic sorting by age, and that sorting is dampened by significant costs of moving for retirement. I estimate that the reduction in these costs observed between 1960 and 2010 allowed individuals to increasingly separate locations for work and retirement which generated 3.7 percent welfare gains. I further evaluate local and aggregate effects of place-based policies that promote the relocation of retirees to less productive cities.

*PhD student Department of Economics, New York University; kristina.komissarova@nyu.edu. I am very grateful to my advisors Simon Gilchrist, Ricardo Lagos, and Sharon Traiberman for their guidance and support throughout the project. I also owe special thanks to Katarína Borovičková, Evgenii Fadeev, Jingting Fan (discussant), Sophie Mathes (discussant), Felix Wellschmied, and participants at the NYU Stern Macro Lunch, NYU Macro Workshops, NYU Applied Micro Workshop, 15th and 10th Urban Economics Association Meetings, and 2021 Summer School in Urban Economics.

1 Introduction

An individual's location choice is shaped by the trade-off between job opportunities, housing prices, and the access to amenities (Rosen (1979), Roback (1982), Allen & Arkolakis (2014), Desmet & Rossi-Hansberg (2013)). The determinants of this trade-off can vary significantly over the life cycle: after retirement individuals no longer directly benefit from a location's productivity, but they are still affected by local prices. Moreover, younger and older individuals might value local amenities differently due to changes in preferences and spending patterns with age. Yet, the literature on economic geography has mostly focused on a static location choice, abstracting from modelling the life cycle.

Understanding the life cycle component of location choices becomes especially important with the ageing population trend. Indeed, the growing population of retirees might fuel the development of cities attractive for the old: for instance, the retirement community, The Villages, FL, was the fastest-growing metro over the period 2010-2020. Some states and cities put in place a range of policies aimed at attracting relocating retirees.[1] However, local and aggregate effects of such policies have been largely unexplored.

In this paper, I study the role of the life cycle for individuals' location choices and for the distribution of economic activity across space. First, I provide new empirical evidence on sorting by age across U.S. commuting zones (CZs). Second, I develop a structural framework of location choice over the life cycle: individuals live for two periods and make a work-location decision in the first period, and a retirement-location decision in the second. I use my framework to estimate the values of CZs separately for its younger and older residents, and to quantify the gains from declining mobility costs for moving for retirement. Finally, I quantify the effects of place-based policies that incentivize the relocation of the old to less productive cities.

I start by documenting sorting patterns by age and CZ productivity, proxied by its wage: a 1% higher CZ wage is associated with a 0.9% lower CZ old-to-young ratio.[2] I show that this cross-sectional relationship is related to the changes in the mobility patterns with age: more productive locations receive an inflow of individuals in the younger age cohort and an outflow of individuals in the older age cohort. Moreover, I document that a CZ age composition is correlated with the composition of its establishments: CZs with a higher old-to-young ratio have a higher number of health-care and recreation centers per capita, but a lower number of clothing stores, movie theatres, and childcare centers per capita. Using consumer expenditure

[1] For example, Alabama's state employee retirement system finances a series of golf courses in local cities. City Tallahassee in Florida gives direct reallocation subsidies for retirees to move there: "Florida city Wants more Retirees and is Going After Them"(NYT article)

[2] The old-to-young ratio is defined as the ratio of 60+ years old to $25-59$ years old. I adjust CZ wages for the differences in their demographic composition to proxy for location-specific wage component. I will further discuss the connection between spatial heterogeneity in wages and productivity in the model.

data, I show that younger and older age groups differ in their demand for the services provided by these establishments.[3]

The empirical evidence on the geographic sorting by age emphasizes two key changes over the life cycle that might affect the location choice: a decline in working hours, which makes a location's productivity less important, and changes in expenditures on local services, which affect preferences for local amenities. Also, young and old individuals might differ in their preferences for location's natural and climate amenities.[4] I develop a theoretical framework by introducing these life cycle changes into a spatial equilibrium model, building on Allen & Arkolakis (2014), Monte et al. (2018), and Redding & Sturm (2008). The key model ingredients are the following: (i) individuals choose locations for both work and retirement periods subject to mobility costs of moving for retirement; (ii) workers and retirees differ in their preferences for exogenous amenities of locations and for endogenous non-tradeable services and (iii) local services are provided by amenity developers and adjust endogenously to local expenditures. I also assume that the productivity of the tradeable sector is subject to agglomeration spillovers as in Allen & Arkolakis (2014). In the model, dynamics arise for two reasons: first, the young anticipate costly relocation for retirement, and second, workers only earn wages in the first period and so there is a savings decision embedded in their work decision.

I estimate the model using cross-sectional data on U.S. CZs, data on mobility patterns around a retirement age, and data on the consumer spending patterns. In particular, I estimate a model-implied gravity equation for moving for retirement, and I find that mobility costs for the old increase significantly with the distance between CZs. I further calibrate the fixed component of mobility costs for the old and the composite amenities for the young and the old across CZs to match data on different locations' young and old residents, wages, and rents, and to match the aggregate share of individuals who relocate for retirement. I document significant heterogeneity in CZ amenities for the old relative to the young, which is partially explained by differences in CZs geography, climate, and composition of non-tradeable establishments. Next, I quantify the endogenous adjustment of amenities to the location's age composition by estimating the heterogeneity between younger and older age groups in expenditure shares on local services.

Using the fully calibrated model, I show that differences in exogenous productivities and exogenous amenities across locations lead to geographic sorting by age. Sorting is further amplified by an endogenous adjustment of location's productivity and amenities to location's age composition. To build intuition for this result, consider the CZs of New York (NY) and

[3] I show that older individuals spend more on health care services but less on clothing, childcare and food outside of home.

[4] I show that the CZ old-to-young ratio is positively correlated with better climate (warmer winters and cooler summers) and natural amenities (percentage of water area and topography). This is consistent with the findings in Lu (2020) and Schneider & Green (1992) that older individuals increasinly move to locations with better climate and natural amenities

Sarasota (FL). According to the model estimates, New York is endowed with a higher exogenous productivity, whereas Sarasota is endowed with better exogenous amenities for the old. Because retirees no longer benefit from local wages, they are less willing to pay high prices in exchange for productivity. Moreover, retirees are attracted by better amenities for the old in Sarasota. Therefore, the model predicts a net reallocation of retirees from a more productive to a more amenable city. As a result, housing constraints relax in New York, tighten in Sarasota, and push more young towards New York. This further magnifies productivity differences in two cities due to the presence of productivity spillovers. Additionally, the composition of local establishments shifts with local expenditures, which makes New York even more amenable for the young (e.g., more bars and restaurants), and Sarasota even more amenable for the old (e.g., more health care facilities).

I further illustrate how the model explains the documented empirical facts on geographic sorting by age. I compute that if the only change over the life cycle was the decline in working hours after retirement, that is if the old had the same preferences as the young, then the elasticity of the old-to-young ratio with respect to the CZ wage would be equal to -0.7 as compared to -0.9 in the data. Therefore, most of the observed negative relationship between CZ wages and demographics is attributed to the decline in working hours after retirement, and the rest is attributed to relative amenities for the old being better in less productive locations. Next, I find that the endogenous adjustment of local services to the differences in expenditure shares between the young and the old explains half of the documented correlation between the CZ old-to-young ratio and the CZ composition of establishments.

The model estimates show that the value of a given location differs significantly for its younger and older residents. This implies large gains from separating locations for work and retirement periods. I estimate that moving from an area at the 25th percentile of the value for the old to an area at the 75th percentile is equivalent to increasing the individual's income by a factor of two. This large dispersion in the location-specific utilities of the old is sustained by significant mobility costs that retirees face. This also implies that location decisions for the work period depend on how easy it is to move from the location to places that are valuable for retirement. In particular, the model estimates show that at least thirty percent of the spatial variation in working age population across U.S. CZs can be explained by the variation in the location's value for the old.

Using the MSA-level data, I show that sorting by age and location's productivity significantly increased from 1960 to 2010. I also document that the share of individuals who moved between states around the retirement age increased by 20 percent from 1960 to 2000.[5]

[5] Using U.S. Census samples for years 1960 and 2010, I compute the share of individuals in the age group $57-77$ who changed their state of residence over the 5 year window. Same for the age group $25-45$. The change in the mobility of the younger age group was around 0 percent. The comparison is available only for

Using the modelling framework, I decompose the changes in geographic sorting by age over time into the changes caused by mobility frictions and into the changes caused by location-specific factors. I compute that the model implied decline in mobility frictions explains around 50 percent of the increase in sorting over the 1960-2010 period. I estimate that the resulting welfare gains from higher life cycle mobility are equal to 3.7 percent. The key source of these gains is that lower mobility frictions allow the old to increasingly choose locations that are cheaper and more amenable for the old, and also allow the young to increasingly choose locations that are more productive and more amenable for the young.

Next, I evaluate place-based policies that subsidize the reallocation of retirees to less productive cities through a universal lump-sum tax paid by workers. I show that such policies contribute to a more equal distribution of amenities across cities, and, as compared to the place-based policies targeted at workers, lead to an increase in the aggregate productivity. Another key difference between policies targeted at workers and the ones targeting retirees is that the old bring their accumulated savings with them and in this way contribute more to local expenditures. However, as in the case with workers, policies that incentivise the old to relocate to less productive cities are not welfare improving, because the efficient allocation does not feature higher sorting of younger and older individuals. Finally, I show that locally financed policies aimed at attracting retirees cannot generate growth in locations' young and old population.

Relationship to the literature. The paper builds on the literature on quantitative spatial models that include static environments Roback (1982), Redding & Sturm (2008), Allen & Arkolakis (2014), Ahfeldt et al. (2015), Redding & Sturm (2020), and recent dynamic models Desmet & Rossi-Hansberg (2014), Eckert & Peters (2018) and Allen & Donaldson (2018). I extend existing frameworks by modelling the individual's life cycle: I introduce the location choices for both work and retirement periods, and allow for the transfer of individual's accumulated savings across locations and time.[6] I show that the extended model remains tractable, and allows for the usage of model inversion techniques to recover the amenities for the young and the old across locations.

The paper is related to the recent studies that model the location choice as a part of the individual's dynamic problem (Bilal & Rossi-Hansberg (2018), Giannone et al. (2019) and Heise & Porzio (2022)). These papers focus on understanding the role of financial constraints, moving constraints, and housing for location decisions of individuals and firms. In my paper, I contribute to this line of work by studying how the changes in working hours and preferences over the life

1960 and 2000 Censuses due to the data availability issues.

[6] In the models introduced in Desmet & Rossi-Hansberg (2014), Eckert & Peters (2018) and Allen & Donaldson (2018) individuals consume all of the earnings made during the period, and there are no changes in working hours and preferences over time.

cycle affect location choices. Bilal & Rossi-Hansberg (2018) do not model changes in working hours and preferences with age. Giannone et al. (2019) model features the decline in working hours after retirement, but younger and older individuals have same preferences for locations' amenities. Another major difference is that the model in this paper has closed form solutions for the distribution of individuals across locations, so I can invert the model to solve for unobserved amenities for the young and the old across U.S. CZs. Model tractability also allows me to analyze efficiency properties of my model in order to discuss optimal sorting of the young and the old, and policies geared toward that.

A growing literature considers how heterogeneous preferences for consumption amenities reinforce sorting of high and low skilled households within or across cities. (Diamond (2016), Guerrieri et al. (2013), Couture et al. (2019), Su (2020), Hoelzlein (2019), Eckert (2019)). Closest to this paper, Almagro & Dominguez-líno (2019) identify heterogeneous response of amenities to the demographic composition of city's neighbourhoods. This paper further studies whether city-level amenities adjust to a city's age composition, and quantifies the contribution of this mechanism to sorting of the young and the old across cities. I follow the recent studies that microfound amenity spillovers through the endogenous entry of establishments producing non-tradable goods (Almagro & Dominguez-líno (2019), Couture et al. (2019), Hoelzlein (2019)). I use the life cycle changes in the individual's spending profiles (Aguiar & Hurst, 2013) to quantify the endogenous adjustment of amenities to the differences in the preferences of the young and the old.

The paper complements the research on the local and aggregate effects of place-based policies. Kline & Moretti (2014) and Glaeser & Gottlieb (2008) argue that due to the constant elasticity of spatial spillovers there are no gains from reallocating workers across cities. Fajgelbaum & Gaubert (2020) and Rossi-Hansberg et al. (2019) show that due to the differences in spatial spillovers that heterogeneous workers generate, there is a room for place-based policies that target sorting of workers across space. In this paper, I extend their results for the dynamic life cycle setting, where individuals contribute to local spillovers differently as they age. In addition, I model endogenous amenities through the entry of local establishments, rather than introducing them in a reduced form way. Consistent with the findings in the literature, I show that the presence of productivity spillovers and location preference shocks requires optimal place-based policies targeted at workers. I further show that differences between the young and the old in spatial spillovers don't require additional policies targeted at sorting by age and location.

The paper relates to the literature that studies the determinants of the mobility flows for retirement using individual-level mobility data around the retirement age (Lu (2020), Schneider & Green (1992), Duncombe et al. (2001)). Retirees were found to be attracted to places with favourable climate and natural amenities and to places with lower taxes and lower living costs. Rosenthal & Strange (2004) show that couples near retirement age tend to move away from places with favourable business environments and towards places with high values of amenities. To the best of my knowledge, this paper is the first to document the pattern of geographic sorting by age and location's average wages in the U.S and its changes over time. Further, as estimated in Koşar et al. (2021), younger and older individuals value some of location's attributes differently.[7] In this paper, I use aggregate data on the demographics of U.S. CZs combined with the spatial structure of the model to recover the differences in CZ amenities for younger and older residents.

The remainder of the paper proceeds as follows. Section 2 describes data and definitions used in the paper. Section 3 presents empirical analysis with the description of empirical facts. Section 4 introduces the theory and discusses its efficiency properties. Sections 5 and 6 describe the calibration of the model and its quantitative analysis. Section 7 describes counterfactual exercises. Section 8 concludes.

2 Data

This paper uses the 5 percent samples of the U.S. Censuses for the years 1960, 1970, 1980, 1990, 2000 and 1 percent samples of American Community Survey for the years 2012 -2016 which I aggregate into a single observation corresponding to year 2014. The data is from Integrated Use Microdata Series (IPUMS) (Ruggles et al., 2015). The IPUMS data provide detailed economic and demographic information at the individual level, including data on wages, housing costs, age, education, and geographic location of residence.

To proxy for amenities, I use establishment-level data from the County Business Patterns for year 2014 and climate and geography data from United States Department of Agriculture.[8] The expenditure-level data come from the NBER extracts of the Consumer Expenditure Survey (CEX), which includes all waves from 1980 through 2003. The CEX data provide household-level observations on expenditures made in a wide set of consumption categories.

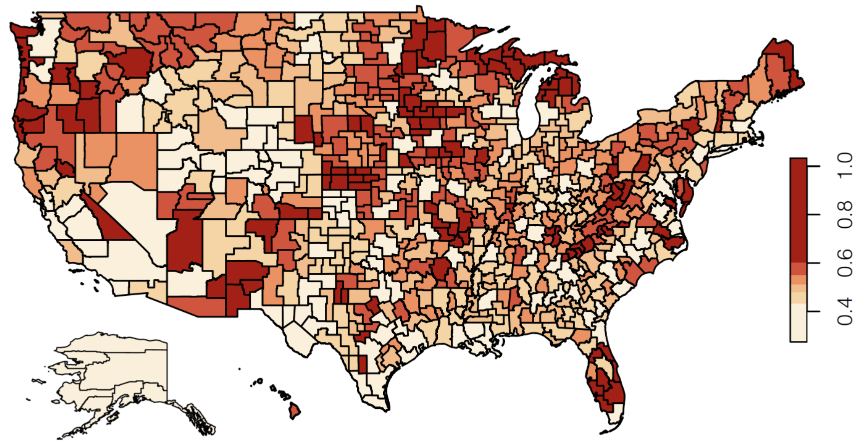
As a unit of georgraphy, I use commuting zones (CZs) which are designed to proxy for local labor markets. To proxy for CZs productivities, I use IPUMS data on average wages earned by CZs workers.[9] I adjust wages for the differences in worker's age, sex, race, nativity and highest completed education. To proxy for CZs prices, I use IPUMS data on average rents paid by CZ

[7] Koşar et al. (2021) employ a stated-preference approach using two waves from the New York Fed's Survey of Consumer Expectations to estimate individual's preferences for location's attributes. The find that older individuals have higher willingness to pay for lower crime and lower taxes.

[8] https://www.ers.usda.gov/data-products/natural-amenities-scale

[9] By workers I define individuals of age 17-64, who work more than 44 weeks and more than 34 hours a week.

Figure 1: Demographic composition of US CZs (age 60+ to age $[25, 60)$)



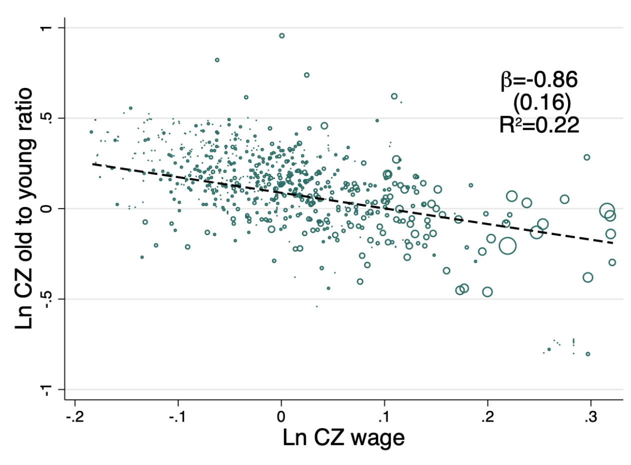
Notes: The figure shows the distribution of the ratio of older residents (age 60+) to younger residents (age 25-59) across 741 U.S. CZs for the year 2014 (average across 2021-2016 ACS)

households. I adjust rents for the observed quality characteristics of the rented unit so that to compare prices of similar units of housing across locations. To evaluate CZs age composition, I construct two age groups: the younger age group includes individuals whose age is between 25 and 59 years, and the older age group includes individuals older than or equal to 60.[10] More details on the data are in Appendix A.

3 Motivating empirical evidence

This section documents new facts on geographic sorting by age. First, I show that there exists significant dispersion in the old-to-young ratios across U.S. CZs. Second, I document that the CZ wage is negatively related to its old-to-young ratio. I show that this cross-sectional relationship is consistent with the changes in mobility patterns with age: high wage CZs receive an inflow of individuals in the younger age cohort and an outflow of individuals in the older age cohort. Third, I show that the CZ old-to-young ratio is correlated with the composition of its establishments in a way that is related to the differences in expenditure patterns between the young and the old.

Figure 2: Sorting of the young and the old by productivity



Notes: The figure plots CZ ratio of individuals older than 60 to individuals between ages 25 and 60 against CZ average wages earned by workers. Both variables are logs of the averages of demeaned values across 1980, 1990, 2000 and 2014 samples. CZ wages are adjusted for the differences in CZ demographics. β stands for the OLS regression coefficient and robust standard errors are in parenthesis. Regressions are weighted by location's average population across same samples. The size of each CZ circle is proportionate to its average population.

3.1 Geographic sorting by age across U.S. CZs

Figure 1 shows the distribution of the older age group relative to the younger age group across U.S. CZs. There exists significant heterogeneity in the old-to-young ratios which vary from around 0.28 to almost 1. The highest ratio is in CZ Ocala in Florida, which is the home to one of the largest retirement communities in the U.S., Villages, with lots of amenities specifically designed for older residents.[11] The lowest ratio is in CZ Kodiak in Alaska.

To evaluate the magnitude of the geographic sorting by age, I compute the average dissimilarity index for the younger and the older age groups across Census samples of 1980, 1990, 2000 and 2014.[12] The resulting value is 0.085, which indicates that 8.5% of older individuals need to be reallocated in order to distribute them evenly across all U.S. CZs. The same index for high and low skill workers, geographic sorting of which has been widely studied in the literature, is equal to 0.143. While the degree of geographic sorting by age is lower than the degree of geographic sorting by skill, it is still large, and its magnitude might be

[10] The age cut-off 60 is lower than average retirement age of 64 in the US to capture the relocation of individuals in anticipation of retirement.

[11] Vilages is an age-restricted community, that legally discriminates on the basis of age, and limits residency only to individuals older than 55. The available amenities include golf and country clubs, swimming pools and other recreational activities.

[12] The dissimilarity index is defined as $\frac{1}{2}\left|\frac{L_i^o}{\sum_i L_i^o} - \frac{L_i^y}{\sum_i L_i^y}\right|$ where L_i^o is the size of the older age group in the location i, and L_i^y is the size of the younger age group in i

dampened by mobility frictions that older individuals face.

3.2 Sorting by age and CZ wages

One potential reason behind geographic sorting by age is that after retirement individuals no longer benefit from locations' job opportunities. In this section, I explore whether geographic sorting by age is systematically related to the differences in locations' wages. For each CZ, I compute its old-to-young ratio and the average wages earned by its residents. I adjust wages for the observable workers' characteristics to proxy for the location-specific wage component.[13] As the goal is to study the long-term relationship between the two variables, I compute their averages across Census samples of 1980, 1990, 2000, and 2014. Figure 2 shows that the CZ old-to-young ratio is significantly and negatively related to the CZ wage with the elasticity of -0.86. As shown in Appendix B, the result is robust in the specification without weighting, robust to considering retirees instead of old, and to considering different age cut-offs. I also show that the pattern of sorting remains strong even when the CZ wage is proxied by its wage 44 years ago, suggesting that persistent component of differences in location-specific productivities is driving the sorting pattern.[14]

Next, I study whether the observed cross-sectional relationship between the CZ wage and its age composition is supported by the data on the changes in the mobility patterns with age. First, I explore the reallocation patterns of different age cohorts across U.S. CZs over a ten year period. I consider a cohort of older individuals (ages between 47 and 56) and follow them in the next decade when they potentially retire (ages between 57 and 66). Similarly, I consider a cohort of younger individuals (ages between 15 and 24) and follow them in the next decade when they potentially start working. Denote by $L^c_{m,t}$ the population size of a cohort c in location m and Census year t. For each CZ m, I compute the ratio $L^c_{m,t+10}/L^c_{m,t}$ to proxy for the reallocation patters of the cohorts: a higher ratio implies that CZ m has a higher net inflow of individuals of cohort c over the ten year period. I use average values for 1980-1990, 1990-2000 and 2000-2010 Censuses to partially smooth the time effects for a particular cohort. Figure 3 shows that high wage CZs on average experience an outflow of individuals from the older cohort and an inflow of individuals from the younger cohort, suggesting that individuals move to high wage locations when young and then relocate to lower wage locations for retirement.

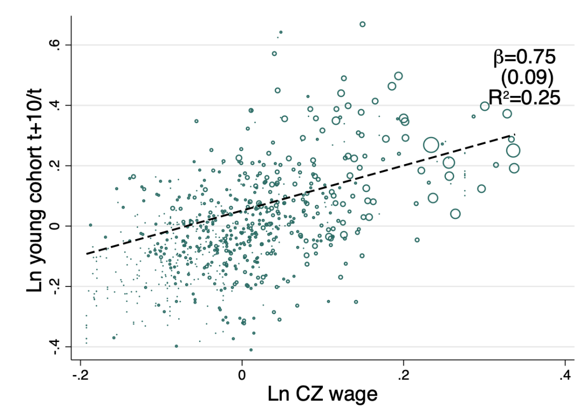
In Appendix C, I use data on individual's location five years ago to show that older movers have a lower probability of moving to a more productive location but a higher probability of moving to a location with a higher share of the old. I also show that the largest change in

[13] Controls include age, sex, race, nativity, and highest completed education. See Appendix A for the details.

[14] In Appendix B, I show that the CZ old-to-young ratio in 2014 is significantly and negatively related to the CZ wage in 1970 with the elasticity of -0.6.

Figure 3: Reallocation of young and old cohorts across CZs

(a) An older cohort



(b) A younger cohort

Notes: The figure plots the CZ's growth in the population of a given cohort over the ten year period against the CZ's average wages earned by workers in periods t and $t+10$. Both variables are logs of the averages of demeaned values across 1980-1990, 1990-2000 and 2000-2010 samples. An older cohort denotes individuals who fall into the age category 47-57 in the Census year t and a younger cohort denotes individuals who fall into the age category 15-25 in the Census year t. CZs' wages are adjusted for the differences in CZs' demographics. β stands for the OLS regression coefficient and robust standard errors are in parenthesis. Regressions are weighted by CZ's average population across same samples. The size of each CZ circle is proportionate to its average population.

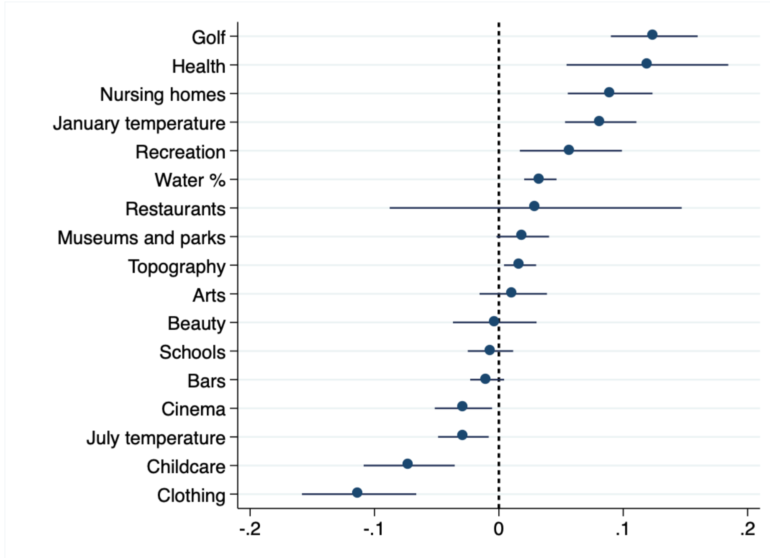
mobility patterns happens around the retirement age (age group 56-65).

3.3 CZ age composition and its amenities

Another potential reason behind geographic sorting by age is that younger and older individuals value locations' amenities differently. This section explores whether locations' age composition is related to the observable measures of its amenities. I consider climate and natural variables from USDA and measures of the variability of local services. The latter is approximated by the number of CZ establishments in different non-tradeable industries per thousand residents. Figure 4 shows the results from the regression of the CZ old-to-young ratio on a set of CZ amenities controlling for CZ wages, population, and prices. I find that CZs with a higher ratio of old to young residents have a significantly higher number of health-care centers, nursing homes and recreational facilities (such as golf and country clubs) per capita, but a lower number of clothing stores, childcare centers and bars per capita. In line with the results of previous studies (Lu (2020), Schneider & Green (1992)), CZs with a higher old-to-young ratio tend to have more favourable climate (warmer winters and cooler summers) and more favourable natural amenities (higher percentage of water area and higher topography).

There might be two potential reasons why the composition of local establishments correlates

Figure 4: CZ age composition and its amenities



Notes: The figure plots point estimates and 95% confidence intervals from the single regression of the log of CZ share of old (residents of $age \geq 60$ relative to residents $25 \leq age < 60$) on the set of local amenities. Logs of CZ population, wages and rents are included as controls. Standard errors are robust to heteroskedasticity. Demographics data is from 2012-2016 ACS. Establishment counts are from CBP year 2014. Data on natural amenities is from USDA. Details on the construction of establishment categories are in appendix A.4

with demographics. First, locations can differ in exogenous factors that affect the variety of some services more than the others. For example, some places can be more suitable for the entry of health care services which makes them relatively more attractive for older residents. Second, the variety of local services can adjust endogenously to local demographics through the entry of firms in response to local expenditures. One of the goals for the structural model in the paper is to separate these two effects to better understand the relationship between CZ age and its amenities.

3.4 Differences in expenditure shares by age

If younger and older individuals differ in their preferences for locally provided services, these differences should be reflected in their expenditure patterns. Using the consumer expenditure data (CEX), I compute expenditure shares at the household level on a set of consumption categories roughly corresponding to local industries considered in the previous section. The available groups are: food outside of home, alcohol consumption outside of home, clothing, personal care services, health care services, low level education, childcare services, recreational services, and nursing home services. More details on the data construction and the composition of the categories are in the Appendix A.3. To evaluate the effect of the household's head age

TABLE 1

AGE AND EXPENDITURE SHARES ON LOCAL SERVICES

	(1) Childcare	(2) Clothing	(3) Bars	(4) Restaurants	(5) Recreation	(6) NursingH	(7) Beauty	(8) Health
Age $\geq$ 60	-0.96	-0.73	-0.63	-0.54	-0.58	0.22	0.23	1.26
	(0.014)	(0.012)	(0.02)	(0.015)	(0.015)	(0.008)	(0.014)	(0.014)
Obs	77842	77842	77842	77842	77842	77842	77842	77842
R squared	0.05	0.18	0.12	0.16	0.17	0.3	0.03	0.17
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Robust standard errors are given in parenthesis. Dependent variable is the log of the expenditure share on a particular spending category. For the cases when expenditure shares are zero, I bottom code the expenditure data at 1 dollar. Regressions include a set of controls for sex, race, education, and marital status. Category restaurants stands for the food outside of home, nightclubs stands for the alcohol consumption outside of home, and beaty stands for personal care services. More details on the construction of spending categories are in Appendix A.3

on the expenditure shares, I estimate the following specification for each service category:

$$log(s^n_{it}) = \alpha + \beta \text{age}_{\geq 60} + \delta X_{it} + \psi_t + \epsilon_{it}$$

where $s^n_{i,t}$ is the share of the consumption category n in individual's i total expenditures at year t, $\text{age}_{\geq 60}$ is the dummy for the age 60 or older, ψ_t is the time fixed effect and X_{it} if the set of demographic controls for race, sex, marital status and education.

Table 1 presents the estimation results. In line with the cross-sectional evidence on CZ amenities and its age composition, I find that the older demographic group spends a significantly higher share of their expenditures on health-care services, nursing homes, and significantly lower share on childcare, restaurants , nightclubs and clothing. The magnitude of the differences is large: for example, older individuals spent nearly twice more on health-care services and twice less on childcare services. These effects can be directly related to changes in family composition and health care needs with age.

Declining expenditure shares on clothing and restaurants are in line with the findings in Aguiar & Hurst (2013), where the authors argue that these services are complementary to working hours. Similarly, expenditures on other categories, such as personal care services, might increase with age due to their complementarity with leisure time. The main discrepancy with the previous findings is that the share spent on recreation declines with age. This result might mask significant heterogeneity in subcomponents of the recreation consumption group in the data. [15]

The empirical results above are consistent with the evidence documented in Su (2019) on the frequency of visits to local amenities by different age groups. She finds that the older age group (≥ 40) visits medical facilities and personal care facilities more often, while restaurant and bars less often as compared to the younger age group.

4 Theoretical Framework

This section introduces a model to formalize how changes in working hours and preferences for locations' amenities over the life cycle affect location choices and spatial equilibrium. The set-up shares many features of the models in Allen & Arkolakis (2014), Allen & Donaldson (2018), Redding & Sturm (2008) and Monte et al. (2018). The key new element is that individuals live for two periods: "young" and "old", and they choose locations for both periods subject to mobility frictions for relocating when old. The model allows for the heterogeneity in individual's preferences, in locations' amenities for the young and the old, and in locations' productivities. Further, I model an endogenous adjustment of locations' amenities to expenditures of locations' residents, similar to Almagro & Dominguez-líno (2019) and Couture et al. (2019), and endogenous adjustment of locations' productivities to local employment as in Allen & Arkolakis (2014). The sections below describe the model set-up, how the spatial equilibrium is determined, and discuss the efficiency properties of the model.

4.1 Setup

The economy consists of a set of discrete locations, which are indexed by $i = 1, ..., J$. Locations differ in terms of their exogenous land endowments, exogenous productivities in the tradeable sector, and exogenous amenities.

There are three types of producers in each location. First, perfectly competitive firms produce a single freely tradeable good using local labor with the production function featuring agglomeration externalities. Second, each location has construction firms who use a local land endowment and a final good to produce housing with a decreasing returns to scale production technology. Finally, each location has a free entry of amenity developers who compete for local housing to produce differentiated varieties in two categories of non-tradeable services.

The economy is populated by overlapping generations of individuals. Each cohort has an initial size L, and is replaced by a new cohort after T periods. I assume that individuals can be in two states: "young" and "old". In the young state individuals inelastically supply one unit

[15] While it includes recreational activities such as fees for sporting events and country clubs, it also includes expenditures on pets, repair of recreational equipment e.t.c.

of labor in the tradeable sector in return for the location-specific wages, and in the old state they retire and receive no labor income. Individuals in the two states differ in their preferences for location-specific exogenous amenities and in their preferences for two categories of locally provided services.

Agents live for $T = t_y + t_o$ periods in total, where the first t_y periods correspond to the young state, and the rest t_o periods to the old state. They choose a pair of locations (i, j), where i is the location for the first t_y young periods and j is the location for the following t_o old periods. The first location choice is free of costs, but reallocation for the old period is subject to location pair specific mobility costs μ_{ij}.

The model structure implies that in each period of time the economy is populated by $t_y L$ young individuals and $t_o L$ old individuals. I denote by $\rho = t_o/t_y$ the relative length of the old period, which is also the relative number of old individuals in the economy. The model is static and aims at capturing the long-run steady state of the economy. For simplicity, there is no capital accumulation and the gross saving rate is equal to one.

4.1.1 Consumer preferences

The preferences of a consumer ω who lives in location i when young and in location j when old are determined as:

$$U_{ij\omega} = t^y ln(u_i^y \epsilon_{i\omega}^y) + t^o \ln(u_j^o \epsilon_{j\omega}^o) - t^o \ln(\mu_{ij}) \tag{1}$$

where u_i^y and u_i^o are the common components of the per period utility of young and old residents in location i, μ_{ij} is the cost of moving from i to j in terms of the consumption units of the old period, $\epsilon_{i\omega}^y$ and $\epsilon_{j\omega}^o$ are idiosyncratic preferences of individual ω for living in location i when young and in location j when old. This formulation implicitly assumes that the discount rate and the gross interest rate are both equal to one, so that the consumption choices within the young and the old states are the same. For more details see Appendix D.1.

Define the preference shock for each location path $\epsilon_{ij\omega}$ as $\ln \epsilon_{ij\omega} = t^y \left(\ln \epsilon_{i\omega}^y + \rho \ln \epsilon_{j\omega}^o \right)$. I assume that $\epsilon_{ij\omega}$ is drawn independently from an extreme-value (Fréchet) distribution:

$$F(\epsilon_{ij\omega}) = e^{\epsilon_{ij\omega}^{-\theta}}, \qquad \theta > 1$$

where $\theta > 1$ controls the dispersion of idiosyncratic utility. I assume that individuals learn their realisation of the preference shocks for each path (i, j) in the young state before making location choices.

The per period common utility of young and old consumers in location i is:

$$u_{it}^y = A_i^y (c_t^y)^{\delta_c} (h_t^y)^{\delta_h} (Q_{it}^y)^{\delta_b} \qquad u_{it}^o = A_i^o (c_t^o)^{\delta_c} (h_t^o)^{\delta_h} (Q_{it}^o)^{\delta_b} \tag{2}$$

where A_i^y and A_i^o are the location-specific utility shifters for the young and the old, c is the consumption of the tradeable good, h is the consumption of housing, and Q_i is the consumption of the bundle of non-tradeable services in location i. I assume that $\delta_h + \delta_c + \delta_b = 1$.

In section 3.4, I showed that younger and older demographic groups systematically differ in their expenditures on local services. To account for this, I assume that there are two categories of services $s = 1, 2$ that enter the bundle of non-tradeables in each location. The first one ($s = 1$) corresponds to the services that are valued more by the young, and the second one ($s = 2$) corresponds to services that are valued more by the old :

$$Q_i^y = Q_{1i}^{\alpha_1^y} Q_{2i}^{1-\alpha_1^y} \qquad Q_i^o = Q_{1i}^{\alpha_1^o} Q_{2i}^{1-\alpha_1^o}$$

where $\alpha_1^y > \alpha_1^o$. Within a group of services competitive amenity developers in location i offer varieties that are imperfect substitutes. Individuals have CES preferences over the varieties with the common elasticity of substitution parameter σ:

$$Q_{si} = \left(\int_{m=1}^{N_{si}} q_{ms}^{\frac{\sigma-1}{\sigma}} \, dm \right)^{\frac{\sigma}{\sigma-1}}$$

where q_{ms} denotes the quantity of the variety m consumed in category s, and N_{si} denotes the number of varieties in category s in location i.

4.1.2 Production of the tradeable good

A continuum of firms in each location produces under perfectly competitive conditions a single homogeneous product that is freely tradeable across locations. The production function is linear in labor with location-specific productivity level z_i:

$$y_i(\omega) = z_i l_i(\omega)$$

The productivity level z_i is determined as:

$$z_i = Z_i \left(\int l_i(\omega) d\omega \right)^{\gamma^P} = Z_i \left(L_i^y \right)^{\gamma^P} \tag{3}$$

where Z_i is the exogenous component that can vary across locations, and γ^P is the parameter that governs any potential agglomeration externalities operating through the size of the local production, modelled in a structural way as in Allen & Arkolakis (2014) and Allen & Donaldson (2018). As only young individuals supply labor, the agglomeration spillover is determined by the number of the young in location i - L_i^y. The price of the final good is the same across all locations and is normalised to be equal to one.

4.1.3 Housing sector

Producers in the construction sector in location i use intermediary input I_i and the fixed supply of land $\overline{H}_i$ to produce housing H_i according to the following decreasing returns to scale production function:

$$H_i = \overline{H}_i I_i^{\frac{1}{1+\epsilon_h}}$$

Intermediary input is produced one to one with the final consumption good $I_i = c_{ih}$. Housing units are rented to location's residents and to amenity developers at the rental rate R_i. The profit maximisation problem of the producers is:

$$\max_{I_i} R_i \overline{H}_i I_i^{\frac{1}{1+\epsilon_h}} - I_i$$

I assume that construction firms are owned by local landlords who spend profits on the consumption of the tradeable good.

4.1.4 Production of non-tradeable services

There is an infinite mass of potential amenity developers outside of the city who can produce a single variety in sector s. To produce q units of a variety, a developer needs to rent kq units of housing, so the marginal cost of production is location-specific and equal to kR_i. A developer in location i producing variety m in category s solves the following profit maximisation problem:

$$\max_{p_{msi}} q_{msi}(p_{msi})(p_{msi} - kR_i)$$

Amenity developers can freely enter each category of services in each location. To produce a variety they need to pay location-service specific operating costs F_{si} in terms of the rented housing units. By N_{si} I denote the equilibrium number of firms in service category s in location i.

4.2 Consumer problem and indirect utility

Consider an individual ω who chooses location i for the work period and location j for the retirement period. When young the individual inelastically supplies one unit of labor in a location of choice and earns per period location-specific wage w_i. This income is spent on the consumption of the tradeable good, housing and local services during young and old periods.

By solving consumer problem, I derive the following expression for the indirect utility:

$$V_{ij\omega} = \ln\left(\frac{w_i a_i^y}{R_i^{(1-\delta_c)}}\right) + \rho \ln\left(\frac{w_i a_j^o}{R_j^{(1-\delta_c)}}\right) - \rho \ln(\mu_{ij}) + \ln(\epsilon_{ij\omega}) \tag{4}$$

The utility increases in wage w_i which determines consumption levels in both periods, and it decreases in rental prices R_i and R_j which determine housing and non-tradeable prices in young and old periods correspondingly. The indirect utility is further shifted by location-specific composite amenities for the young and the old a_i^y and a_j^o. These amenities include both exogenous amenities of locations and endogenous variety of local services:

$$a_i^y = A_i^y \left(N_{1i}^{\alpha_1^y} N_{2i}^{(1-\alpha_1^y)}\right)^{\frac{\delta_b}{\sigma-1}} \qquad a_i^o = A_i^o \left(N_{1i}^{\alpha_1^o} N_{2i}^{(1-\alpha_1^o)}\right)^{\frac{\delta_b}{\sigma-1}} \tag{5}$$

Endogenous number of establishments N_{si} affects consumer utility through a love of variety effect. The strength of this effect is governed by the elasticity of substitution σ and by the share of income spent on non-tradeable group $\delta_b \alpha_s^a$ by age group $a = y, o$. Using the free entry condition, I derive the following expression for the equilibrium number of firms:

$$N_{si} = \frac{\frac{t^y}{1+\rho}\left(L_i^y \delta_b \alpha_s^y w_i + \rho \sum_j L_{ji} \delta_b \alpha_s^o w_i\right)}{F_{si} R_i \sigma} \tag{6}$$

where L_{ij} denotes the number of individuals choosing location path $\{i, j\}$. The number of firms in each service category is proportional to local expenditures on this service category and inversely proportional to local housing prices, similar to Krugman (1979). As the young and the old differ in their preferences, location's demographic composition affects the composition of its establishments. Additionally, the location of the origin of the older residents affects the number of local firms by determining the retirees' expenditures.

Full derivations of indirect utility and the equilibrium number of firms are in Appendix D.2.

4.3 Consumer location choice problem

I will further work with the exponential transformation of the original utility function. Denote by W_{ij} the exponent of the deterministic part of the indirect utility (30) :

$$W_{ij} = \frac{w_i^{(1+\rho)} a_i^y (a_j^o)^\rho}{R_i^{1-\delta_c} R_j^{\rho(1-\delta_c)} \mu_{ij}^\rho}$$

Individuals choose a location path that maximizes their utility:

$$\max_{ij} W_{ij}\epsilon_{ij\omega}$$

Due to the Freshet assumption on the distribution of the preference shocks, the share of individuals that chooses location path (i, j) is given by:

$$\frac{L_{ij}}{L} = \frac{W_{ij}^{\theta}}{\sum_{k=1}^{J}\sum_{m=1}^{J} W_{km}^{\theta}} \tag{7}$$

The expected lifetime utility of individuals can be computed as:

$$U = E(\max_{ij} W_{ij}\epsilon_{ij\omega}) = (\sum_{k=1}^{J}\sum_{m=1}^{J} W_{km}^{\theta})^{\frac{1}{\theta}} \tag{8}$$

Combining optimal location choices (7), indirect utility (30), and expected utility (8), I can write bilateral migration flows for retirement as follows:

$$L_{ij} = \mu_{ij}^{-\theta\rho} \underbrace{\left(\frac{w_i^{(1+\rho)} a_i^y}{R_i^{1-\delta}}\right)^{\theta}}_{\text{Value of i for y}} \underbrace{\left(\frac{a_j^o}{R_j^{(1-\delta)}}\right)^{\rho\theta}}_{\text{Value of j for o}} L U^{-\theta} \tag{9}$$

Migration flows are higher into destination locations j that have better amenities for the old and lower rents, out of origin locations that are more attractive for the young, and among location pairs that have lower mobility costs.

4.4 Housing market

Profit maximisation problem of firms in the construction sector leads to the following housing supply function:

$$H_i = d_h \overline{H}_i \left(R_i\right)^{\frac{1}{\epsilon_h}} \tag{10}$$

where $d_h = (1 + \epsilon_H)^{\frac{1}{\epsilon_h}}$. Parameter ϵ_h governs the inverse elasticity of the housing supply.

Given rental rate R_i, the amount of housing demanded by consumers is:

$$H_i^h = t^y \sum_j L_{ij} h_{ij}^y + t^o \sum_j L_{ji} h_{ji}^o = \frac{t^y}{(1+\rho)R_i} \left(L_i^y \delta_h w_i + \rho \sum_j L_{ji}\delta_h w_j\right)$$

The amount of housing demanded by amenity developers is equal to the housing used in the production of local services including fixed entry costs:

$$H_i^d = t^y L_i^y k \sum_{s=1}^{2} \int_{m=1}^{N_{si}} q_{ms,ij}^y + t^o \sum_j L_{ji} k \sum_{s=1}^{2} \int_{m=1}^{N_{si}} q_{msji}^o + \sum_{s=1}^{2} N_{si} F_{si}$$

The housing clearing condition is then:

$$H_i = H_i^h + H_i^d$$

Using (26) and (29), I simplify the condition to:

$$H_i = (1 - \delta_c) \frac{t^y}{(1+\rho) R_i} \left(L_i^y w_i + \rho \sum_j L_{ji} w_j \right) \tag{11}$$

Fixed fraction $\delta_h + \delta_b = (1 - \delta_c)$ of local expenditures is spent on local housing either directly through housing consumption or indirectly through the consumption of local services.

4.5 Equilibrium

Definition 1. *For any time-invariant geography vector* $\{\overline{H}_i, Z_i, A_i^y, A_i^o, \mu_{ij}\}$ *a steady-state equilibrium is a vector of endogenous variables* $\{L_i^y, L_i^o, R_i, w_i, N_{1i}, N_{2i}, a_i^y, a_i^o\}$ *such that for any* i*:*

1. *Firms in the tradeable sector earn zero profits:* $w_i = z_i$*. Local productivity* z_i *satisfies equation (3);*

2. *The mobility flows between locations are consistent with optimal location choices (9);*

3. *The number of the young and the old in each location is determined as:*

$$L_i^y = t_y \sum_j L_{ij} \quad L_i^o = t_o \sum_j L_{ji} \tag{12}$$

4. *Housing supply* H_i *satisfies (10);*

5. *Housing markets clear such that equation (11) holds;*

6. *The number of firms in each service category is determined by the free entry condition of amenity developers such that equation (29) holds;*

7. *Local amenities satisfy equation (31).*

From equilibrium conditions, I can express the number of the young and the old in each location as a function of location's amenities, wages and rents:

$$\frac{L_i^y}{L} = t_y \underbrace{\left(\frac{w_i^{(1+\rho)} a_i^y}{R_i^{1-\delta_c} U}\right)^{\theta}}_{\text{Value for young}} \underbrace{\sum_j \left(\frac{(a_j^o)^\rho}{R_j^{\rho(1-\delta_c)} \mu_{ij}^\rho}\right)^{\theta}}_{\text{Retirement option}} \tag{13}$$

$$\frac{L_i^o}{L} = t_o \underbrace{\left(\frac{a_i^o}{R_i^{(1-\delta_c)}}\right)^{\theta\rho}}_{\text{Value for old}} \underbrace{\sum_j \left(\frac{w_j^{(1+\rho)} a_j^y}{R_j^{1-\delta_c} \mu_{ji}^\rho U}\right)^{\theta}}_{\text{Proximity to young}} \tag{14}$$

The presence of mobility costs for retirement make location choices for the young and the old periods a part of the single dynamic problem. The number of the young in each location depends not only on the attractiveness of this location for the young, but also on how easy it is to move from this location to locations that are valuable for retirement. Similarly, the number of the old in each location depends both on the attractiveness of this location for the old, determined by its amenities for the old and rents, and on the proximity of this location to the locations valuable for the young.

As argued in the literature, the presence of spillovers might lead to multiplicity of equilibria. Unfortunately, the existence and uniqueness techniques developed in Allen et al. (2015) do not directly apply to the model developed in the paper. Therefore, I proceed with numerical analysis and show that for the calibrated parameter values the solution exists and is stable.

4.5.1 Adjusted fundamentals

Combining equations (29) and (31), I derive the following equations for the endogenous amenities:

$$a_i^y = \frac{A_i^y}{\sigma F_{1,i}^{\alpha_1^y} F_{2,i}^{\alpha_2^y}} \left(\frac{E_{1,i}^{\alpha_1^y} E_{2,i}^{\alpha_2^y}}{R_i}\right)^{\frac{\delta_b}{\sigma-1}} \qquad a_i^o = \frac{A_i^o}{\sigma F_{1,i}^{\alpha_1^o} F_{2,i}^{\alpha_2^o}} \left(\frac{E_{1,i}^{\alpha_1^o} E_{2,i}^{\alpha_2^o}}{R_i}\right)^{\frac{\delta_b}{\sigma-1}} \tag{15}$$

where $E_{s,i}$ denotes total expenditures on category s in location i and is determined by $E_{s,i} = t^y(L_i^y \delta_b \alpha_s^y \frac{w_i}{1+\rho} + \rho \sum_j L_{ji} \delta_b \alpha_s^o \frac{w_j}{1+\rho})$.

Define by adjusted amenities the exogenous components of the equations (15):

$$\tilde{A}_i^y = \frac{A_i^y}{\sigma F_{1,i}^{\alpha_1^y} F_{2,i}^{\alpha_1^y}} \qquad \tilde{A}_i^o = \frac{A_i^o}{\sigma F_{1,i}^{\alpha_1^o} F_{2,i}^{\alpha_1^o}} \tag{16}$$

Adjusted exogenous amenities are determined both by location's exogenous characteristics such as climate and nature (A_i), and by location specific fixed costs of entry in non-tradable industries

$(F_{s,i})$. Therefore, spatial variation in the relative exogenous amenities for young and old can be rationalised not only by the differences in the valuation of climate and nature, but also by the differences in the preferences for local services. For example, if due to some exogenous reasons local entry costs into the health care industry are low, this would be reflected as a relatively higher exogenous amenity for the old.

Given the exogenous fundamentals $\{\tilde{A}_i^y, \tilde{A}_i^o, Z_i, \overline{H}_i, \mu_{ij}\}$, I can solve for the equilibrium set of variables $\{L_{ij}, w_i, R_i, a_i^y, a_i^o\}$ using equations (3), (12), (9), (11), (15), and (16). Therefore, I don't need to estimate the fixed cost of entry $F_{s,i}$ to compute the main equilibrium outcomes of the model, and also to compute the counterfactual changes in the number of firms.

4.6 Efficient Allocation and Policy

As argued in Fajgelbaum & Gaubert (2020) and Rossi-Hansberg et al. (2019), differences in productivity and amenity spillovers between demographic groups can lead to their inefficient sorting across space. I explore this idea in my setting, where differences in spillovers arise due to the individual's life cycle, and where the adjustment of amenities is modelled through the entry of variety producers. [16]

Consider a planning problem where a planner can redistribute the final good across consumers and construction producers, and reallocate housing units between local residents and local non-tradable sectors to maximise the expected lifetime utility of the individuals. The planner can't directly choose location of consumers, so the planner is bounded by the spatial mobility constraint (9). The planner also can't condition on the realization of the individual's preference shocks for location pairs. The planner problem is defined formally in Appendix D.3.

Proposition 1. *In an efficient allocation the following should hold:*

$$\underbrace{x_{ij}}_{consumption\ cost} + \underbrace{\lambda}_{opport} = \underbrace{z_i^* + \gamma^P z_i^*}_{marginal\ product\ in\ i} - \frac{x_{ij}}{\theta(1+\rho)} \tag{17}$$

$$\underbrace{R_i^* N_{si} F_{si}}_{housing\ cost} = \underbrace{\frac{\sigma}{\sigma-1}\frac{1}{\sigma}\frac{\delta_b}{1+\rho}\left(\alpha_s^y \sum_j x_{ij} L_{ij} + \rho\alpha_s^o \sum_j x_{ji} L_{ji}\right)}_{utility\ gain} \tag{18}$$

for $s = 1,2$, where $x_{ij} = c_{ij}^y + \rho c_{ij}^o + R_i^*(h_{ij}^y + \sum_{s=1,2} N_{si} k q_{sij}^y) + \rho R_j^*(h_{ij}^o + \sum_{s=1,2} N_{sj} k q_{sij}^o)$ is the cost along the path, R_i^* is the multiplier under the housing constraint of the city i and λ is the multiplier under the labor allocation constraint

[16] Rossi-Hansberg et al. (2019) consider only differences in productivity spillovers in the tradable sectors. Fajgelbaum & Gaubert (2020) consider both differences in amenity spillovers and productivity spillovers, but both are modelled in a reduced form way.

Efficient allocation can be implemented by the spatial subsidy for workers $s_j^w = \frac{\gamma^p - \frac{(1+\rho)}{\theta}}{1+\frac{1+\rho}{\theta}}$, *a lump-sum transfer* $T = -\lambda^* \frac{\theta}{1+\rho+\theta} - \Pi^*$, *where* Π^* *- is the per capita profit in the rental market in the optimal point, and a subsidy for the consumption of non-tradeables* $s^q = \frac{1}{\sigma}$.

Condition (17) shows that social planner equalises the marginal cost and benefit of allocating an individual to location path $\{i, j\}$. The marginal cost is the consumption levels along that path, the tightening of the housing constraints and the opportunity cost of allocating the individual on a different path. The marginal benefit is the marginal productivity of the individual when she is working in location i, including the productivity spillover γ^P. In equilibrium, the individual's expenditures x_{ij} are equal to location i wage: $w_i = z_i$. So the planner's optimality condition with respect to L_{ij} differs from the equilibrium one due to the presence of productivity spillovers γ^P, and also due to the dispersion in the location-specific preference shocks θ. The intuition for the latter is that in decentralised equilibrium workers face uninsurable preference shocks: some individuals receive high preference draws for working in less productive cities which makes them consume less throughout their life time. The planner cannot condition on the realisation of preference shocks, but by redistributing the final good towards less productive cities the planner can increase the expected utility of individuals.

Condition (18) shows that the planner chooses the optimal mass of firms in two sectors by equalising the housing costs associated with entry with the utility gains from a higher variety of services. As compared to the equilibrium outcome, the planner internalises the love of variety gains and wants to redistribute the housing away from residential use and towards services in the same proportion across locations equal to $\frac{\sigma}{\sigma-1}$. Therefore, the presence of endogenous amenities does not contribute to an inefficient allocation of young and old across cities. This result differs from the findings in Fajgelbaum & Gaubert (2020), where differences in reduced-form amenity spillovers require policies targeted at sorting of heterogeneous agents.

Overall, depending on the trade-off between productivity spillovers and distributional inefficiency due to the preference shocks, the planner either wants to increase or decrease the concentration of young individuals in more productive cities. As location choices of the young are affected by location choices of the old through housing prices, one way to affect the allocation of workers is through the reallocation of retirees. In the quantitative section, I further explore whether policies that incentivize the reallocation of retirees to less productive cities are welfare improving.

5 Model calibration

This section discusses the quantification of the model parameters. Section 5.1 describes the data and the mapping of the endogenous variables of the model to their empirical counterparts. Section 5.2 describes calibration of the structural parameters. Section 5.3 discusses the tree-step estimation procedure to recover the exogenous geography parameters.

5.1 Data

The model is calibrated to the mean outcomes across 1980, 1990, 2000 Censuses and 2012-2016 ACS. As in the empirical section, I define locations by 741 CZs in the U.S. I approximate L_i^o by the number of CZ residents older than 60.[17] Similarly, I approximate L_i^y by the number of CZ residents between ages 25 and 60. I use average composition adjusted wages of local workforce as my measure of w_i, and I use average quality adjusted rents paid by local residents as my measure of R_i. Therefore, for 741 commuting zones in the U.S. I observe the vector of wages, rents and location choices $\{w_i, R_i, L_i^y, L_i^o\}$. More details on the construction of the variables are in the Appendix A.

To estimate the mobility costs μ_{ij}, I additionally need data on location choices for each location path L_{ij}. In 2000 Census sample, I observe individual's age, current location and location five years ago. As the retirement decision is not directly observed, I assume that the reallocation for retirement reasons occurs in a 20 year window around the average retirement age of 64. So I approximate the location choices L_{ij} with $\hat{L}_{ij} = \sum_{a \in 57}^{a=77} L_{ij}^5(a)$, where $L_{ij}^5(a)$ is the number of individuals of age a who were in location i five years ago and currently are in the location j. While $\hat{L}_{ij}$ represents only a subsample of all moves for retirement due to the limited 5 year window, I show in Appendix E.3 that if individuals move for retirement with some constant probability for each year within the reallocation window, then this data limitation does not affect the estimation of mobility costs.

Finally, I use the consumer expenditure data (CEX) provided by NBER CEX extracts to estimate the differences in the expenditure shares on local services by age group. More details on the CEX dataset and the construction of the consumption groups are in Appendix A.3.

[17] The average retirement age in the US is 64, but as shown in section E.3, the highest change in the direction of mobility happens in the age group 56-65. So individuals might decide on changing location for retirement reasons earlier than actual retirement.

TABLE 2
CALIBRATED PARAMETERS

	Value	Description	Strategy	Source
γ^P	0.06	productivity spillover	Literature	Rosenthal & Strange (2004)
θ	3.35	preference shocks shape	Calibrated	Hornbeck & Moretti (2018)
δ_h	0.36	share spent on housing	Calibrated	CEX data
δ_b	0.224	share spent on local goods	Calibrated	based on Moretti (2013)
σ	5	Substitution elasticity for amenities	Literature	Hoelzlein (2019)
$1/\epsilon_h$	1.75	elasticity of housing supply	Literature	Saiz (2010)
α_1^y	0.711	share spent on local good 1 by young	Calibrated	CEX data
α_1^o	0.464	share spent on local good 1 by old	Calibrated	CEX data
ρ	0.5	length of old period	Calibrated	Old to young ratio

5.2 Calibrated parameters

I use the CEX data for the period from 1980 to 2000 to compute the average share of housing in consumers expenditures $\delta_h = 0.36$. [18] I calibrate the expenditure share on local non-housing goods δ_b to match the estimate from Moretti (2013) that 35 percent of non-housing costs varies systematically with housing prices.[19]

Parameters α_1^y and α_1^o govern the degree of heterogeneity in the expenditure shares on local goods between young and old age groups. I calibrate these parameters following the analysis in section 3.4. In particular, I assume that expenditure patterns on local goods can be approximated by 8 consumption categories corresponding to local services in the CEX data (see Appendix A.3 for the details). Then I compute predicted expenditure shares on these categories for the younger and the older age groups. Table 3 presents the results and describes the estimation procedure. In line with the results in the section 3.4, there exists significant heterogeneity between younger and older demographic groups in their expenditure patterns: e.g. the older group spends much more on healthcare services, while the younger group spends more on clothing, restaurants and childcare. I further aggregate the expenditure categories into two broader groups. Group 1 includes categories where the younger group spends a higher share of their income than the older group: restaurants, bars, clothing, childcare services, and recreation. Other categories are included in group 2: healthcare, nursing homes and beauty services. The resulting parameters $\alpha_1^y = 0.711$ and $\alpha_1^o = 0.464$ are the expenditure shares on group 1 of the young and the old age groups correspondingly.

Parameter σ governs the elasticity of demand for consumption amenities. Dolfen et al. (2019)

[18] Total expenditures exclude savings and contributions to retirement plans. Housing costs include both housing rent (imputed rent for homeowners) and utilities.

[19] The paper estimates the local component of non-housing costs by regressing changes in consumer price indices for individual cities (reported by BLS) on housing price changes within cities.

TABLE 3
PREDICTED EXPENDITURE SHARES ON LOCAL SERVICES

	(1) Restaurants	(2) Bars	(3) Clothing	(4) Childcare	(5) Recreation	(6) Beauty	(7) Health	(8) NursingH
$25 \leq age < 60$	0.223	0.025	0.262	0.035	0.165	0.05	0.24	0
$age \geq 60$	0.17	0.014	0.155	0.004	0.121	0.056	0.476	0.004

Notes: Table shows the predicted expenditure shares for younger and older age groups. First, I compute the total expenditures on 8 categories and the shares corresponding to each separate group. Then, I regress the resulting shares on age dummy (≥ 60), time fixed effects and a set of controls for sex, race, education and marital status. Finally, I compute the predicted values for the two age group dummy variables, keeping all the controls at their average values.

estimate elasticity of demand $\sigma = 6.1$ for the offline retail stores in the U.S., Couture et al. (2019) estimate $\sigma = 6.5$ for the U.S. restaurants, gyms, movie theaters, and outdoor amenities, Su (2019) reports values between 3.69 and 16 for disaggregated local sectors, and estimates $\sigma \in [4.26, 7.11]$ for medical facilities. As I use aggregated sectors, I follow Hoelzlein (2019) and choose $\sigma = 5$ which is towards the lower end of the empirical estimates.

In the model, parameter θ governs the elasticity of the individual's location choices with respect to location fundamentals. I calibrate θ so that the average employment response to exogenous local TFP shocks matches the estimates from Hornbeck & Moretti (2018).[20]

Finally, I set $\rho = 0.5$ to match the aggregate ratio of old-to-young in the data. Parameters L and t_y are normalised, as they only affect the scaling of the model variables.

Parameters from the Literature I set productivity spillover $\gamma^P = 0.06$ based on the empirical estimates of productivity spillovers summarized in Rosenthal & Strange (2004).[21] Saiz (2010) estimates developed land supply elasticities for 95 metropolitan statistical areas in the United States. I set $1/\epsilon_H$ to the population-weighted average of these elasticities equal to 1.75.

5.3 Mapping of the model to data

I now describe the three step estimation procedure to recover the exogenous geography parameters $\{\mu_{ij}, \tilde{A}_i^y, \tilde{A}_i^o, Z_i, \overline{H}_i\}$ for the U.S. CZs. First, I estimate the gravity for retirement flows to evaluate how mobility costs vary with the distance between locations. Second, I calibrate composite amenities $\{a_i^y, a_i^o\}$ and fixed component of mobility costs, so that the model exactly matches data $\{L_i^y, L_i^o, w_i, R_i\}$ and the share of the old who relocate for

[20] Hornbeck & Moretti (2018) estimate that a 1% increase in city manufacturing TFP during the period from 1980 to 1990 is associated with an average local employment increase by 4.16%.

[21] the empirical estimates are found to be in the rage of $3\% - 8\%$.

retirement. Finally, I decompose the model implied composite amenities, observed wages and housing prices into exogenous and endogenous components using the calibrated structural parameters described in section 5.2.

5.3.1 Gravity for retirement flows

I assume that mobility costs are a constant elasticity function of the physical distance between locations and a stochastic error for all pairs of CZs with positive commuting flows:

$$\ln \mu_{ij} = \begin{cases} \ln \kappa + \xi \ln dist_{ij} + \epsilon_{ij} & \text{if} \quad i \neq j \\ 0 \quad \text{if} \quad i = j \end{cases} \tag{19}$$

where κ is the fixed cost of reallocation, ξ is the elasticity of the mobility cost with respect to the distance, and ϵ_{ij} is an unobserved idiosyncratic component of commuting costs specific to a CZ pair. I won't be able to separately identify mobility costs from the amenities, instead I estimate the elasticity of mobility costs with respect to the distance between locations, and then calibrate the fixed component of mobility costs together with location-specific amenities for the young and the old.

In my baseline specification, I follow Redding & Sturm (2020) and assume prohibitive mobility costs ($\mu_{ij} \to \infty$) for the pairs of commuting zones with zero retirement flows. So I restrict my sample to CZs with positive retirement flows. Combining the specification for mobility costs (19) and the gravity equation for migration flows (9), I derive the gravity equation for the log retirement flows:

$$\ln \left(\frac{L_{ij}}{L} \right) = -\theta \rho \xi \ln dist_{ij} + \psi_i + \pi_j + e_{ij}$$

where the location of origin fixed effect includes local wage, amenities for young and rents, and expected utility and fixed component of the mobility cost $\psi_i = (1+\rho)\theta \ln w_i + \theta \ln a_i^y - (1-\delta_c)\theta \ln R_i - \theta \ln(U) - \theta \rho \ln \kappa$, the location destination fixed effect includes amenities for old and rents $\pi_j = \rho\theta \ln a_j^o - \rho\theta(1-\delta_c) \ln R_j$, and the error term represents idiosyncratic error in mobility costs between commuting zones $e_{ij} = -\theta \epsilon_{ij}$

In my baseline specification, I estimate the composite parameter $-\theta\xi\rho$ using OLS with origin and destination fixed effects. As described in data section 5.1, I approximate flows L_{ij} with $\hat{L_{ij}} = \sum_{a \in 57}^{a=77} L_{ij}^5(a)$, where $L_{ij}^5(a)$ is the number of individuals of age a who were in the location i five years ago and currently are in the location j from the US Census data for the year 2000. I observe the non zero migration flows for 101445 pairs of commuting zones.[22].

[22]In the Census data I observe mobility at the PUMA level. I then redistribute the migration flows across PUMAs to migration flows across CZs using the probabilistic matching based on the Autor & Dorn (2013)

TABLE 4

Gravity equation for retirement flows

	(1)	(2)
Dependent Variable	$\ln\left(\frac{L_{ij}}{L}\right)$	$\ln\left(\frac{L_{ij}}{L}\right)$
$\ln dist_{ij}(-\theta\rho\xi)$	-0.92	-1.42
	(0.005)	(0.014)
Estimation	OLS	Poisson PML
Observations	101445	548340
R^2	0.72	-
Origin and destination fixed effects	✓	✓

Notes: The dependent variable $\frac{L_{ij}}{L}$ is the share of individuals of age 57-77 who lives in the CZ i 5 years ago and live in the CZ j in 2000. The variable $dist_{ij}$ denotes the distance in miles between commuting zones i and j. Poisson PML is Poisson pseudo maximum likelihood estimator. Standard errors are clustered at the commuting zone level.

Estimation results for the baseline specification are presented in table 4 column (1). The elasticity of retirees migration flows with respect to the distance is equal to -0.92 and is significant at the 1% level. From the regression R^2, the gravity equation explains around 71% percent of the variation in bilateral mobility patterns, meaning that gravity framework works well for explaining the migration patterns of retirees. In column (2), I relax the assumption of infinite mobility costs for the commuting zone pairs with zero mobility flows, and estimate the gravity equation using the Poisson pseudo maximum likelihood estimator of Santos & Tenreyro (2006) on a full sample of CZ pairs. I find a similar pattern of results with a higher elasticity of -1.42.

The estimation above provide the first component of the mobility costs ${dist_{ij}}^{-\theta\rho\xi}$. In the next section, I discuss how to calibrate the fixed component of the mobility costs κ to match the share of the old that reallocates for retirement.

5.3.2 Recovering the composite amenities a_i^y and a_i^o and the fixed mobility cost κ

Given the estimated mobility costs μ_{ij}, I prove that I can uniquely recover composite location amenities a_i^y and a_i^o by inverting the equilibrium conditions for the distribution of the young and the old across locations:

Proposition 2. *Given observed data on $\{L_i^y, L_i^o, w_i, R_i\}$ and estimated mobility costs μ_{ij}, there exist unique (up to scale) values of $\{a_i^y, a_i^o\}$ that satisfy model equilibrium conditions (20) and*

crosswalk. More details are in AppendixE.2

(21):

$$\frac{L_i^y}{L} = \left(\frac{w_i^{(1+\rho)} a_i^y}{R_i^{1-\delta_c} U}\right)^{\theta} \sum_j \left(\frac{(a_j^o)^{\rho}}{R_j^{\rho(1-\delta_c)} \mu_{ij}}\right)^{\theta} \tag{20}$$

$$\frac{L_i^o}{L} = \rho \left(\frac{a_i^o}{R_i^{(1-\delta_c)}}\right)^{\theta\rho} \sum_j \left(\frac{w_j^{(1+\rho)} a_j^y}{R_j^{1-\delta_c} \mu_{ji} U}\right)^{\theta} \tag{21}$$

Proof. See Appendix D.4 □

The model predicts high amenities for the young in places that have high young population despite high housing prices, low wages and a bad access to amenable places for retirement. Similarly, the model predicts high amenities for the old in places that have high old population despite a bad access to the locations valuable for young and high housing prices.

Using the results from the proposition 2, for each value of the fixed cost κ I can uniquely recover the composite amenities a_i^y and a_i^o. Using the solution for the location choice problem 9, I compute the model-implied aggregate share of individuals who moves for retirement. Parameter κ is calibrated so that the model-implied reallocation share matches the corresponding share in the data. Using Census 2000 sample, I compute that 7.5% of individuals in the age group 60-65 moved between CZs within a 5 year period. This reallocation share misses all the moves that individuals make after the age of 65. Finkelstein et al. (2019) document that 10% of Medicare enrollees between ages 65 and 99 moved between CZs at least once over the 15 year period.[23] Assuming that 20% of individuals move more than once[24], this implies that roughly 14% of individuals in the older age group move within 20 year window. Therefore, I use this reallocation share as my target in the calibration.

I find that there exists significant dispersion in the relative amenities for old a_o/a_y across the U.S. CZs. Tables 5 and 6 list CZs with lowest and highest relative amenities for the old:

[23] Finkelstein et al. (2019) use administrative data on Medicare enrollees for a 100% panel of Medicare beneficiaries (69 milllion enrollees) for the period 1999-2014. Over this period of time 7 million enrollees changed their CZ

[24] Finkelstein et al. (2019) report that 2.4 million out of 7 million movers move more than twice

CZ Name	State	a_o/a_y
Anchorage	Alaska	0.07
Fairbanks	Alaska	0.1
Los Angeles	California	0.14
Bethel	Alaska	0.15
Houston	Texas	0.16

Table 5: CZ with lowest a_o/a_y

CZ Name	State	a_o/a_y
Port St. Lucie	Florida	16.8
Sarasota	Florida	15
Tampa	Florida	13.3
Cape Coral	Florida	11.3
Miami	Florida	9.8

Table 6: CZ with highest a_o/a_y

The CZ with the worst relative amenities for old is Anchorage in Alaska, followed by Fairbanks in Alaska. These cities might be less attractive for older residents due to severe climate conditions. Places with the best amenities for the old are Port St. Lucie, FL, followed by Sarasota, FL. These cities not only have much nicer natural amenities and better climate, but also lots of recreational activities for older residents such as golf and country clubs and swimming pools.

To study whether the model implied amenities are systematically related to the observable measures of amenities, I regress the ratio a_o/a_y on a set of climate and geography variables and on densities of different local establishments. Table J3 shows results. As predicted by the model, amenities are related to the composition of local establishments: a_o/a_y is significantly higher in locations with higher densities of golf and country clubs, health and nursing care services and personal care services, but lower in locations with higher density of restaurants, movie theaters and childcare services. Moreover, relative amenities for old are significantly higher in locations with higher percentage of water area and with warmer winter temperatures.

Finally, in Appendix F I show how I can recover the exogenous geography parameters $\{A_i^y, A_i^o, \overline{H}_i, Z_i\}$ given the recovered endogenous amenities, data, and the structure of the model.

6 Quantitative Analysis of the Model

This section studies the importance of the life cycle for the spatial equilibrium and the key sources of the observed patterns of sorting by age and location.

TABLE 7

GINI COEFFICIENTS FOR THE MAIN LOCAL VARIABLES

	Population	Rents	Amenities young
Full model /Data	1	1	1
No mobility for retirement	1.01	1.13	1.01
No life cycle	1.20	2.09	1.04

Notes: The table shows the Gini coefficients for the local variables in the data (full model), in the model with no life cycle ($T^y = T$ and $t^o = 0$) and in the model with no mobility for retirement (I recompute local population, rents and amenities assuming that $L_i^o = \rho L_i^y$ and keeping L_y fixed at its observed level). Coeffients are normalised by the full model levels. The sample includes 741 U.S. CZs

6.1 Life cycle and spatial equilibrium

First, I study how the life cycle affects the distribution of workers across locations. Consider the equation for the distribution of young individuals (13):

$$\frac{L_i^y}{L} = t_y \underbrace{\left(\frac{w_i^{(1+\rho)} a_i^y}{R_i^{1-\delta_c} U}\right)^\theta}_{\text{Value for young}} \underbrace{\sum_j \left(\frac{(a_j^o)^\rho}{R_j^{\rho(1-\delta_c)} \mu_{ij}^\rho}\right)^\theta}_{\text{Retirement option}}$$

I compute that the R^2 from the regression of the observed L_i^y on the estimated value for the young is equal to 0.7. This implies that the retirement option contributes to at least 30 percent of the variation in the CZ working age population size. This result is driven by the fact that locations that location's value for the young is different from its value for the old, and by the fact that the young anticipate the costly relocation for retirement.

To further study the importance of the life cycle for the distribution of economic activity across cities, I compare the Gini coefficients for the main endogenous local variables in the full model (as in the data) with the restricted version of the model where all individuals are young workers who make one static location choice ($t^y = T$ and $t^o = 0$). To isolate the role of reallocating retirees, I additionally compute the Gini coefficients for the case when older individuals don't move for retirement. Table 7 presents the results. Overall, the life cycle contributes to a more equal distribution of location population sizes, rents and amenities. First, some cities have an exogenous advantage in their amenities for old, and this allows them to attract both young and old residents, which contributes to a more equal distribution of economic activity across space. Moreover, older individuals move out of more expensive cities towards smaller and cheaper places. Additionally, they bring their higher earnings with them, which further contributes to a more equal distribution of prices and amenities across locations.

In the model, young individuals are fully mobile, so their expected life time utility is equalised

TABLE 8
DECOMPOSITION GEOGRAPHIC SORTING BY AGE

Full model	Restricted model $A_i^o = A_i^y$	Restricted model $a_i^o = a_i^y$
1	0.242	0.998

Notes:

across space. However, this is no longer true for the utility of older residents, who face mobility costs of reallocation across cities. I find that the per period utility of older residents varies significantly across the U.S. CZs: moving from an area at the 25th percentile of the value for the old to an area at the 75th percentile is equivalent to increasing the individual's income by a factor of 2.07.[25] These differences are driven by the fact that locations that are most valuable for the work period are often different from locations that are best for the retirement, so a worse utility for the older residents is offset by a higher utility during the work period and vice versa.

6.2 Decomposing sorting patterns by age and location

What are the drivers of sorting of younger and older residents across locations? Equilibrium conditions (13) and (14) imply that if older individuals can move across locations (mobility costs are finite), sorting of the young and the old across locations is governed by the dispersion in the differences of location values for young and old residents. This dispersion arises from the fact that older individuals no longer benefit from locations' wages and from the fact that they value location's exogenous and endogenous amenities differently. To separate these channels, I simulate the restricted version of the model where the only life cycle change is retirement.[26] Then I consider a version of the model where the old differ in their preferences for locations' exogenous amenities but not for locations' services. In each simulation, I compute the dissimilarity index for the distribution of the young and the old across CZs. The results are in table 8. Differences in CZs exogenous amenities is the key driver of geographic services by age followed by the differences in preferences for wages. Endogenous adjustment of amenities to differences in preferences for services has limited role in explaining sorting by age.

Next, I use the model to study drivers of the observed relationships between CZs old-to-young ratios, wages and establishments. First, consider the correlation between the composition of location establishments and location's demographics documented in section 3.3. To directly relate the empirical evidence to the model outcomes, I aggregate establishments into two broader service categories as in calibration section 5.2, and then I compute the elasticity of the composition establishments with respect to CZ demographic

[25]Per period utility of older residents in a location i who lived in j when young is $u^o_{ji} = \frac{w_j a^o_i}{R_i^{1-\delta_c}}$

[26]I set $A^o_i = A^y_i$ and $\alpha^o_s = \alpha^y_s$ for $s = 1, 2$.

TABLE 9

Composition of local establishments: model vs data

Dependent Variable	(1) N_2/N_1 data	(2) N_2/N_1 model
L_i^o/L_i^y	0.46	0.26
	(0.07)	(0.002)
Controls	✓	✓

Notes: p-value $^{***}p \leq 0.01$, $^{**}p \leq 0.05$, $^{*}p \leq 0.1$.Robust standard errors are given in parenthesis. All variables are in logs. N_2 includes establishments in hospital, nursing care homes and beaty services. N_1 includes restaurants, clothing, and childcare services. Both regressions include CZ wages and rents as controls, and include observations for 722 U.S. continental CZs.

composition by running the following OLS regression:

$$\ln \frac{N_{2i}}{N_{1i}} = \beta_o + \beta_1 \ln \frac{L_i^o}{L_i^y} + X_i + \epsilon_i$$

where X_i include controls for log location's wages and rents. Then I run the same specification on the model-simulated data, where I assume that the fixed costs of entry $F_{s,i}$ are symmetric across locations. In this counterfactual, the differences in the exogenous entry conditions are eliminated, and the only reason why the number of firms correlates with demographics is the endogenous response of firms to local expenditures. Figure 9 presents the results and shows that the model implied elasticity $\beta_1 = 0.26$ is more than half of the empirical estimate $\beta_1 = 0.46$ in the data. Therefore, half of the observed relationship between CZ establishments and old-to-young ratios is attributed to the differences in exogenous entry conditions for local services and another half is attributed to the endogenous response of local services to CZ age composition.

Another empirical fact documented in the paper is that CZ wage is negatively correlated with CZ old-to-young ratio. To isolate the role of differences in working hours in generating this pattern, I consider a version of the model where the only change over the life cycle is retirement.[27] In the restricted model, the elasticity of the ratio of old to young w.r.t. CZ wages is-0.7 versus -0.9 in the data.[28] Therefore, most of the observed negative relationship between CZ wages and demographics is attributed to the decline in working hours after retirement, and the rest is attributed to relative amenities for the old being better in less productive locations.

[27] I assume that $\tilde{A}_i^o = \tilde{A}_i^y$ and that $\alpha^o = \alpha^y$

[28]I run the same regression as on figure 2 using the data simulated from the restricted version of the model

7 Counterfactual Exercises

This section describes several counterfactual exercises. Section 7.1 studies local and aggregate effects of place-based policies that target older individuals to relocate to less productive cities. Section 7.2 shows that geographic sorting by age has increased over 1960-2010 period and argues that declining mobility frictions for the old was the main factor behind these changes.

7.1 Place-based policies targeted at retirees

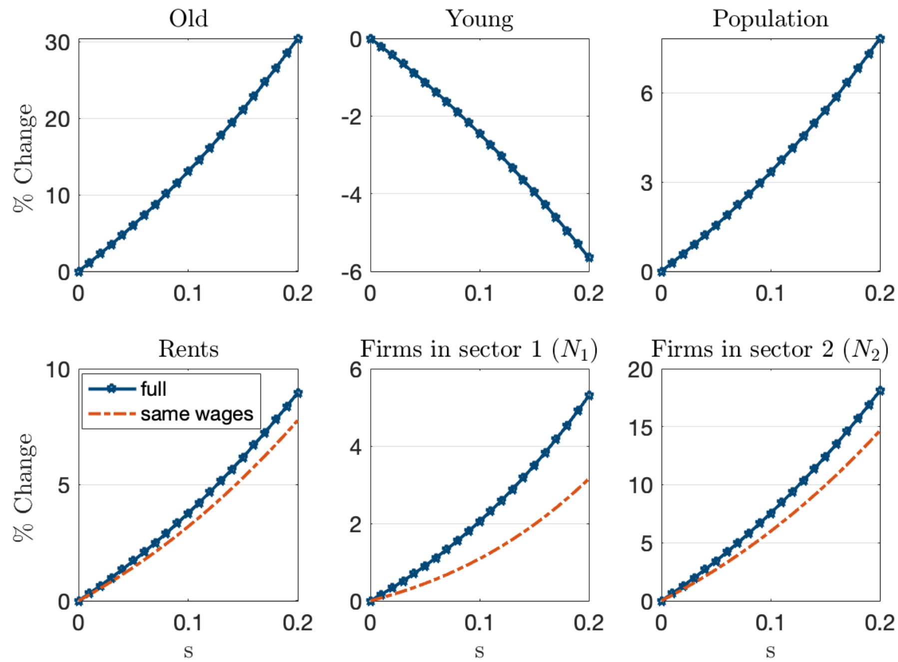
Should policy promote the development of smaller/less-productive cities by making them more attractive for retirees? Existing literature has largely focused on the effects of place-based policies that subsidize firms to locate in less developed cities or regions. These types of policies aim at reducing the spatial disparities and increasing the aggregate efficiency. However, as argued in Kline & Moretti (2014), positive local impacts of such policies are counterbalanced by losses in other locations and lead to the aggregate losses in the welfare and TFP. Gaubert (2018) further shows that such policies actually increase the spatial disparities in GDP per capita in the U.S. Related to the questions in this paper, some states indeed try to boost their less developed areas by attracting the retirees to reallocate there.[29] However, the efficiency of such policies depends on whether the optimal allocation features higher or lower sorting of the young and the old given the calibrated parameters.

In this section, I study quantitatively local and aggregate effects of the policies that incentivise the old to reallocate to less productive cities. I consider counterfactuals in which older individuals are given a subsidy s to move to the least productive cities in the U.S (cities in the 25th quantile by productivity). Figure 5 shows the local effects of the policy as a function of the subsidy level s. For example, a 10% subsidy in terms of the average wages in the economy which costs 0.14% of GDP increases the total population of treated cities by 3.4%, increases rents by 3.8% and increases the number of firms in sectors 1 and 2 by 2.09% and 7.58% respectively. The red dotted line on figure 5 shows what response of local variables would be lower if relocating retirees had the same wages as the workers in the treated city. The difference between the two lines shows that local policy effects are amplified by higher earnings of relocating retirees.

The counterfactual exercise allows me to compute the general equilibrium effect of the policy on the welfare and TFP. Appendix figure K9 shows that aggregate welfare is monotonically decreasing in the subsidy level s, while the reverse holds for the aggregate TFP.

[29] For example, Alabama provides state assistance to rural communities in retiree attraction, including planning and technical assistance coordinated marketing efforts, financial assistance, and amenity development. More examples are in Reeder (1998).

Figure 5: Local effects of the policy that subsidises the reallocation of old to less productive cities



Notes: Simulated local changes due to the introduction of a subsidy s measured as a percentage of individuals' average wages in the economy given to any old individual moving to a treated city. The policy is finances via a lump sum tax paid by all workers in the economy. Treated cities are the smallest CZs in U.S. that represent 1 percent of total population.

While reallocation of retirees improves productive efficiency, this is achieved by very costly mobility of the old, which does not offset the TFP gains. From the optimal policy perspective, it turns out that for the calibrated parameters, gains from productivity spillovers are offset by a worse distributional efficiency due to the preference shocks. So higher sorting by age and productivity is not welfare improving.

I further study how the policy affects the dispersion of spatial outcomes according to the Gini coefficient for the endogenous spatial variables. I find that the type of the place-based policy that I study contributes to a more equal distribution of amenities and rents across CZs, but a more unequal distribution of wages in the tradeable sector. However, as I don't model the employment in the local service sector, the analysis might miss the additional response from employment and wages there.

Before I considered the national policy which relocates older individuals to less productive cities. But can smaller cities grow by attracting retirees using locally financed policies? In Appendix H, I show that such policies are not successful at increasing the location's population size and its amenities. The reason is that subsidising retirees is costly for workers, as a result less workers and future retirees choose to live in the city. This force offsets the positive effect of the increase in local amenities from expenditures of incoming retirees. So the major effect of the policy is just the transfer from the local young to the local old.

7.2 Changes in geographic sorting by age, 1960-2010

Figure 6 shows that a pattern of sorting by age and location's productivity documented in recent decades was not present in earlier Census sample of 1960.[30] I further compute that the dissimilarity index for the younger and the older age groups has increased by 38% over this period from around 0.06 in 1960 to around 0.08 in 2010. [31] These changes might be related to an increase in the mobility around the retirement age: the share of individuals in the age group 57-77 who moved between states within a 5 year window has increased by around 20% over the 1960-2000 period.[32] By comparison, the share of individuals in the younger age group (25-45) who moved between states has remained nearly the same. Appendix G shows that the interstate mobility for the older age group has been increasing throughout 1960-2000 period, while the interstate mobility for the younger age group was rising from 1960 to 1980 and declining afterwards. Appendix G also discusses how the results on the increase in the interstate mobility for the older age group relate to the literature on the declining interstate mobility.

This section studies the drivers behind the changes in geographic sorting by age over time. In particular, I recalibrate my model to the MSA-level data for 1960 and 2010 years, treating these two time periods as different steady states of the model. I keep all the structural parameters as in the main quantification part, except for the fixed component of the mobility cost (κ) and the relative length of the old period (ρ). I recalibrate ρ to match the increase in the old-to-young ratio over the considered period of time. I recalibrate κ and the exogenous location-specific parameters for year 1960 so that to match the distribution of individuals and prices in 1960 and to match the fact that the share of individuals who reallocate between states for retirement in 1960 is by 20 percent lower than in 2010.[33] More details on the calibration procedure are in Appendix I.1. I find that the mobility costs in 1960 were by 28% lower than in 2010. One of the reasons behind this large model-implied change in mobility costs is that in 1960 population

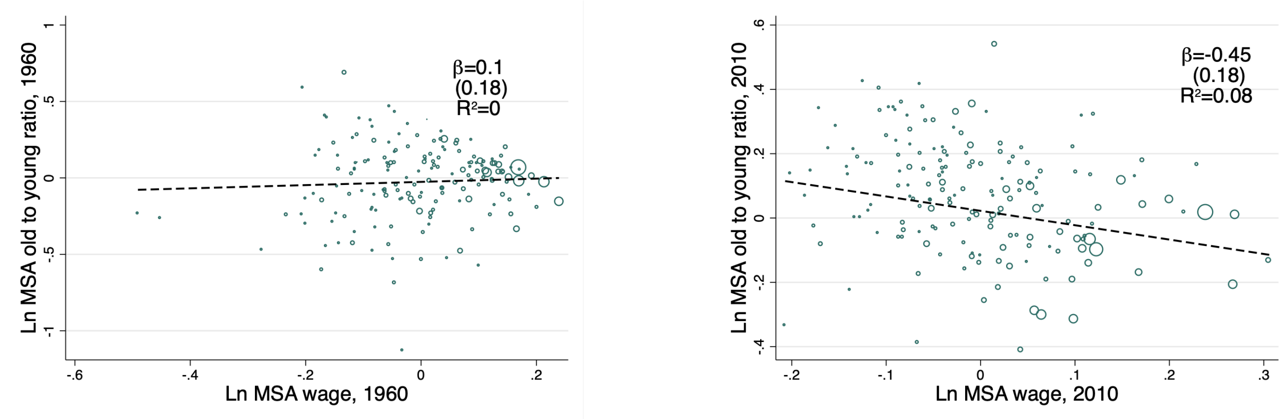
[30] As there is no mapping between geography units and CZs for historical samples for 1960 sample, I chose to make historical comparisons at the MSA level.

[31] The dissimilarity index is defined as $\frac{1}{2}\left|\frac{L_i^o}{\sum_i L_i^o} - \frac{L_i^y}{\sum_i L_i^y}\right|$ where L_i^o is the size of the older age group in the location i, and L_i^y is the size of the younger age group in i. I include 169 common MSAs and one location corresponding to all the areas outside of the MSAs in the computation

[32] I compute that 4.12% of individuals in the age group 55-77 moved between U.S. states in 1960 within a five year window, while the corresponding number for 2000 is 4.92%.

[33] I map each MSA to a single state and compute the share of the old who relocates between states

Figure 6: Sorting by age and productivity, MSA level



(a) 1960

(b) 2010

Notes: Regressions weighted by the MSA total population. Older age group includes individuals of age 60+ and younger age group includes individuals of age $25-60$. Wages are adjusted for the location's demographics. Regressions include 169 MSAs present in both years. Wages and old-to young ratios are in logs and normalised by the MSA mean values in the corresponding year.

was much more concentrated in the most productive U.S. MSAs, and most of these locations actually had old-to-young ratios above national average. So to rationalise low mobility flows for retirement in that period the model requires relatively large mobility costs.

Next, I decompose the drivers behind the changes in sorting by age and location's productivity from 1960 to 2010. Table 10 shows that if I take the model calibrated to 1960 and change mobility costs alone to their calibrated 2010 level, I can explain around 51% of the increase in the correlation between MSA old-to-young ratio and MSA wages. Appendix I.2 performs the full decomposition, and shows that changes in location-specific factors explain the remaining increase in sorting, while the ageing shift actually has an opposite effect on sorting by age and productivity. Finally, I compute that the decline in mobility costs between 1960 and 2010 generated a 3.7 percent increase in welfare, where the welfare is measured by the expected lifetime utility of the individuals. I further compute that this change in mobility costs has an equivalent effect on utility as an increase in individual's income by 2.6 percent. There are two main sources for these gains. First, individuals now can increasingly separate locations for work and retirement periods. So the young can increasingly live in locations that are more productive and more amenable for the young, and the old can increasingly live in locations that are cheaper and more amenable for the old. Second, the old who move now pay lower mobility costs. [34]

One concern might be that changes in the old-to-young ratio are not perfectly aligned with

[34] I find that welfare gains from the fact that movers in 1960 now pay lower mobility cost are only 0.02%

TABLE 10
CHANGES IN SORTING 1960-2010

	1960	Simulation	2010
$corr(W_i, \frac{L_i^o}{L_i^y})$	0.01	-0.13	-0.27

Notes: All variables are normalized by the year means. Column two computes the correlation of the simulated W_i and $\frac{L_i^o}{L_i^y}$ in a model that was calibrated to year 1960 at the MSA level and where mobility costs where set at 2010 level.

the changes in the relative length of the old period in the consumer's problem. First, individuals now work longer, so their perception of the retirement period could be different. Second, the model does not take into account the differences in the initial sizes of the cohorts that are old in 1960 and in 2010. In Appendix I, I conduct a series of robustness checks for different changes in the relative length of the old period ρ over 1960-2010 period. I find that main results are robust to ρ changes.

8 Conclusion

This paper argues on the importance of the life cycle for location choices and for the distribution of economic activity across space. By building and estimating the model of location choices for work and retirement periods, I quantify the differences in locations' values for their younger and older residents. I find that these differences arise from changes in working hours and preferences for local amenities with age, and that they imply large gains from life cycle mobility. I also estimate that mobility frictions preclude retirees from moving to better places and lead to a significant dispersion in location-specific utilities of older individuals.

I document that geographic sorting by age has increased dramatically over 1960-2010 period. Using the model structure, I argue that this increase in sorting is related to the decline in mobility frictions of relocating for retirement, which allowed individuals to increasingly separate locations for work and retirement periods. I further show that policies that target higher sorting of young and old through place-based transfers are not welfare improving. This suggests that it is important to further study the frictional components of the mobility costs that older individuals face and how and why they have changed over time.

The paper also shows that policies that aim to relocate older individuals to less productive cities can be considered for the purpose of reducing spatial disparities, and they differ from similar policies targeted at workers as they do not distort production efficiency ,and because relocating retirees bring their savings from more productive cities with them.

References

Aguiar, M. and E. Hurst (2013). Deconstructing life cycle expenditure. *Journal of Political Economy 121*(3), 437–492.

Ahfeldt, G., S. Redding, D. Sturm, and N. Wolf (2015). The economics of density: Evidence from the berlin wall. *Econometrica 83*(6), 2127–2189.

Allen, T. and C. Arkolakis (2014). Trade and topography of the spatial economy. *The Quarterly Journal of Economics 129*(3), 1085–1140.

Allen, T., C. Arkolakis, and X. Li (2015). On the existence and uniqueness of trade equilibria. *mimeo, Dartmouth and Yale Universities.*

Allen, T. and D. Donaldson (2018). The geography of path dependence. *Unpublished manuscript.*

Almagro, M. and T. Dominguez-líno (2019). Location sorting and endogenous amenities: Evidence from amsterdam.

Autor, D. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review 103*(5), 1553–97.

Bilal, A. and E. Rossi-Hansberg (2018). Location as an asset. *National Bureau of Economic Research No. w24867.*

Couture, V., C. Gaibert, J. Handbury, and E. Hurst (2019). Income growth and distributional effects of urban spatial sorting. *National Bureau of Economic Research* (No. w26142).

Desmet, K. and E. Rossi-Hansberg (2013). Urban accounting and welfare. *American Economic Review 103*(6), 2296–2327.

Desmet, K. and E. Rossi-Hansberg (2014). Spatial development. *American Economic Review 104*(4), 1211–43.

Diamond, R. (2016). The determinants and welfare implications of us workers diverging location choices by skill: 1980-2000. *American Economic Review 106*(3), 479–524.

Dolfen, P., L. Einav, P. J. Klenow, B. Klopack, J. D. Levin, and W. Best (2019). Assessing the gains from e-commerce. *National Bureau of Economic Research* (No w.25610).

Duncombe, W., M. Robbins, and D. A. Wolf (2001). Retire to where? a discrete choice model of residential location. *International Journal of Population Geography 7*(4), 281–293.

Eckert, F. (2019). Growing apart: Tradable services and the fragmentation of the us economy. *mimeograph, Yale University*.

Eckert, F., A. Gvirtz, J. Liang, and M. Peters (2020). A method to construct geographical crosswalks with an application to us counties since 1790. *NBER Working Paper No. 26770*.

Eckert, F. and M. Peters (2018). Spatial structural change. *Unpublished manuscript*.

Fajgelbaum, P. D. and C. Gaubert (2020). Optimal spatial policies, geography, and sorting. *The Quarterly Journal of Economics 135*(2), 959–1036.

Finkelstein, A., M. Gentzkow, and H. L. Williams (2019). Place based drivers of mortaility: Evidence from migration. *National Bureau of Economic Research* (No. w25975).

Gaubert, C. (2018). Firm sorting and agglomeration. *American Economic Review 108*(11), 3117–53.

Giannone, E., Q. Li, N. Paixao, and X. Pang (2019). Unpacking moving.

Glaeser, E. L. and J. D. Gottlieb (2008). The economics of place-making policies. *National Bureau of Economic Research* (No, w14373).

Guerrieri, V., D. Hartley, and E. Hurst (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics*, 45–60.

Heise, S. and T. Porzio (2022). Labor misallocation across firms and regions. *National Bureau of Economic Research* (No. w30298.).

Hoelzlein, M. (2019). Two-sided sorting and spatial inequality in cities.

Hornbeck, R. and E. Moretti (2018). Who benefits from productivity growth? direct and indirect effects of local tfp growth on wages, rents and inequality. *National Burreau of Economic Research* (No. w24661).

Kaplan, G. and Sam-Schulhofer-Wohl (2017). Understanding the long-run decline in interstate migration. *International Economic Review 58*(1).

Kline, P. and E. Moretti (2014). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authoruty. *The Quarterly Journal of Economics 129*(1), 275–331.

Koşar, G., T. Ransom, and W. V. der Klaauw (2021). Understanding migration aversion using elicited counterfactual choice probabilities. *Journal of Econometrics*.

Krugman, P. R. (1979). Increasing returns, monopolistic competition, and international trade. *Journal of International Economics 9*(4), 469–479.

Lu, J. (2020). Household residential location choice in retirement: The role of climate amenities. *Regional Science and Urban Economics 84*.

Molloy, R., C. L. Smith, and A. Wozniak (2011). Internal migration in the united states. *Journal of Economic Perspectives 25*(3), 173–196.

Monte, F., S. Redding, and E. Rossi-Hansberg (2018). Commuting, migration and local employment elasticities. *American Economic Review 108*(12), 3855–90.

Moretti, E. (2013). Real wage inequality. *American Economic Journal: Applied Economics 5*(1), 65–103.

Redding, H. S. S. J. and D. M. Sturm (2020). The making of the modern metropolis: evidence from london. *The Quarterly Journal of Economics*.

Redding, S. J. and D. M. Sturm (2008). The costs of remoteness: Evidence from german division and reunification. *American Economic Review 98*, 1766–1797.

Reeder, R. J. (1998). Retiree-attraction policies for rural development. *Agriculture Information Bulletin* (741).

Roback, J. (1982). Wages, rents, and the quality of life. *the Journal of Political Economy*, 1257–1278.

Rosen, S. (1979). Wages-based indexes of urban quality of life. *Current Issues in Urban Economics*.

Rosenthal, S. S. and W. C. Strange (2004). Evidence on the nature ans sources of agglomeration economies. *Handbook of regional and urban economies 4*(2119-2171).

Rossi-Hansberg, E., P.-D. Sarte, and F. Schwartzman (2019). Cognitive hubs and spatial redistribution. *National Bureau of Economic Research* (w26267).

Ruggles, S., K. Genadek, R. Goeken, J. Grover, and M. Sobek (2015). Integrated public use microdata series: Version 6.0 (machine readable database).

Saiz, A. (2010). The geographic determinants of housing supply. *Quarterly Journal of Economics 125*(3), 1253–96.

Santos, S. J. and S. Tenreyro (2006). The log of gravity. *The Review of Economics and Statistics 88*(4), 641–658.

Schneider, M. and B. Green (1992). A demographic and economic comparison of nonmetropolitan retirement and nonretirement counties in the us. *Journal of Applied Sociology*, 63–84.

Su, Y. (2019). Measuring the value of urban consumption amenities: A time use approach. *Available at SSRN 3454631*.

Su, Y. (2020). The rising value of time and the origin of urban gentrification.

Appendices

For Online Publication

A Data Appendix

Definitions

Commuting Zones. The unit of geography is a commuting zone (CZ), which is an aggregation of the U.S. counties that partitions the U.S. and is designed to proxy for the local labor markets. I use time consistent mapping between Census Pumas and 741 US commuting zones provided by Autor & Dorn (2013).

Age groups. I define by the younger age group individuals whose age is between 25 and 59 years, and by the older age group individuals older than or equal to 60. Using IPUMS data and age group definitions, I compute the young and the old population of each CZ in the U.S.

Composition adjusted wages. To proxy for the CZ productivity, I use IPUMS data on the average wages earned by CZ workers.[35] I adjust wages for the differences in worker's age, sex, race, nativity and highest completed education. More details are in Appendix A. The resulting composition adjusted wages control for the differences in observable characteristics of the CZ workforce, and, therefore, proxy for the location-specific component of individual's wages.

Quality adjusted rents. To proxy for the local prices, I use IPUMS data on the average rents paid by the households in a given location. I adjust rents for the observed quality characteristics of the rented unit, so that to compare prices of similar units of housing across locations. More details are in Appendix A.

Amenities. I follow Diamond (2016) and Almagro & Dominguez-líno (2019) in measuring the availability of local amenities by using the densities of establishments in different service industries. Details on the construction of the industry groups are in Appendix A. To proxy for CZ natural and climate amenities, I use data from USDA on January hours of sun, January average temperature, July average temperature, topography and percentage of water area. This data is available only for 722 continental CZs.

[35]By workers I define individuals of age 17-64, who work more than 44 weeks and more than 34 hours a week.

A.1 Composition adjusted wages

I compute individual's log hourly wages in a subsample of workers: individuals of age 17-64, who work more than 44 weeks and more than 34 hours a week. For each sample year I regress log hourly wages on the set of demographic controls: age, sex, race, nativity and highest completed education:

$$\ln w_{i,t} = \alpha + \beta X_{i,t} + \epsilon_{i,t}$$

Composition adjusted wages for residents in each location j can then be constructed as:

$$\ln w_{i,j,t}^{adj} = \ln w_{i,j,t} - \sum_k \beta_t^k \left(X_{i,t}^k - \overline{X}_t^k \right)$$

where $w_{i,j,t}$ is the individual i wage in a location j at the time t, and $\overline{X}_t^k$ is the average of the demographic control k in year t across all individuals. CZ j composition adjusted wage is computed as the average of the composition adjusted wages of its workers.

A.2 Quality adjusted rents

To compute a proxy for local prices, I consider rental payments reported by a subsample of households who rent housing. I use gross rent, which is monthly rental cost of the housing unit, including contract rent plus additional costs for utilities (water, electricity, gas) and fuels (oil, coal, kerosene, wood, etc.). In order to be able to compare rental prices across CZs, I adjust them for the differences in the quality of the rented housing. The observable characteristics of the housing units are the following:

- **Year build**. Six indicator variables for the year build (0-1 years, 2-5 years, 6-10 years, 11-20 years, 21-30 years, 31+)
- **Number of bedrooms**. Six indicator variables (0 bedrooms, 1, 2, 3, 4, 5+)
- **Kitchen facilities**. Indicator for the presence of the kitchen facilities
- **Plumbing facilities**. Indicator for the presence of the plumbing facilities

For each sample year for the subsample of renters I regress household's log rents on the set of housing controls defined above:

$$\ln R_{i,t} = \alpha + \beta X_{i,t} + \epsilon_{i,t}$$

Quality adjusted rents for residents of each location j can then be constructed as:

$$\ln R_{i,j,t}^{adj} = \ln R_{i,j,t} - \sum_k \beta_t^k \left(X_{i,t}^k - \overline{X}_t^k \right)$$

where $R_{i,j,t}$ is the individual i rent paid in a location j at the time t, and $\overline{X}_t^k$ is the average value of the quality control k in year t across all individuals. City j quality adjusted rents are computed as the average of the quality adjusted rents of its residents who rent housing.

A.3 CEX data

I use data extracts from the Consumer Expenditure Surveys quarterly interview survey compiled by Ed Harris and John Sabelhaus and provided by NBER. The dataset includes households interviewed between the first quarter of 1980 and the second quarter of 2003. The interviews are conducted during five consecutive quarters: first quarter collects demographic data for each household member, and subsequent four interviews collect data on household level expenditures. The NBER extract aggregates the four interviews into a single observation per household, summing all quarterly expenditures.

As expenditure level data is provided at the household level, I restrict the sample to the single households, married couples and married couples with children under the age of 21. For each household I define a household's head using the procedure developed in Aguiar & Hurst (2013): a head is defined as an individual who identifies herself or himself as a household head in the survey, and if there are multiple heads, the head is identified by sex, employment, age and marital status. To match expenditure data to demographics, I use the demographic data for the identified household's head.

Harris and Sabelhaus aggregate expenditures into 47 broad categories. I compute total expenditure using all available categories, and then consider a subsample of categories that is most closely related to locally provided services:

- Food outside of home: includes food and non-alchoholic beverages at restaurants, cafes and fast food places, food at school and boarding places.
- Nightclubs: includes alcoholic beverages at restaurants, cafes and bars.
- Clothing: includes expenditures on clothing and shoes, jewelry and watches and tailoring.
- Personal care services: includes expenditures on barbershops, beaty parlors and health clubs.

- Health care services: includes expenditures on drugs and hospitals both out of pocket and included in medical insurance.
- Nursing Homes: includes care in Convalescent or Nursing Home
- Recreation (other recreation category in NBER extract): includes expenditures on repair and rental of recreational equipment, expenditures on pets, expenditures on country clubs, health clubs, swimming pools and other clubs, admission fees for entertainment activities and sporting event.
- Childcare services: includes expenditures on day care centers, nursery schools, elementary and secondary education.

I further restrict my sabsample to households that report postive expenditures for the sum of 8 local service categories, and restrict household's head to be older than 25. This leaves a sample of 77844 households.

A.4 Amenities

I use establishments counts at the county level in the following NAICS codes as proxies for local amenities:

- Food Services and Drinking places (NAICS 722 except 7224)
- Drinking places (NAICS 7224)
- Clothing and Accessory stores (NAICS 448)
- Golf Courses and Country clubs (NAICS 71391)
- Other recreation (NAICS 7139 except 71391).
- Performing Arts, Spectator Sports, and Related Industries (NAICS 711)
- Motion picture theatres (NAIC 512131)

I use crosswalks from counties to pumas and from pumas to commuting zones provided by Autor & Dorn (2013). To compute establishment density I scale total number of establishments by commuting zone total population size to get establishment counts per 1000 residents.

Additionally, I use geographic amenities data constructed by the United States Department of Agriculture.[36] These measures are warm winter, winter sun, temperate summer, low summer

[36] https://www.ers.usda.gov/data-products/natural-amenities-scale

humidity, topographic variation, and water area. The data is available at the county level. I use crosswalks provided by Eckert et al. (2020) to map the data to Commuting zones. The data is available for 722 CZs out of 741.

B Robustness Checks for the Empirical Results

This section provides the robustness checks for the main empirical results in the paper.

B.1 Robustness Checks for the Results in Figure 2

First, I show that the pattern of sorting presented on Figure 2 is robust to considering retirees instead of the old. I define retirees as individuals of age 50 and older who is out of the labor force. Figure B1 shows the elasticity of retiree-to-old ratio with respect to CZ wage. The results are stronger than in case of the old-to-young ratio, which might be driven by workers retiring earlier in less productive CZs

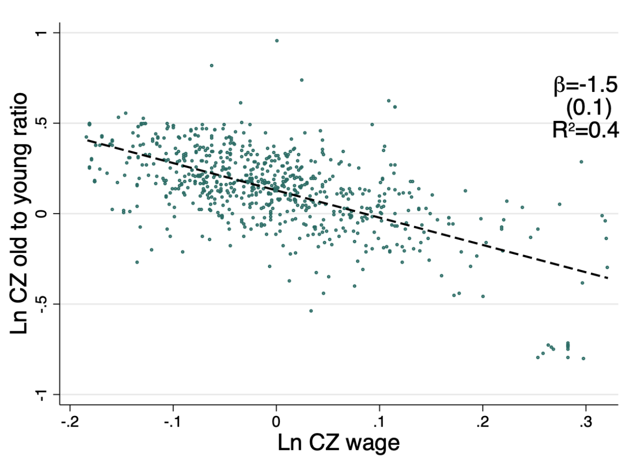
Figure B1: Sorting of retirees and workers across by productivity

Notes: The figure plots CZ's ratio of retirees to workers against CZ's average wages earned by workers. Retirees are defined as individuals of age 50 and older who is out of the labor force. Workers are individuals between ages 17 − 64, who work more than 44 weeks and more than 34 hours a week. Both variables are logs of the averages of demeaned values across 1980, 1990, 2000 and 2014 samples. CZs' wages are adjusted for the differences in CZs' demographics. β stands for the OLS regression coefficient and robust standard errors are in parenthesis. Regressions are weighted by CZ average population across same samples. The size of each CZ circle is proportionate to its average population.

Figure B2 shows that pattern of sorting by age and productivity is the same if we consider an age cut-off of 65+ instead of 60+. I also check if the results are robust to the weighting: the original regression uses observations weighted by the CZ population size. If observations are not weighted, the correlation between location productivity and share of old is even stronger, as shown in the figure B3.

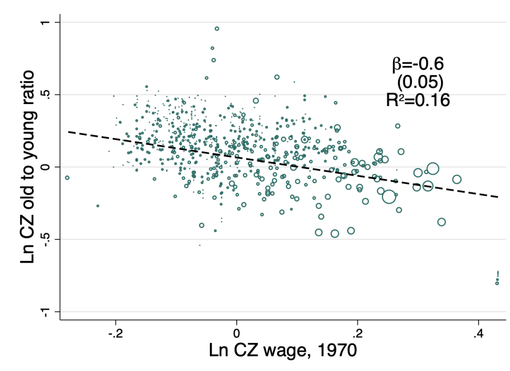
Finally, I provide an additional check that the original evidence on sorting is mostly driven by the fact that the young and the old value locations differently, rather than by the fact that they respond to location-specific shocks in a different way. In particular, I compute whether the lagged location productivity (year 1970) is correlated with location's current old-to-young (year 2014). Figure B4 shows that the resulting sorting patterns by age and lagged productivity is strong and significant, suggesting that time persistent differences in CZ productivities are driving the sorting pattern.

Figure B2: Sorting of the young and the old (65+) by productivity



Notes: The figure plots CZ's ratio of individuals older than 65 to individuals between ages 25 and 65 against CZ's average wages earned by workers. Both variables are logs of the averages of demeaned values across 1980, 1990, 2000 and 2014 samples. CZs' wages are adjusted for the differences in CZs' demographics. β stands for the OLS regression coefficient and robust standard errors are in parenthesis. Regressions are weighted by location's average population across same samples. The size of each CZ circle is proportionate to its average population.

Figure B3: Sorting of the young and the old by productivity (no weights)



The figure plots CZ's ratio of individuals older than 60 to individuals between ages 25 and 60 against CZ's average wages earned by workers. Both variables are logs of the averages of demeaned values across 1980, 1990, 2000 and 2014 samples. CZs' wages are adjusted for the differences in CZs' demographics. β stands for the OLS regression coefficient and robust standard errors are in parenthesis.

Figure B4: Sorting of young and old across the US CZs: lagged productivity

The figure plots CZ's ratio of individuals older than 60 to individuals between ages 25 and 60 in 2014 against CZ's average wages earned by workers in 1970. CZs' wages are adjusted for the differences in CZs' demographics. β stands for the OLS regression coefficient and robust standard errors are in parenthesis. Regressions are weighted by location's average population across same samples. The size of each CZ circle is proportionate to its average population.

C Changes in the individual-level mobility patterns with age

This section studies how the direction of individual's mobility between CZs changes with age. I use 2000 Census sample to identify the subsample of movers: individuals who changed their CZ within a 5 year window. For each mover, I compute the difference between the fundamentals of the destination and origin CZs. I study how the probability of moving up (to a commuting zone with at least 1% higher fundamental of interest) depends on the individual's age:

$$\text{Up}_{i,j} = \alpha + \beta X_{i,t} + \sum_{i=1}^{5} \gamma_i \text{age group}_{i,t} + \psi_j + \epsilon i$$

Where $X_{i,t}$ is a set of demographic controls for race, sex, nativity and education and ψ_j are location fixed effects. Age groups are $25-35, 36-45, 46-55, 56-65, 66+$.

Table C1 shows the results. As hypothesised, older individuals are significantly less likely to move to a more productive location as compared to young. The highest difference happens around the retirement age (age group 55-65). Older individuals are also less likely to move to a more populated and expensive location, except for the oldest group of 66+. This group does not move to cheaper places, but still moves to places with a significantly higher share of old, indicating the reallocation for amenity reasons.

TABLE C1
Direction of individual mobility by age

	Higher Wage	Higher Pop.	Higher Rent	Higher old share
$[36-45]$	-0.023***	-0.022***	-0.021***	0.02***
	(0.002)	(0.002)	(0.002)	(0.002)
$[46-55]$	-0.047***	-0.04***	-0.035***	0.043***
	(0.003)	(0.003)	(0.003)	(0.003)
$[56-65]$	-0.088***	-0.066***	-0.047***	0.099***
	(0.005)	(0.005)	(0.003)	(0.006)
66+	-0.043***	-0.027***	-0.004	0.084***
	(0.005)	(0.004)	(0.004)	(0.007)
Obs.	1,091,989	1,091,989	1,091,989	1,091,989
R-squared	0.307	0.335	0.299	0.299
CZ FE	YES	YES	YES	YES
Controls	YES	YES	YES	YES

Results relative to the youngest group 25-35. Standard errors in parentheses are clustered at the CZ level

D Proofs for the Theoretical Part

D.1 Formulation of the utility function

Individuals choose a pair of locations $\{i, j\}$, where i is where they live for the t^y young periods, and j is where they live for the t^o old periods. Conditional on the location path $\{i, j\}$, the individual's problem is:

$$\max_{c_t^y, c_t^o, h_t^y, h_t^o, q_{m,s,t}^y, q_{m,s,t}^o} \sum_{t=1}^{t=t^y} u(c_t^y, h_t^y, q_{m,s,t}^y) + \sum_{t=t^y+1}^{t=T} u(c_t^o, h_t^o, q_{m,s,t}^o)$$

$$\text{s.t.} \quad \sum_{t=1}^{t=t^y} \left(c_t^y + R_{i,t} h_t^y + \sum_{s=1}^{2} \sum_{m=1}^{N_{s,i,t}} p_{m,s,i,t} q_{m,s,t}^y \right) + \sum_{t=t^y+1}^{T} \left(c_t^o + R_i h_t^o + \sum_{s=1}^{2} \sum_{m=1}^{N_{s,i,t}} p_{m,s,i,t} q_{m,s,t}^o \right) = \sum_{t=1}^{t=t^y} w_i$$

where w_i is the wage that individual receives each period that she is young from supplying one unit of labor and where I assumed that the discount factor and the interest rate are equal to one.

As the model is stationary, in equilibrium all the prices will be constant through time. Then, in the optimal solution to this problem individuals will choose same consumption levels within young and old period: $c_t^y = c_y$, $h_t^y = h^y$, $c_t^o = c^o$, $h_t^o = h^o$, $q_{m,s,t}^y = q_{m,s}^y$ and $q_{m,s,t}^o = q_{m,s}^o$. Then the problem is equivalent to:

$$\max_{c^y, h^y, c^o, h^o} t_1 u(c^y, h^y, a_i) + t_2 u(c^o, h^o, a_j)$$

$$\text{s.t.} \quad t_1 \left(c^y + R_i h^y + \sum_{m=1}^{N_{s,i}} p_{m,s,i} q_{m,s}^y \right) + t_2 \left(c^o + R_i h^o + \sum_{m=1}^{N_{s,i}} p_{m,s,i} q_{m,s}^o \right) = t_1 w_i$$

D.2 Indirect Utility Derivation

Consider the consumption choice problem of an individual who chose the location path (i, j):

$$\max_{c^y, c^o, h^y, h^o, q_{ms}^y q_{ms}^o} \ln A_i^y (c^y)^{\delta_c} (h^y)^{\delta_h} (Q_i^y)^{\delta_b} + \rho \ln A_j^o (c^o)^{\delta_c} (h^o)^{\delta_h} (Q_j^o)^{\delta_b} - \rho \ln \mu_{ij} + \ln \epsilon_{ij\omega} \tag{22}$$

s.t.

$$c^y + R_i h^y + \sum_{s=1}^{2} \int_{m=1}^{N_{si}} p_{msi} q_{ms}^y dm + \rho (c_{ij}^o + R_j h_{ij}^o + \sum_{s=1}^{2} \int_{m=1}^{N_{sj}} p_{msj} q_{ms}^o dm) = w_i \tag{23}$$

where p_{msi} denotes the price of variety m of category s in location i, and $\rho = t^o / t^y$ is the relative length of the old period.

The resulting solution for the consumption of the final good and housing is:

$$c_{ij}^y = c_{ij}^o = \frac{\delta_c w_i}{1+\rho} \qquad h_{ij}^y = \frac{\delta_h w_i}{R_i (1+\rho)} \qquad h_{ij}^o = \frac{\delta_h w_i}{R_j (1+\rho)} \tag{24}$$

The resulting demand for each variety m in categories $s = 1, 2$ is:

$$q^y_{msij} = \alpha^y_s \delta_b \frac{w_i}{1+\rho} \frac{1}{P_{si}} \left(\frac{p_{msi}}{P_{si}}\right)^{-\sigma} \qquad q^o_{msij} = \alpha^o_s \delta_b \frac{w_i}{1+\rho} \frac{1}{P_{sj}} \left(\frac{p_{msj}}{P_{sj}}\right)^{-\sigma} \tag{25}$$

where the price index is $P_{si} = \left(\int_{m=1}^{N_{s,i}} p_{msi}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$.

Denote by L_{ij} the number of individuals who choose location path (i, j). Then, the total number of young and old individuals in location i is:

$$L^y_i = t^y \sum_j L_{ij} \qquad L^o_i = t^o \sum_j L_{ji}$$

The total demand for variety m of a group s in location i is the sum of quantities demanded by young and old residents:

$$q_{msi} = t^y \sum_j L_{ij} q^y_{msij} + t^o \sum_j L_{ji} q^o_{msji} = \frac{1}{P_{si}} \left(\frac{P_{si}}{p_{msi}}\right)^{\sigma} t^y \left(L^y_i \delta_b \alpha^y_s \frac{w_i}{1+\rho} + \rho \sum_j L_{ji} \delta_b \alpha^o_s \frac{w_j}{1+\rho}\right)$$

The resulting elasticity of the demand for each variety is invariant across locations and is equal to $-\sigma$.

Amenity developers will optimally charge a constant markup over their marginal costs:

$$p_{msi} = \frac{\sigma}{\sigma - 1} k R_i \tag{26}$$

Given that all firms in the service group s in location i are symmetric, in equilibrium $q_{msi} = q_{si}$ and $p_{msi} = p_{si}$. Then, the quantity demanded can be simplified as:

$$q_{si} = \frac{t^y}{N_{si} p_{si}} \left(L^y_i \delta_b \alpha^y_s \frac{w_i}{1+\rho} + \rho \sum_j L_{ji} \delta_b \alpha^o_s \frac{w_j}{1+\rho}\right) \tag{27}$$

The number of firms in each category adjusts so that the operating profit of the developer just offsets the fixed operating cost:

$$q_{si}(p_{si} - kR_i) = F_{si} R_i \tag{28}$$

Substituting the equilibrium prices (26) and quantities (27) into free entry condition (28), I can get an expression for the equilibrium number of firms:

$$N_{si} = \frac{\frac{t^y}{1+\rho}\left(L^y_i \delta_b \alpha^y_s w_i + \rho \sum_j L_{ji} \delta_b \alpha^o_s w_i\right)}{F_{si} R_i \sigma} \tag{29}$$

The number of firms in each service category is proportional to local expenditures on this service category and inversely proportional to local housing prices, similar to Krugman (1979). As the young and the old differ in their preferences, location's demographic composition affects the composition of its establishments. Additionally, the location of the origin of the older residents affects the number of local firms by determining the retirees' expenditures.

Combining the solution to the consumer problem (24)-(25) and the equilibrium number of firms (29), I derive the indirect utility for each location path (i,j)[37]:

$$V_{ij\omega} = \ln\left(\frac{w_i a_i^y}{R_i^{(1-\delta_c)}}\right) + \rho \ln\left(\frac{w_i a_j^o}{R_j^{(1-\delta_c)}}\right) - \rho\ln(\mu_{ij}) + \ln(\epsilon_{ij\omega}) \tag{30}$$

where local amenities are determined as:

$$a_i^y = A_i^y\left(N_{1i}^{\alpha_1^y} N_{2i}^{(1-\alpha_1^y)}\right)^{\frac{\delta_b}{\sigma-1}} \qquad a_i^o = A_i^o\left(N_{1i}^{\alpha_1^o} N_{2i}^{(1-\alpha_1^o)}\right)^{\frac{\delta_b}{\sigma-1}} \tag{31}$$

D.3 Social planner problem

Social planner problem

Definition 2. *Social planner chooses consumption* $\left\{c_{ij}^{y,o}, h_{ij}^{y,o}, q_{1ij}^{y,o}, q_{2ij}^{y,o}\right\}$ *and a mass of firms in two non-tradeable sectors* $\{N_{1i}, N_{2i}\}$ *to solve*

$$\max LU$$

subject to

- *spatial mobility constraint:*

$$\ln\left(\frac{L_{ij}}{L}\right)^{\frac{1}{\theta}} = \ln\left(A_i^y (c_{ij}^y)^{\delta_c} (h_{ij}^y)^{\delta_h} (Q_i^y)^{\delta_b}\right) + \rho\ln\left(A_i^o (c_{ij}^o)^{\delta_c} (h_{ij}^o)^{\delta_h} (Q_j^o)^{\delta_b}\right) - \ln U$$

  where $Q_i^a = \left(N_{1i}^{\frac{\sigma}{\sigma-1}} q_{1ij}\right)^{\alpha_1^a} \left(N_{2i}^{\frac{\sigma}{\sigma-1}} q_{2ij}\right)^{\alpha_2^a}$ *for* $a = y, o$

- *Housing production function*

$$H_i = \overline{H}_i I_i^{\frac{1}{1+\epsilon_h}}$$

- *Resource constraint:*

$$\sum_j \sum_i L_{ij}\left(c_{ij}^y + \rho c_{ij}^o\right) + \sum H_i I_i = \sum_j \sum_i L_{ij} Z_i \left(L_i^y\right)^{\gamma_P}$$

[37] the results are equal up to a constant, which is a monotonic transformation that does not affect the final results

- *individual's allocation constraint:*

$$\sum_j \sum_i L_{ij} = L \tag{32}$$

- *Housing constraints:*

$$H_i = \sum_j L_{ij}\left(h^y_{ij} + \sum_{s=1,2} N_{s,i} k q_{sij}\right) + \rho \sum_j L_{ji}\left(h^o_{ji} + \sum_{s=1,2} N_{s,i} k q_{sji}\right) + \sum_{s=1,2} N_{s,i} F_{s,i}$$

Proposition 3. *In an efficient allocation the following should hold:*

$$\underbrace{x_{ij}}_{consumption\ cost} + \underbrace{\lambda}_{opport} = \underbrace{z^*_i + \gamma^P z^*_i}_{marginal\ product\ in\ i} - \frac{x_{ij}}{\theta(1+\rho)} \tag{33}$$

$$\underbrace{R^*_i N_{si} F_{si}}_{housing\ cost} = \underbrace{\frac{\sigma}{\sigma-1}\frac{1}{\sigma}\frac{\delta_b}{1+\rho}\left(\alpha^y_s \sum_j x_{ij} L_{ij} + \rho \alpha^o_s \sum_j x_{ji} L_{ji}\right)}_{utility\ gain} \tag{34}$$

for $s = 1, 2$, where $x_{ij} = c^y_{ij} + \rho c^o_{ij} + R^*_i(h^y_{ij} + \sum_{s=1,2} N_{si} k q^y_{sij}) + \rho R^*_j(h^o_{ij} + \sum_{s=1,2} N_{sj} k q^o_{sij})$ is the cost along the path, R^*_i is the multiplier under the housing constraint of the city i and λ is the multiplier under the labor allocation constraint

Proof. The Lagrangian of the planning problem is:

$$\begin{aligned}
L = \ln U - \sum_{ij} \omega_{ij} (\ln &\left(A^y_i (c^y_{ij})^{\delta_c} (h^y_{ij})^{\delta_h} \left(N_{1i}^{\frac{\sigma}{\sigma-1}} q_{1ij}\right)^{\delta_b \alpha^y_1} \left(N_{2i}^{\frac{\sigma}{\sigma-1}} q_{2ij}\right)^{\delta_b \alpha^y_2}\right) + \\
\rho \ln &\left(A^o_i (c^o_{ij})^{\delta_c} (h^o_{ij})^{\delta_h} \left(N_{1j}^{\frac{\sigma}{\sigma-1}} q_{1ij}\right)^{\delta_b \alpha^a_1} \left(N_{2j}^{\frac{\sigma}{\sigma-1}} q_{2ij}\right)^{\delta_b \alpha^a_2}\right) - \ln U - \ln\left(\frac{L_{ij}}{L}\right)^{\frac{1}{\theta}})\\
+R^*_i \left(\overline{H}_i I_i^{\frac{1}{1+\epsilon_h}} - \sum_j L_{ij}\left(h^y_{ij} + \sum_{s=1,2} N_{s,i} k q^y_{sij}\right) - \rho \sum_j L_{ji}\left(h^o_{ji} + \sum_{s=1,2} N_{s,i} k q^o_{sji}\right) - \sum_{s=1,2} N_{s,i} F_{s,i}\right)\\
+P^* \left(\sum_j \sum_i L_{ij} Z_i \left(\sum_j L_{ij}\right)^{\gamma_P} - \sum_j \sum_i L_{ij}\left(c^y_{ij} + \rho c^o_{ij}\right) - \sum H_i I_i\right)\\
+\lambda \left(L - \sum_j \sum_i L_{ij}\right)
\end{aligned}$$

Where $P^*, \lambda, \omega_{ij}, R^*_i$ are Lagrangian multipliers. Normalize $P^* = 1$. The first order conditions are:

$$[L_{ij}] : R_i^*(h_{ij}^y + \sum_{s=1,2} N_{s,i} k q_{sij}^y) + \rho R_j^*(h_{ij}^o + \sum_{s=1,2} N_{s,j} k q_{sij}^o) + c_{ij}^y + \rho c_{ij}^o = \quad z_i(1+\gamma^p) - \frac{\omega_{ij}}{\theta L_{ij}} - \lambda \tag{35}$$

$$\left[c_{ij}^y\right] : \frac{\delta_c \omega_{ij}}{c_{ij}^y} = L_{ij} \qquad \left[c_{ij}^o\right] : \frac{\delta_c \omega_{ij}}{c_{ij}^o} = L_{ij} \tag{36}$$

$$\left[h_{ij}^y\right] : \frac{\delta_h \omega_{ij}}{h_{ij}^y} = R_i^* L_{ij} \qquad \left[c_{ij}^o\right] : \frac{\delta_c \omega_{ij}}{h_{ij}^o} = L_{ij} R_j^* \tag{37}$$

$$\left[q_{sij}^y\right] : \frac{\delta_b \alpha_s^y \omega_{ij}}{q_{sij}^y} = L_{ij} R_i^* k N_{si} \qquad \left[q_{sij}^o\right] : \frac{\delta_b \alpha_s^y \omega_{ij}}{q_{sij}^o} = L_{ij} R_j^* k N_{sj} \tag{38}$$

$$[N_{si}] : \frac{\sigma}{\sigma-1} \delta_b \alpha_s^y \sum_j \omega_{ij} + \rho \frac{\sigma}{\sigma-1} \delta_b \alpha_s^o \sum_j \omega_{ji} - R_i^* \sum_j L_{ij} q_{1ij}^y k - R_i^* \rho \sum_j L_{ji} q_{1ji}^o k - R_i^* F_{1i} = 0 \tag{39}$$

$$[I_i] : R_i \frac{1}{1+\epsilon_H} \overline{H}_i I^{-\frac{\epsilon_H}{1+\epsilon_H}} - H_i = 0 \tag{40}$$

Substituting 38 into 39, I get the optimality condition for the number of firms in non-tradeable sector s:

$$\frac{1}{\sigma-1} \left(\delta_b \alpha_s^y \sum_j \omega_{ij} + \rho \delta_b \alpha_s \sum_j \omega_{ji} \right) = R_i^* F_{si} N_{si} \tag{41}$$

Denote by x_{ij} the total cost of putting an individual on a location path ij in terms of the consumption of the final good and tightening of the housing constraints (either directly or through consumption of non-tradeables):

$$x_{ij} = c_{ij}^y + \rho c_{ij}^o + R_i^*(h_{ij}^y + \sum_{s=1,2} N_{si} k q_{sij}^y) + \rho R_j^*(h_{ij}^o + \sum_{s=1,2} N_{sj} k q_{sij}^o)$$

From the FOCs the following holds the following holds:

$$\omega_{ij} = \frac{x_{ij} L_{ij}}{(1+\rho)}$$

D.4 Proof of Proposition 4

Denote $x_i = (a_i^y)^{-\theta}$ and $y_i = (a_i^o)^{-\beta\theta}$. Then the system of equilibrium conditions transforms into:

$$x = \frac{L}{U^\theta} \sum_j K_{ij} y_j^{-1}$$

$$y = \frac{L}{U^\theta} \sum_j K_{ji} x_j^{-1}$$

where K_{ij} are known positive values. The corresponding matrices of coefficients are:

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \Gamma = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

Then matrix A can be found as:

$$A \equiv \Gamma B^{-1} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

taking the absolute value of matrix coefficients we get that:

$$A^p = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

The spectral radius of the matrix is equal to 1. Applying the results from Allen et al. (2015), I get that the solution of the system exists and is unique up to the scale. So is the solution for a_i^y, a_i^o.

E Mobility Costs Estimation

E.1 Distance between CZs

I use the data from the county distance database[38] that computes great-circle distances calculated using the Haversine formula based on internal points in the geographic area for all pairs of counties in the US. To compute the distances between the CZs, I map each 1990 county to a single 1990 CZ using the crosswalk in Eckert et al. (2020) based on the largest area of the intersection.[39]. Then I compute a distance between CZs as an area weighted average distance between counties in two CZs.

E.2 Mobility Data

For the main regression I use Census IPUMS data for the year 2000 that provides individual level data on the current location and location 5 years ago at the PUMA level. I aggregate all migration PUMAs to commuting zones and compute $L_{ij}^5(a)$ - the number of individuals of age a who were in the CZ i five years ago and are currently in the CZ j using the crosswalk weights between migration PUMAs and CZs. As the crosswalks weights provide many to many

[38]https://www.nber.org/research/data/county-distance-database

[39]because 1990 CZs were constructed as a collection of 1990 counties, a one to one mapping works well

mapping, for each PUMA to PUMA migration flows, I find all CZs corresponding to origin and destination PUMAs, and I allocate the weight of the observation to CZ flows according to the area intersection weights provided in Autor & Dorn (2013).

E.3 Estimating mobility regression

In the data I observe $L_{ij}^5(a)$. I use five year mobility data for individuals of age $a \in [63, 72]$. I further assume that conditional on reallocation, the probability that it happens at age a is p_a:

$$p(a) = p\lambda^{(a-58)} \qquad \frac{p(1-\lambda^{15})}{1-\lambda} = 1$$

That is probabilities decline with age at the rate λ, and the sum of probabilities between ages 60 and 75 should be equal to one.

Let Π_5 be the probability that an individual is in the same location as 5 years ago. Then:

$$1 - \Pi_5(a) = (1 - \Pi)p_a^*$$

where $p_a^* = p\lambda^{(a-65)}(1 + \lambda + \lambda^2 + \lambda^3 + \lambda^4)$ is the probability of reallocation within the last 5 years.

Given the assumptions made above, the probability that an individual of age $a \in [63, 72]$ is in the same location as 5 years ago is:

$$\Pi_5(a) = \Pi(a) + (1 - \Pi(a))(1 - p_a^*)$$

$$\Pi(a) = \frac{1}{p_a^*}(\Pi_5 - (1 - p_a^*))$$

Given the assumptions made in the previous section:

$$\frac{L_{ij}^5(a)}{L} = \pi_{ij} p_a^*$$

$$\sum_{a=63}^{a=72} \frac{L_{ij}^5(a)}{L} = \pi_{ij} \sum_a p_a^*$$

Therefore, the share of individuals who moved from i to j within a 5 year window corresponds to the share of individuals who move from i to j up to the scale. As the scale parameter will be absorbed in the fixed effects, I can use $\sum_{a=63}^{a=72} L_{ij}^5(a)$ as a proxy for L_{ij} in the main regression.

F Exogenous productivities, amenities and housing supplies

This section shows how given the structural parameters calibrated in section 5.2, composite amenities $\{a_i^y, a_i^o\}$ recovered in section 5.3.2 and data $\{L_i^y, R_i, w_i\}$, I can invert the remaining exogenous fundamentals of the model $\{\tilde{A}_i^y, \tilde{A}_i^o, Z_i, \overline{H}_i\}$. First, I compute model-implied migration flows $\hat{L}_{ij}$ according to equation (9) . This allows me to calibrate exogenous adjusted amenities for young and old across locations using their model-based definition (16):

$$\tilde{A}_i^y = a_i^y \left(\frac{E_{1i}^{\alpha_1^y} E_{2i}^{(1-\alpha_1^y)}}{R_i} \right)^{\frac{-\delta_b}{\sigma-1}} \qquad \tilde{A}_i^o = a_i^o \left(\frac{E_{1i}^{\alpha_1^o} E_{2i}^{(1-\alpha_1^o)}}{R_i} \right)^{\frac{-\delta_b}{\sigma-1}}$$

where E_{si} is the total expenditure in the city i on category s and can be approximated as:

$$\hat{E}_{s,i} = \left(L_i^y \delta_b \alpha_s^y \frac{w_i}{1+\rho} + \rho \sum_j \hat{L}_{ji} \delta_b \alpha_s^o \frac{w_j}{1+\rho} \right)$$

Given the observed wage vector $\{w_i\}$ and calibrated productivity spillover value γ^P, I can invert the exogenous productivities across locations:

$$Z_i = w_i \left(L_i^y\right)^{-\gamma^P}$$

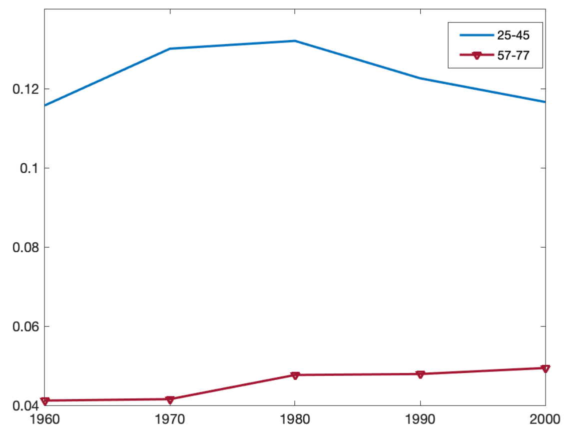
Finally, given the data $\{R_i, w_i, L_i^y\}$ and the predicted flows $\hat{L}_{ij}$, I recover exogenous housing supplies $\{\overline{H}_i\}$ using equations (10) and (11):

$$\overline{H}_i = \frac{(1-\delta_c)}{d_h(1+\rho)} R_i^{-(1+\frac{1}{\epsilon_h})} (L_i^y w_i + \rho \sum \hat{L}_{ji} w_j)$$

G Changes in interstate mobility rates, 1960-2000

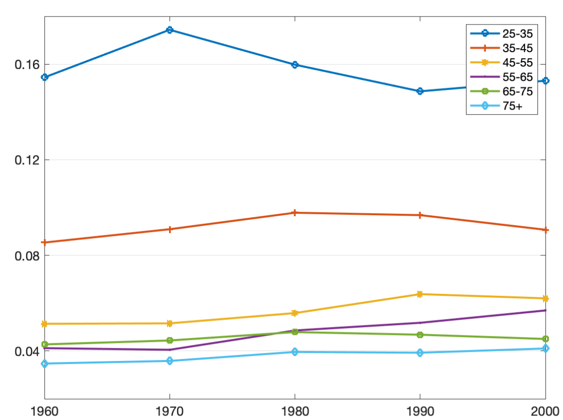
In Census samples for years 1960, 1970, 1980, 1990 and 2000, I observe whether individuals live in a different state as compared to their state of residence five years ago. I consider a subsample of individuals for whom the data on the state of residence five years ago is available, and who resided in the US five years ago. I assume that reallocation for retirement reasons occurs within a 20 year window around the average retirement age of 64. Therefore, I focus on the age group 57-77. For comparison, I also consider the mobility of the younger age group 25-45. Figure G5 shows the dynamics of the interstate mobility rates for two age groups. For the older age group, the mobility has been increasing throughout the sample period from around 0.04 in 1960 to almost 0.05 in 2000. In contrast, the mobility rate for the younger age groups has been rising during 1960-1980 period and has been declining afterwards. Figure G6 breaks the data into the more detailed age groups, and shows that the increase in the interstate mobility starts for the age group 45-55, and peaks for the age group 55-65.

Figure G5: Interstate migration rates over a five year window



Notes: The figure plots the interstate migration rates for younger and older age groups. In each Census year, I consider a subsample of individuals for whom the migration information is known and for whom their place of residence 5 years ago is available. I also exclude individuals who lived abroad 5 years ago.

Figure G6: Interstate migration rates over a five year window



Notes: The figure plots the interstate migration rates for different age groups. In each Census year, I consider a subsample of individuals for whom the migration information is known and for whom their place of residence 5 years ago is available. I also exclude individuals who lived abroad 5 years ago.

The finding on an increasing interstate mobility around the retirement age might be in contrast to the existing evidence on declining interstate mobility in the US. Kaplan & Sam-Schulhofer-Wohl (2017) document a declining interstate yearly migration rate for the working age individuals over 1990 - 2010 period using CPS data. Molloy et al. (2011) use Census data to show that 5-year migration rate was rising over 1960-1980 period and declining afterwards. Using same dataset as Molloy et al. (2011), I also find that the interstate mobility has been declining over 1980-2000 period for the younger age group. However, I show that 5-year mobility rates have been rising for the older age group over the same time period. As argued in Kaplan & Sam-Schulhofer-Wohl (2017), the major driver of the migration decline was a decline in the geographic specificity of returns to occupations, and also the increase in workers ability to learn about other locations before moving. Around retirement age work-related reasons are no longer crucial for moving, which might explain why older individuals have a different mobility dynamics.

H Local policies to attract retirees

This section considers locally financed policies aimed at attracting retirees. I consider the least productive U.S. CZs (cities in the 25th quantile by productivity). Then I compute counterfactual equilibria where all relocating retirees to these CZs are given a subsidy s. This subsidy is paid as a lump-sum tax on local workers. Figure H7 illustrates the local effects of the policy as a function of the subsidy size. Overall, the policy is not successful at increasing the location's population size and its amenities. The reason is that subsidising retirees is costly for workers, as a result less workers and future retirees choose to live in the city. So the major effect of the policy is the transfer from local young to local old.

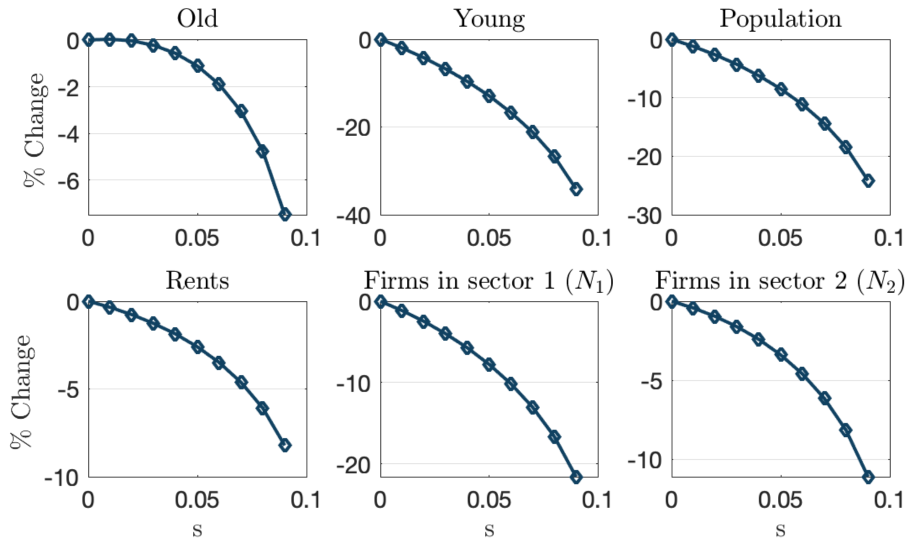
I Changes in sorting, 1960-2010

I.1 MSA calibration

Due to the data availability issues, historical comparison between 1960 and 2010 is available only at the MSA level. This section provides additional details on the mapping of the model to the MSA data for the years 1960 and 2010. I select 169 MSAs available in the Census data for both time periods.[40] Using the Census data I construct the vectors of wages, rents, young and old location choices $\{W_{i,t}, R_{i,t}, L^y_{i,t}, L^o_{i,t}\}$ corresponding to 169 US MSAs for years $t = 1960, 2010$.

[40] 169 MSAs available in both time periods cover 58 percent of the population in 1960 and 67 percent of the population in 2010

Figure H7: Effects of the policy that subsidises the reallocation of old to less productive cities financed locally



Notes: Simulated local changes due to the introduction of a subsidy s measured as a percentage of individuals' average wages in the economy given to any old individual moving to a treated city. The policy is financed via a lump sum tax paid by all workers in the treated cities. Treated cities are the least productive U.S. CZs.

[41] I keep all the structural parameters fixed at their calibrated values described in section 5.2, except for the relative length of the old period ρ and the fixed mobility cost κ. I recalibrate ρ_t for each year $t = 1960, 2010$ so that to match the ratio of the sizes of the older and younger age groups. This implies that ρ_t is by 20% lower in 1960 as compared to 2010.

I take mobility costs for year 2010 to be equal to the calibrated mobility costs for the full model (see section 5.3), where the quantification was based on the 5 year window mobility data for the year 2010. Due to the data limitations, I cannot reestimate the elasticity of mobility costs with respect to distance in 1960. So I assume that the change in mobility costs happens through the fixed cost of moving κ. I further calibrate the fixed mobility costs for 1960 to match the evidence on an increase in the older population mobility during 1960-2000 period.[42] Due to the data limitations, I can only compare the individual level mobility data across U.S. states within a 5 year window. I compute that mobility around retirement age increased by around 20% over 1960-2000. I further calibrate the fixed component of mobility costs for retirement model in 1960 together with the amenities for young and old $\{a^y_{60}, a^o_{60}\}$ so that the share of older

[41] Construction of the samples $\{W_{i,t}, R_{i,t}, L^y_{i,t}, L^o_{i,t}\}$ follows the same steps as in the main calibration, see section A

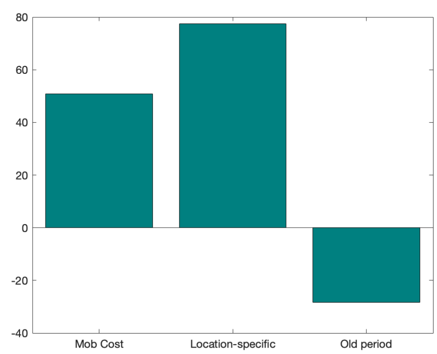
[42] The analysis can be done only for 1960-2000 period because in 2010 Census there is no comparable mobility variable. 1960 and 2000 Census have data on inidividual's location 5 years ago, while 2010 Census has data on individual's location 1 year ago.

individuals who reallocate between states is 20 percent lower than in 2010.[43].The calibration procedure follows the same steps as in the main text of the paper. This allows me to compute that a model implied decline in mobility costs is equal to 36 percent. The following steps are the same as in the main calibration, except for the different vectors of the data.

I.2 Decomposing changes in sorting, 1960-2010

The correlation between the MSA's old-to-young ratios and MSA's productivity changed from 0.01 in 1960 to -0.27 in 2010. This section describes the decomposition of this change. I start with an equilibrium in 1960 which by construction fully matches the data for that year. Then I start changing exogenous parameters one by one: first mobility cost of moving κ, then exogenous geography parameters A_i, Z_i, H_i, and finally the relative length of the old period ρ. The decomposition results are presented on figure I8.

Figure I8: Decomposition



Notes: The figure decomposes the changes in the correlation between MSA's old-to-young ratios and MSA's productivity from 0.01 in 1960 to -0.27 in 2010. The location-specific factors include the set $\{A_i^y, A_i^o, Z_i\}$.

I.3 Robustness Checks

First, I check whether the main results are robust to alternative specification of the ρ_{1960} parameter. Initially it was calibrated so that to match the 20% increase in the old-to-young ratios over 1960-2010 period. Table I2 shows that the model implied changes in mobility costs are not sensitive to the alternative specifications of ρ_{1960}. The reason is that changes in the relative length of the old period primarily affect the calibrated values of the amenities for the old and the young. Table I2 further shows that the welfare gains from calibrated change in

[43]Here I implicitly assume that an increase in mobility across MSAs over the same time period is the same as across MSAs

mobility costs are higher for the smaller change in ρ. The reason is that a smaller change in ρ implies larger value of ρ in 1960, which makes the retirement period more important at that time and thus increases the importance of mobility costs for the old.

TABLE I2
Robustness of the results to changes in ρ over 1960-2010 period

% Increase in ρ	% Implied change in κ	% Welfare change due to $\Delta\kappa$
10	-27.62	3.68
20	-27.62	3.68
30	-27.62	3.67

J Additional tables

TABLE J3
Determinants of model implied amenities

Dependent Variable	(1) a_i^o/a_i^y	(2) a_i^o/a_i^y
Golf and Country clubs	0.41	1.1
	(0.06)	(0.13)
Recreation	−0.06	0.16
	(0.33)	(0.25)
Hospitals	0.41	1.23
	(0.13)	(0.34)
Personal care	0.35	0.93
	(0.06)	(0.18)
Nursing homes	0.33	0.48
	(0.06)	(0.2)
% Water area	0.09	0.42***
	(0.027)	(0.08)
January temperature	0.16	0.04
	(0.06)	(0.09)
Food & drinking	−0.14	−2.05
	(0.2)	(0.63)
Movie theaters	−0.4	−0.63
	(0.05)	(0.14)
Childcare	−0.17	0.08
	(0.07)	(0.15)
Schools	−0.06	−0.24
	(0.04)	(0.12)
Museums and Parks	−0.06	0.36
	(0.04)	(0.12)
Clothing	−0.07	0.62**
	(0.07)	(0.26)
Arts	0	−0.39
	(0.06)	(0.11)
July temperature	−0.02	0.25
	(0.04)	(0.09)
Topography	−0.008	0.15
	(0.02)	(0.06)
Observations	722	722
R-squared	0.34	0.67
Weighted by CZ population		✓

Notes: Robust standard errors are given in parenthesis. All variables except nature are in logs. USDA nature variables are standardized. Includes 722 mainland commuting zones available in USDA data.

TABLE J4

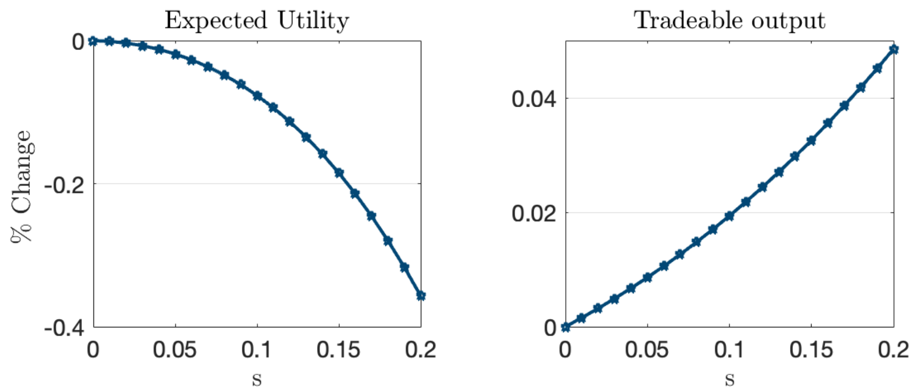
LOCAL EFFECTS OF THE PLACE-BASED POLICIES TARGETED AT RETIREES

L	L^y	L^o	N^y	N^o	a^y	a^o	R
5.36	-3.4	19.43	3.01	10.6	0.3	0.4	3.8

Notes: The table shows the percentage response of the corresponding local variables to a 10% subsidy for retirees for moving into a city. N^y denotes the number of firms in the non-tradeable category 1 that is valued more by young. N^o denotes the number of firms in the non-tradeable category 2 that is valued more by old.

K Additional figures

Figure K9: Aggregate effects of the policy that subsidises the reallocation of old to smaller cities



Notes: Aggregate effects of the policy that introduces of a subsidy s measured as a percentage of individuals' average wages given to any old individual moving to a treated city. The policy is finances via a lump sum tax paid by all workers in the economy.Treated cities are the cities in 25th quantile by productivity