Fertility, Housing Costs and City Growth

Nicolas Coeurdacier
SciencesPo and CEPR

Pierre-Philippe Combes
SciencesPo-CNRS and CEPR

Laurent Gobillon
Paris School of Economics-CNRS and CEPR

Florian Oswald
SciencesPo

This version: December 21, 2022
Very preliminary, please do not circulate.*

Abstract

Developed economies moved from a baby boom to a baby bust, with very low fertility in large urban centers. We develop a spatial theory of demographic change and urban growth, whereby the housing market acts as an endogenous automatic stabilizer of fertility: high housing costs deter fertility while lower demographic growth mitigates future housing price increase. Our theory predicts rich dynamics of fertility and housing costs across space and time. The aging of baby boomers and their sorting across space triggers a later baby bust, more pronounced in larger cities, together with a long-lasting fertility and house price cycle that varies across locations. The predictions of the theoretical model with endogenous demographics and city growth are confronted to the data for French urban areas over the last decades.

Keywords : Fertility, Population Dynamics, Housing Prices, Sorting.
JEL codes: J11, J13, R21, R23.

*We would like to thank our colleagues as well as seminar/conference participants for comments and insights. We thank Hugo Lhuillier for research assistance. We gratefully acknowledge funding under the Banque de France-SciencesPo Partnership and the Alliance Program. Contact: SciencesPo, 28 rue des saint-pères, 75007 Paris, France. Nicolas Coeurdacier nicolas.coeurdacier@sciencespo.fr; Pierre-Philippe Combes ppcombes@gmail.com; Laurent Gobillon laurent.gobillon@psemail.eu; Florian Oswald florian.oswald@sciencespo.fr.

1 Introduction

The centers of large French and U.S. cities are childless and average fertility is lower in more populated cities. In the central part of Paris, the largest urban area, the ratio of children to adults in parental age is about 50% lower than in smaller cities of less than 100,000 inhabitants. The difference is even larger when comparing the center of Paris to far away suburbs of the urban area.[1] More broadly, fertility is significantly lower in denser locations. Figure 1 illustrates the (negative) density-dependence of fertility across space in French urban areas in 2015—the ratio of children to adults in parental age is 0.35 in the densest municipalities, compared to 0.65 in the least dense ones. This fact is very robust across time and holds both for France and the US.

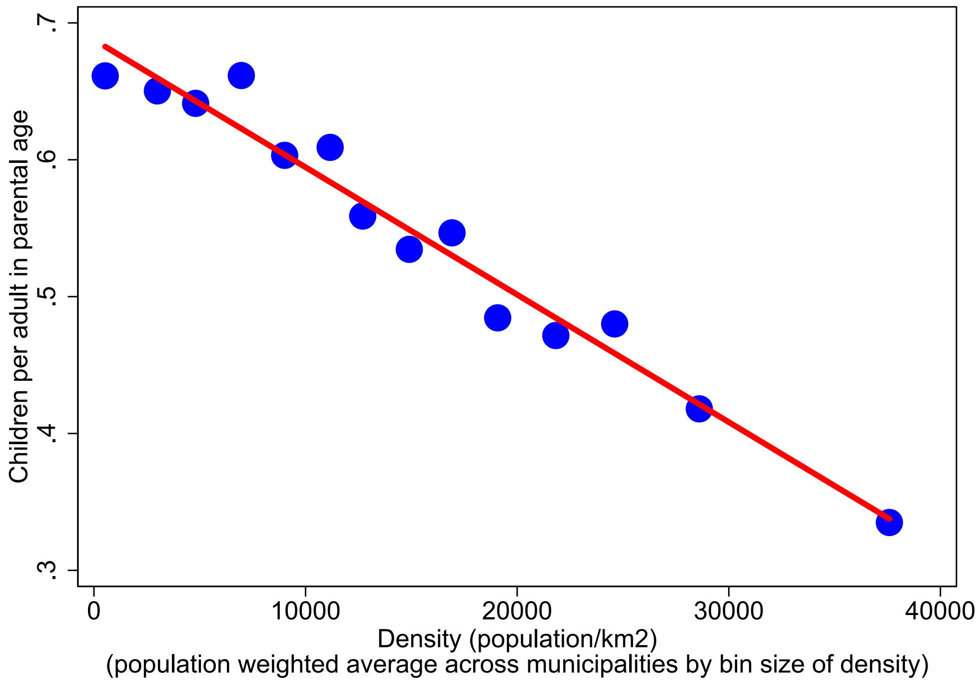


Figure 1: The (negative) density-dependence of urban fertility.

Notes: The figure shows the number of children (age 1-17) per adult in parental age (age 27-53) as a function of the density of the municipalities, binned into 14 categories of density ([0,2000[, [2000,4000], ..., [18000, 20000[, [20000, 23000[, [23000, 26000[, [26000, 30000[, and above 30000). Density (population per km^2) and children per adult in parental age in each bin are defined as the municipal population weighted average of the municipalities in this bin. The red line is a linear regression line. *Source*: French Census 2015.

To account for these facts, we develop a theory of density-dependent fertility based on the scarcity of land, the price of children being tied to housing costs and the price of land. The idea is intuitive: if some locations are more attractive than others, they will be more densely populated and living in these locations will be more expensive—land will be scarce there and housing prices higher. To the extent that raising children implies buying more living space, these locations should also exhibit

[1] Figure B.1 shows how the number of children per adults in parental age depends on urban are population. (resp. B.2) shows how it depends on the distance from the center of Paris within the Parisian urban area.

lower fertility overall. Families will be larger in cheaper and less dense locations. Data for France strongly supports this main prediction of our theory between and within city. A natural implication is that the reallocation of individuals away from central locations ('suburbanization') should go hand in hand with rising fertility. To the opposite, if parents relocate towards the center of large cities, fertility should fall. Both mechanisms also apply between small and larger cities. While identifying the channel through which fertility is higher in cheaper locations turns out to be a challenging task due to the endogeneity of fertility and housing prices and the possibly heterogeneous preferences for children in the society, we make use of the quantitative nature of our model that we structurally estimate to investigate the importance of housing costs in driving fertility across space and time. Doing so, we extend quantitative spatial models (surveyed in Redding and Rossi-Hansberg (2017)) with endogenous fertility and demographics.

Beyond childless centers in large cities, our approach sheds new light on the trend of declining fertility over the last decades. In many developed countries, fertility is now below replacement rates (United Nations, 2019). If this trend were to continue, population will eventually shrink, raising concerns regarding the ability of economies to cope with aging populations and more broadly, the sustainability of welfare states. A shrinking population might also threaten the future of productivity growth and innovation (Jones, 2020). Our theory enlightens a particular mechanism, which can explain the evolutions of fertility, complementing alternative theories à la Becker (1960): rising housing prices over the recent decades (Knoll et al., 2017; OECD, 2022) deter families from having children—a phenomenon particularly acute nowadays in large cities.

However, in the future, an important prediction of our theory is that economies eventually converge towards a stable population. While fertility might fall as long as housing prices are rising, it should eventually converge back towards replacement in the long run. The intuition goes as follows: high fertility and population growth puts pressure on land, rising its value. This, in turn, lowers the affordability of housing for families, who reduce their number of offspring. Population growth falls. If population shrinks with fertility below replacement, the demographic pressure on land diminishes. Housing becomes more affordable, families are willing to increase their size and population growth increases again. The housing market acts as an automatic stabilizer of fertility. This is the reason why our theory predicts that, after long lasting transitional dynamics, average fertility converges towards replacement—with cheaper cities having above replacement fertility while expensive ones exhibit below replacement fertility. Migration across locations over the life-cycle insures a stable spatial long-term distribution of population across cities. We also show how, under some demographic shifts or changes in housing supply conditions, economies can go through periods of very low fertility before eventually converging back towards the replacement rate. More specifically, our theory predicts that the baby-boom post World War II initiated a long-lasting fertility and housing price cycle across space—the ageing of baby-boomers triggering a later baby bust. A last important feature of our theory is that the joint location and fertility decisions dampen the housing prices increase in more productive location, larger cities and city centers, mitigating the effect on aggregate productivity of rising housing costs. More broadly, the theory sheds light on the importance of demographic changes for the spatial sorting of workers and aggregate urban productivity.

Our contribution is threefold. First, we provide an independent empirical contribution documenting how fertility behavior and housing costs are related. While these stylized facts are only suggestive that housing prices matter for fertility decisions, they motivate our quantitative theory of population dynamics, housing costs and city growth. This is the second and main contribution of the paper. Third, we estimate the model using French data across urban areas and simulate the spatial impact of the baby boom since 1960.

On the empirical front, we document a set of novel stylized facts regarding fertility and demographic sorting across space in France that helps motivating the theory. In a given French urban area, we show that fertility is higher in more suburban locations—the ratio of children to adults in parental age increases with distance from the city center. Across urban areas, the same ratio is lower in more populated and more expensive urban areas. More broadly, fertility is significantly higher in cheaper and less dense urban locations—with large spatial variations as shown in Figure 1.[2] These findings hold controlling for a number of observable local characteristics and are robust across time. Although one cannot rule out that unobserved heterogeneity across locations correlated with fertility is driving this pattern (e.g. children friendly amenities), this is consistent with our price mechanism, whereby higher housing costs increase the price of children. We also show how housing expenditures in a given location are increasing with household size—the most crucial theoretical assumption for our story. Lastly, we provide further evidence on the sorting of younger and older households. Younger households sort into more expensive locations (in line with Couture and Handbury (2020) and Moreno-Maldonado and Santamaria (2022) for the U.S.) and, most importantly for us, older households compete with parental households for cheaper locations—their presence crowds out families and children. This latter observation is at the heart of our story: the aging of baby-boomers relocates families towards areas where children are more expensive.

Equipped with this empirical evidence, we develop a quantitative life-cycle spatial equilibrium model able to account for the observed patterns of fertility and demographic composition across cities over time. The economy is populated with overlapping generations of individuals going through four stages of life: children, young, parents and old. At each stage of life, households choose where to live among a discrete set of locations and consume goods and housing space out of their location-specific income. At the parental stage, households also decide on their discrete number of children depending on their preferences for fertility and the housing costs of children—the demand for housing space being increasing in household size. Residential locations differ in their productivity, amenity and endogenous housing prices. Depending on their stage of life, households face different a trade-off for these characteristics. Younger households without children favor the most productive locations, while elderly favor less productive/high amenity locations—their income at retirement

[2] While the spatial variations of fertility are found to be quantitatively large, this can be driven by a large response of fertility to housing costs or by the endogenous sorting of households according to their fertility preferences—households willing to have a large family moving to cheaper locations. The quantitative nature of the theory helps disentangling both dimensions.

being less dependent on location-specific productivities. Most importantly for our story, parents sort into cheaper locations due to the housing costs of children, even more though when they have a stronger preference for fertility. This generates density-dependent fertility across space.

Our model is estimated on census data for France at the municipality level. We show how the demographic composition across space helps retrieving location-specific productivities and amenities. Based on this identification strategy, the model is able to replicate the dispersion of housing values and fertility across and within cities. It also matches the spatial sorting of younger and older households.

It is important to note that housing costs and fertility are jointly determined. Housing costs feedback into fertility decisions and population dynamics implied by fertility also affect subsequent housing values since demographic pressure changes the demand for housing. Thus, our overlapping generations model can shed light on the past and future dynamics of demographics and housing costs across French urban areas—demographic composition at older age being implied by age-specific mortality rates. More specifically, using municipality data since 1960, we simulate the consequences of the post-World War II baby-boom driven by a stronger preference for children and declining commuting costs fostering suburbanization (Baum-Snow (2007), Redding (2021)). While it initially triggered faster growth in suburban areas and smaller cities, the ageing of baby boomers, together with their increased longevity and their sorting across space, generated a later relocation of workers towards more productive and expensive locations. Together with the implied demographic pressure on land, this generates a 'baby bust', more pronounced in larger cities. This is in line with the current demographics of French urban areas. Looking forward, the current drop in fertility mitigates future housing price increases and fertility is expected to rebound, even more so as the mortality of the generation of baby boomers frees up space. Importantly, the evolution in the sorting of workers following these, past and future, demographic changes generates a novel trade off between demographic growth and aggregate urban productivity.

Related literature. Our theory relates to several different domains. It first relates to biology and ecology, which develops density-dependent population dynamics for other species. Many species exhibit rising mortality and falling fertility when density rises (see Sibly and Hone (2002) and Sinclair (1989, 2003), Mills (2012) for references; see discussion in Lutz et al. (2006)). For those species, these Malthusian dynamics are likely driven by the lack of food resources. However, our theory shows that, even though humans might have solved the 'food problem', they might still be subject to similar dynamics to the extent that resources spent on children are linked to the costs of housing. Lee (1987) provide some extensive discussion of density-dependent population dynamics and their relevance for humans.[3]

Our paper also relates to the pioneering literature in sociology and demography that links the pro-

[3] In a Malthusian perspective, Lee (1987) argues that the main source of density-dependence in human population dynamics has severely weakened with the diminishing role of agriculture in modern societies. Our view is that to the extent that housing costs increase with demographics pressure, particularly so in some locations, this might limit fertility in a similar fashion as having less land per person for agriculture did limit population growth in the past.

cess of urbanization to the dynamics of fertility to explain the demographic transition. Starting with Thompson (1916, 1929), Davis (1937) and Notestein (1945) (Caldwell, 2006, for a survey), the 'demographic transition theory' discusses how the decline in fertility mirrors the process of industrialization and urbanization through economic and social changes (declining returns to children relative to agrarian societies, social mobility and new job opportunities, higher returns to education, change in the role of women and modernization of family norms and values, among possible channels).[4] While relevant to account for the fertility decline before WWII, the theory was challenged by the post-war baby boom. Instead, our approach emphasizes the scarcity of urban land and the density in cities to explain fertility decisions. This generates transitional dynamics over generations from lower to higher fertility depending on the availability of urban land.[5] In a similar vein, Easterlin (1961) generates a long-lasting fertility cycle through congestion in labor markets instead of housing markets—whereby smaller cohorts have better job opportunities and can afford larger families.

Our paper also relates to the macroeconomic approaches of fertility in the long run. The source of the observed fertility decline has attracted a considerable amount of attention in economics since the pioneering work of Becker (1960). Most theories surveyed in Jones et al. (2008) and Doepke et al. (2022) rely on the opportunity cost of time to raise children and/or the quantity-quality trade-off in children (Becker and Lewis (1973), Willis (1973), see Hotz et al. (1997) for a survey). Both leading theories arguably forecast that modern economies are slowly going towards shrinking populations and an 'empty world' (Jones (2020)). Our theory linking the price of children to the price of housing provides a complementary approach to explain the evolution of fertility across time—quantitatively relevant since housing is the main component of children costs (see Lino et al. (2017)).[6] It is important to note that the implications of our theory differ drastically since our density-dependent theory of population growth predicts a stable long-term population.[7] Beyond aggregate implications for fertility, our theory also delivers predictions across space. Related to our approach, De la Croix and Gobbi (2017) show empirically that fertility is density-dependent in emerging countries using spatial data (see Lutz et al. (2006) for related findings). The most closely related work to ours is Sato (2007). He identifies in the context of Japan regional variations of fertility that are related to regional variations in population density. He provides a theory based on economic geography where agglomeration and congestion forces determine regional fertility differences through the combination of income and substitution effects. Our approach instead puts the scarcity of land and the price of housing at the heart of fertility variations across space.

A literature investigates the role of demographics on housing prices as pioneered by Mankiw and Weil

[4]See summary in Notestein (1953) which stresses the 'urban industrial society' as the necessary context for these social and economic changes to occur.

[5]Easterlin (1976) also emphasizes land availability to explain fertility patterns in rural America in the 19th century.

[6]In Murphy et al. (2008), children are also costly in terms of housing space and changes in the rental price of housing post WWII contributed to the baby boom. Due to the sorting of individuals, our approach provides predictions across space, which can be mapped into spatial data on demographics and housing prices.

[7]This result is a reminiscent of Eckstein et al. (1988) where the presence of a fixed factor, land, is guaranteeing convergence towards a stable population.

(1989) (see Hiller and Lerbs (2016) and Bolhuis and Cramer (2020) for recent contributions). While similar mechanisms are at play in our theory, we emphasize the joint determination of population dynamics and housing costs. Demographic shifts change the demand for housing space, and, vice-versa, shifts in housing demand and resulting housing price evolutions feedback on fertility decisions and population dynamics. We also emphasize the spatial nature of demographic changes, initially absent in these more macroeconomic studies.

We also contribute to the literature in urban economics focusing on the sorting of individuals across space. Its main focus is the sorting across skills and income (see Glaeser and Mare (2001), Combes et al. (2008), Baum-Snow and Pavan (2011), Eeckhout et al. (2014), Diamond (2016), Roca and Puga (2017), Guerrieri et al. (2013) and Couture et al. (2019) among others) while ours is the sorting of individuals in terms of demographics and fertility preferences. Couture and Handbury (2020), Moreno-Maldonado and Santamaria (2022) and Komissarova (2022) also emphasise the trade-off between job opportunities, prices, and amenities might change over the life-cycle. Couture and Handbury (2020) shows how the location choice of young college-educated individuals contributed to the revival of large U.S. urban centers. Moreno-Maldonado and Santamaria (2022) emphasizes the role of delayed child-bearing in this revival. Komissarova (2022) focuses on the sorting across space of retirees. While the sorting of young workers and retired baby-boomers is quantitatively important for our story, our approach puts more emphasis on the sorting of parents and their fertility decisions across space, which in turn feedback on future demographic and housing market evolutions.

Lastly, our empirical evidence about the effect of housing costs on fertility behavior is related to a series of papers with similar objectives (Simon and Tamura (2009), Lovenheim and Mumford (2013) and Dettling and Kearney (2014)). These approaches differ from ours as they identify the effect of a change in housing values on fertility behavior at shorter horizons, investigating the decision of having an additional child if housing prices increase. Our focus is the change in fertility decision of new cohorts of parents, which face different housing prices compared to previous ones, and we make use of the structure of the quantitative model to tease out the impact of housing costs on fertility. These papers are also relatively silent regarding the sorting of households across space.[8]

The paper is structured as follows. Section 2 provides motivating empirical evidence on the patterns of fertility and demographic sorting across French cities. Section 3 develops a spatial quantitative equilibrium model over the life cycle, which jointly determines the dynamics of urban population and housing prices. Section 4 estimates the model using French data across municipalities and run experiments to shed light on the past and future evolutions of the demographic growth of French urban areas since the Baby-Boom. Section 5 concludes.

[8] A large empirical literature surveyed in Jones et al. (2008) focus on the effect of income on fertility. Our approach focusing on the spatial dimension of fertility choice is complementary and abstracts from income/substitution effects driving fertility decisions. Black et al. (2013) shows how one needs to control for location choice when evaluating if children are 'normal goods'.

2 Empirical evidence on demographics across space

2.1 Data

Spatial units. Cities are defined as urban areas according to their 2010 official definition. Urban areas are made of different municipalities. There are 792 urban areas in France and around 36,000 municipalities. Municipality coordinates are those of the city hall. In a given year, we compute the urban area center as the barycenter of municipalities, weighting by employment from SAPHIR data detailed below. The distance of a municipality to the city center is defined as the greater circle distance between this municipality and its urban area.

Household census data. We use the SAPHIR dataset that contains harmonized individual census data over the 1968-2015 period to compute demographic variables and other average household characteristics at the municipality level. We only keep ordinary dwellings and restrict the sample to households with at most two adults. Demographic variables are defined according to the age of individuals. The household age is mean age of adults in the household. We consider four age brackets consistent with the life-stages of our theoretical model of Section 3: children correspond to individuals with age in the [0,18[interval, young to individuals with age in the [18,27[interval, parents to individuals with age in the [27,54[interval and old to individuals above 54.

The fertility rate in a given municipality is defined as the ratio between the number of children (below 18) and the number of adults in parental age (in the [27,54[interval).[9] The young to adults ratio is defined as the ratio between the number of young (in the [18,27[interval) and the number of adults (above 18). The old to adults ratio is defined as the ratio between the number of old (above 54) and the number of adults (above 18). Municipal averages are computed using weights at the household level.

Housing prices. Housing price indices were constructed from notary databases by Chapelle et al. (2022). Notary data contain information on transactions of second-hand dwellings occurring in mainland France every even year over the 2000-2012 period. They contain the municipality identifier, the selling price, the floor area, the dwelling type (single-family house or flat) and the construction year in interval brackets. For every dwelling type and available year, the logarithm of housing price per square meter is regressed on dummies for the construction period in brackets (ref. after 1991) and for the quarter (ref. 1^{st}). The log-price index for given dwelling type and year is then defined as the municipal average of the constant plus residual for the corresponding subsample. The log-price index for all dwellings is defined as the average of indices for the two dwelling types, weighting by their proportions in the data.

Housing consumption. Data from the Enquête Nationale Logement (ENL) provide data on housing consumption at the household level every four years over the period 1984-2006 and in 2013 for

[9] Our fertility measures differs from the one of demographers who focus on the number of children per women in reproductive age. This measure is a more direct counterpart of the model.

a representative sample of French households.[10] ENL data provides the municipality of residence of the household, the floor space and other housing attributes, housing expenditure, household income, household composition and additional household characteristics (age of the members, education, renters versus owners, ...). The housing budget share is computed as the ratio of housing expenditure over household income.

More details on the data sources are given in Appendix A.

2.2 Housing consumption and demographics

As a starting point, we aim at showing that the price of children is tied to housing costs—a crucial aspect for our story to hold. To do so, we quantify the impact of family size on household housing consumption and budget shares (housing consumption expenditures as a share of household income).

Empirical strategy. Using household level data on housing consumption and expenditure at different dates t, we estimate for an household i residing in municipality ℓ_k of urban area k,

$$h_{i,\ell_k,t} = c_{k,t} + f_k\left(d_{\ell_k}\right) + \sum_{m=1}^{N} \beta_m \cdot \mathbb{1}_{\{i \in \mathbb{S}_m\}} + X_{i,\ell_k,t} \cdot \alpha + \nu_{i,t}, \tag{1}$$

where $h_{i,t}$ is the housing consumption variable of interest (floor space, housing budget share) of household i at date t, $c_{k,t}$ is a fixed-effect for the urban area k at date t, $f_k\left(d_{\ell_k}\right)$ is a parametrized function of the distance d_{ℓ_k} of the municipality of residence from the center of the urban area, and $\mathbb{1}_{\{i \in \mathbb{S}_m\}}$ is a collection of dummies by household type $\mathbb{S}_m$ (single with n children, couple with n children for $n \in \{0, 1, 2, ..., 4+\}$, ..) and $X_{i,\ell_k,t}$ are controls for observed characteristics at the household level (age, income, education, renters versus owners,...) at date t.

Urban area fixed-effects together with the distance from the center of the municipality of residence aim at controlling as much as possible for the location of residence. This is quite crucial to quantify the housing costs of children since large families tend to sort in cheaper urban municipalities as shown the later Section 2.3 (see Facts 2 and 3).[11] Controlling for location, estimates of β_n capture the effect of different household sizes on housing consumption.

In our baseline specification, $f_k\left(d_{\ell_k}\right) = \alpha_d \cdot \log(d_{\ell_k})$ and controls include age, education (3 levels), income and a dummy for renters. Sensitivity analysis is relegated to Appendix B.

Housing consumption and budget shares across family size. We estimate Equation 1 on floor space and on the housing budget share. For sake of space, we only report the effect of household size on those variables. Results are summarized in Figures 2 and 3.

[10] The sample size varies from year to year, starting with about $15,000$ households in 1984 and up to $21,000$ in 2006—about $117,000$ observations across the 7 waves. The same selection criteria than for the census are applied to the ENL data.

[11] The sample size does not allow to control for fixed-effects at the municipality level.

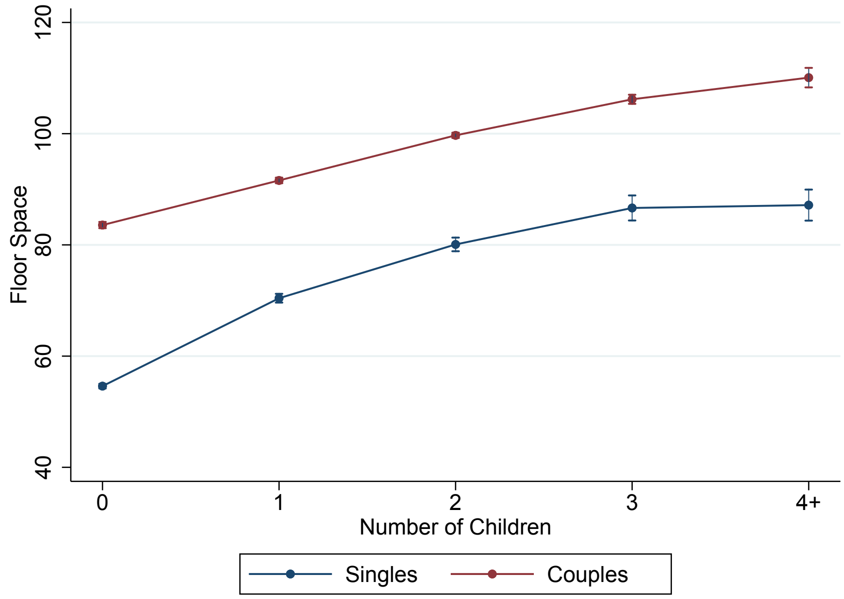


Figure 2: Floor space by household size.

Notes: We report separately results for households with only one adult (singles) or two adults (couples). Estimates are obtained by regressing the floor space at the household level on urban area (UA) dummies, log of distance from the center and a collection of dummies by household types, controlling for age, income, education and renter vs. owner status (Equation 1).

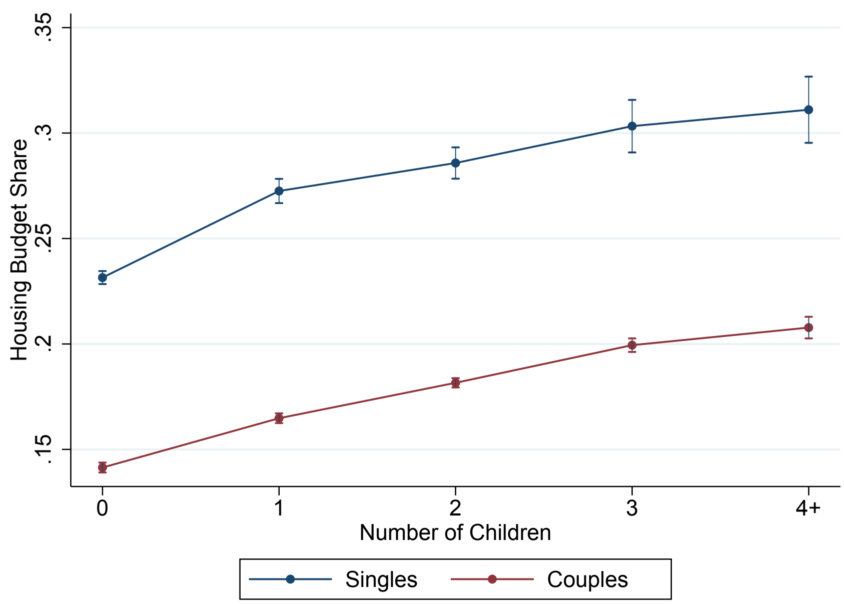


Figure 3: Housing budget share by household size.

Notes: We report separately results for households with only one adult (singles) or two adults (couples). Estimates are obtained by regressing the housing budget share at the household level on urban area (UA) dummies, log of distance from the center and a collection of dummies by household types, controlling for age, income, education and renter vs. owner status (Equation 1).

Fact 1. *Household housing consumption is increasing with the number of children in a given location.*

Households with more children consume more floor space and have a larger budget share of housing for both singles and couples. In terms of magnitude (averaging across couples and singles), having a first child adds about 10 m^2 of floor space (resp. about 3 percentage points to the budget share), the second one adds about 7.5 m^2 (resp. 1.5 p.p), and the third one about 6 m^2 (resp. 1.5 p.p.). Above four children, the additional budget and floor space required becomes less striking. Note that our results are evidence of some economies of scale for housing consumption as function of family size, the first child being more expensive than the second or third ones. In terms of magnitude, in the municipality of Paris, with a price per meter square above 10,000 euros, the first child costs more than $100,000$ euros and the second one adds about $75,000$ euros. This needs to be compared to a smaller municipality of 100,000 inhabitants (e.g. Poitiers or Besançon), where the cost of housing is less $2,500$ euros per meter square and where the housing costs of children are thus significantly lower.

2.3 Urban spatial sorting across demographics

As we have established that children are costly in terms of housing, we now aim at providing novel facts on the spatial sorting across urban locations of households depending on their demographics.

Empirical strategy. Data are the municipality level, with ℓ_k denoting municipality ℓ in urban area k. We perform the following regressions at different census dates (1968, 1975, 1982, 1990, 1999, 2010, 2015),

$$y_{\ell_k} = c_k + f_k\left(d_{\ell_k}\right) + X_{\ell_k} \cdot \alpha + \nu_{\ell_k} \tag{2}$$

where y_{ℓ_k} is a the demographic variable of interest in municipality ℓ_k, c_k is a fixed-effect for the urban area k, $f_k\left(d_{\ell_k}\right)$ is a flexible parametrized function of the distance d_{ℓ_k} of the municipality from the center of the urban area, and X_{ℓ_k} are centered controls for observed characteristics in municipality ℓ_k (amenities, average income, education, ...).

Without loss of generality, $f_k\left(d_{\ell_k}\right)$ is normalized to zero in the most central locations.[12] Thus, the urban area fixed effect c_k measures the variable of interest y_{ℓ_k} at the center of the urban area, while the estimates of f_k, possibly urban area specific, measures how the variable of interest evolves as a function of the distance from the center of the urban area. In other words, the fixed-effect c_k capture the variations across urban areas, while the parametrized function f_k captures variations within urban areas, from central to more suburban municipalities.

Results are shown without controls. Results with controls are relegated to the Appendix B . In our baseline specification, f_k measures the effect of a collection of dummies by distance deciles within the urban area, $f_k\left(d_{\ell_k}\right) = \sum_{m=2}^{10} \beta_m \cdot \mathbb{1}_{\{d_{\ell_k} \in \mathbb{J}_m^k\}}$, with $\mathbb{1}_{\{d_{\ell_k} \in \mathbb{J}_m^k\}}$ a dummy variable equal to

[12] With a continuous measure of distance, d_{ℓ_k}, $f_k\left(0\right) = 0$.

1 if d_{ℓ_k} belongs to the m^{th} decile of distance within the urban area k. Sensitivity regarding the parametrization of f_k is relegated to Appendix B.

Fertility across urban locations. We start by estimating Equation 2 using our measure of fertility at the municipality level, the ratio of children to adults in parental age, as dependent variable.

Fact 2. *In a given urban area, fertility is higher in more suburban locations.*

For each Census wave, Figure 4 shows the average number of children per adult in parental age across deciles of distance from the center of French urban areas. In each year, families are of larger size in more suburban areas. The variations within urban areas are quite large in magnitude: fertility is higher by about 30% in the most suburban locations (relative to the center). Note that, at a given decile of distance, the number of children is falling over time, which mirrors the drop in fertility in France since 1968.

Fact 3. *Across urban areas, central fertility is higher in less populated cities.*

In a second-step, we regress the estimated urban area fixed-effects, c_k against the urban area population (in log) at different Census dates. Results are summarized in Figure 5, which plots, for each Census date, the estimated linear relationship between the average number of children per adults in parental age (at the center) and the population of the urban area. Each estimated slope coefficient is significant at standard confidence intervals.[13] Central fertility is higher in less populated urban areas. This holds across census waves and the magnitude of the cross-sectional dispersion is large—the center of the smallest cities exhibiting a number of children per adult about 25% higher than the center of Paris in recent years, with an even larger difference in earlier years. Note that across Census waves, the number of children per adult in the center of urban areas decreases over time until the early 2000s, but more so in smaller cities.

Due to Facts 2 and 3, fertility exhibits negative density-dependence, whereby more central locations and larger urban areas exhibit lower fertility.

Different mechanisms might explain such a robust pattern in the data. The main focus of our approach hinges on the costs of housing, which increases the price of children in the most expensive locations (Fact 1). Thus, households with a strong preference for a large family move towards municipalities with cheaper housing and, for a given preference for fertility, families ending up in expensive municipalities reduce their family size. Such mechanisms are consistent with Facts 2 and 3, to the extent that suburban locations or smaller cities have lower housing prices as shown in Combes et al. (2018) for France. We provide further suggestive evidence of the price mechanisms by investigating the spatial variations of urban fertility as a function of housing prices, available at the municipality level only for the recent years.The empirical strategy is similar. The number of children per adult (in pa ental age) depends spatially on housing prices within the urban area

[13] See Appendix B for standard errors of the slope coefficients. Scatter plots by census wave of fixed-effects (e.g. fertility at the center) against the log of the population of urban areas are relegated to Appendix B.

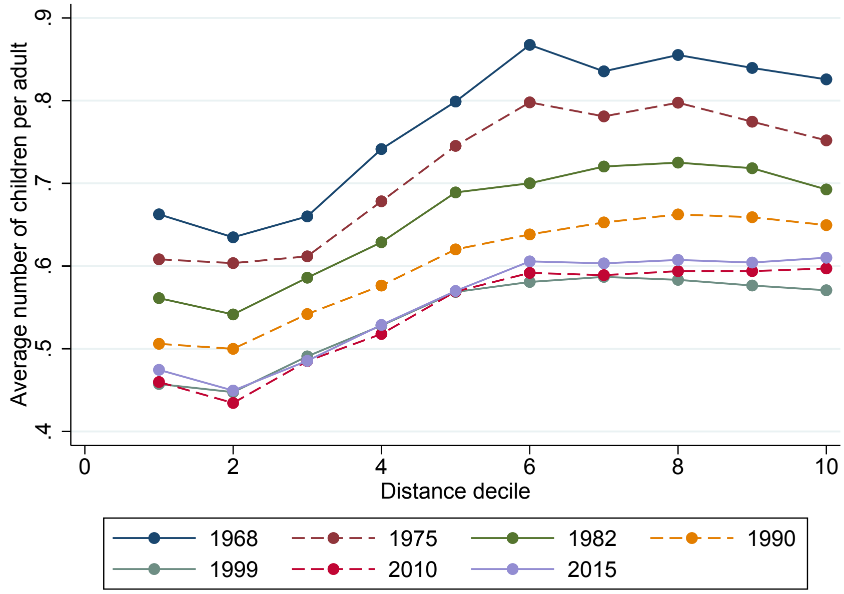


Figure 4: Average number of children per adult in parental age by distance deciles.

Notes: Number of children per adult at different distance deciles are obtained by regressing, at the municipality level, children per adults in parental age on urban area (UA) fixed-effects and dummies by deciles of distance from the center of the UA (Equation 2).

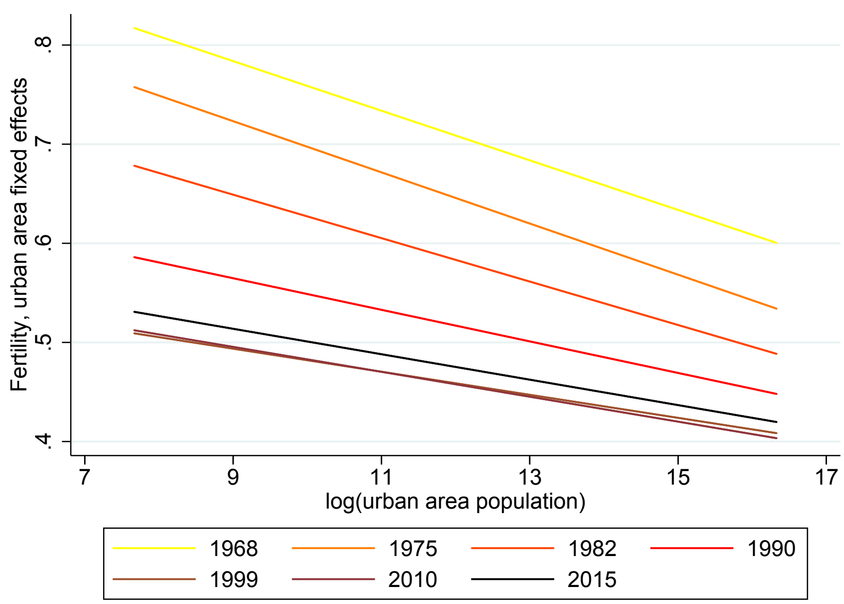


Figure 5: Children per adult in parental age (in the center) and urban area population.

Notes: Each line corresponds to the estimated linear relationship between urban area children per adult fixed effects (c_k) and log of population of the urban area. Urban area children per adult fixed effects (c_k) are obtained by regressing, at the municipality level, children per adults in parental age on urban area (UA) dummies and dummies by deciles of distance from the center of the UA (Equation 2).

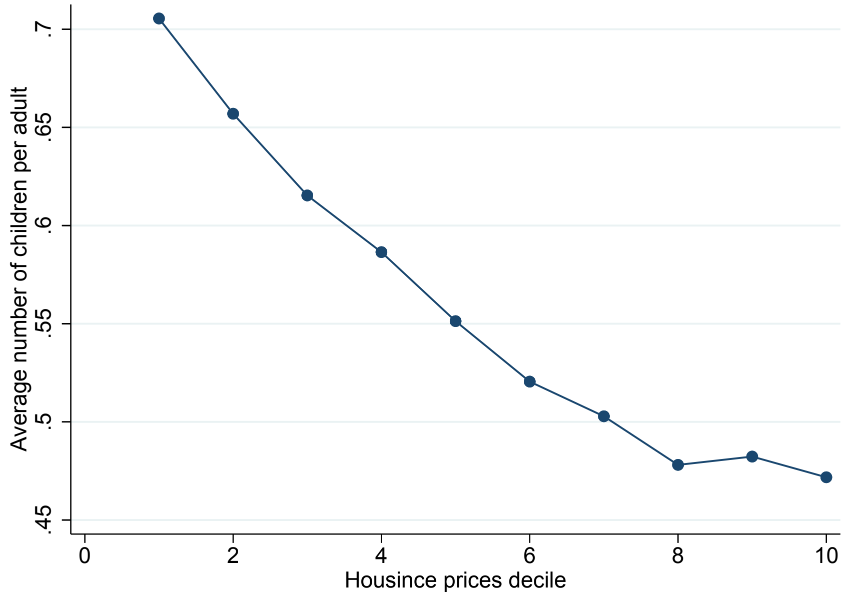


Figure 6: Average number of children per adult in parental age by deciles of housing prices in 2015.
Notes: Number of children per adult at different deciles of housing prices are obtained by regressing, at the municipality level, children per adults in parental age on urban area (UA) fixed-effects and dummies by deciles of housing prices in the UA.

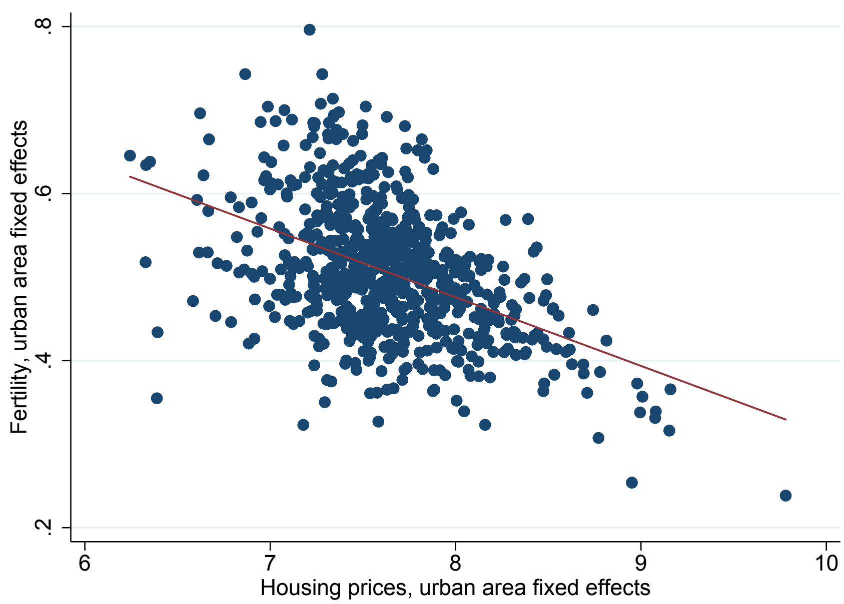


Figure 7: Children per adult in parental age and housing prices (in the center) across urban areas in 2015.
Notes: This plots urban area children per adult fixed effects against urban area housing prices fixed effects. Urban area fixed effects are obtained by regressing, at the municipality level, the variable of interest on urban area (UA) dummies and dummies by deciles of distance for the center of the UA (Equation 2).

using dummies for the deciles of municipality housing prices instead of distance from the center.[14] Figure 6 plots the number of children per adults (in parental age) for the different deciles of housing prices in 2015.[15] Within a city, municipalities with expensive housing prices have significantly fewer children, in line with our price mechanisms.

To explore the relationship between fertility and housing prices across urban areas, we estimate housing prices at the center of each urban area following the same empirical strategy[16] and plot them against the corresponding number of children per adult at the center. This is displayed in Figure 7 which compares the average number of children per adults at the center of a given urban area to the corresponding central housing price for the cross-section of urban areas in 2015. Cheaper urban areas exhibit significantly higher fertility than more expensive ones. Note that the fertility gap between the cheapest and most expensive urban areas is even more pronounced than between the smallest and largest urban areas.

Spatial sorting by age. Following the same strategy, we explore the urban sorting of households depending on their age. We estimate Equation 2 using the average age of adult households (above 18) in each municipality, as dependent variable.

Fact 4. *In a given urban area, the average age of adults increases as we move towards more suburban locations. Across urban areas' centers, the average age of adults is higher in less populated cities.*

Figure 8 plots for each Census date the average age of adult households across deciles of distance from the center. More suburban locations are older on average than more central ones, particularly so in the recent period. In the earlier years, pre-1990, only the most suburban locations are significantly older—with little age difference across the first five deciles of distance. Importantly, at each decile, the average age increases over time, reflecting the aging of the French population since 1968. Aging is particularly visible for suburban municipalities between the fourth and eighth deciles of distance from the center.

Then, we regress the estimated urban area fixed-effects for the average age against the urban area population (in log) at different Census dates. Results are summarized in Figure 9, which plots, for each date, the estimated linear relationship between the average age of adults at the center and the population of the urban area.[17] For all Census waves, larger urban areas are significantly younger, with an age difference of almost 10 years between Paris and a small urban area of 10,000 inhabitants in the recent years. Note that over time, urban areas are aging but mostly the smaller ones—large urban areas stayed relatively young since 1968.

[14] Using the same notations, we perform the following regression, $y_{\ell_k} = c_k + f_k\left(q_{\ell_k}\right) + X_{\ell_k} \cdot \alpha + \nu^q_{\ell_k}$, where the dependent variable y_{ℓ_k} is the number of children per adult in municipality ℓ_k and the function f_k is a collection of dummies by deciles of housing prices, q_{ℓ_k}, within urban area k.

[15] We use the last year of data for housing prices (2012) and census data (2015).

[16] We estimate urban area fixed-effects for housing prices using the log of housing prices as dependent variable in Equation 2.

[17] Each estimated slope coefficient is significant at standard confidence intervals. See Appendix B for standard errors of the slope coefficients. Scatter plots by census wave of fixed-effects (e.g. average age of adult households at the center) against the log of the population of urban areas are relegated to Appendix B.

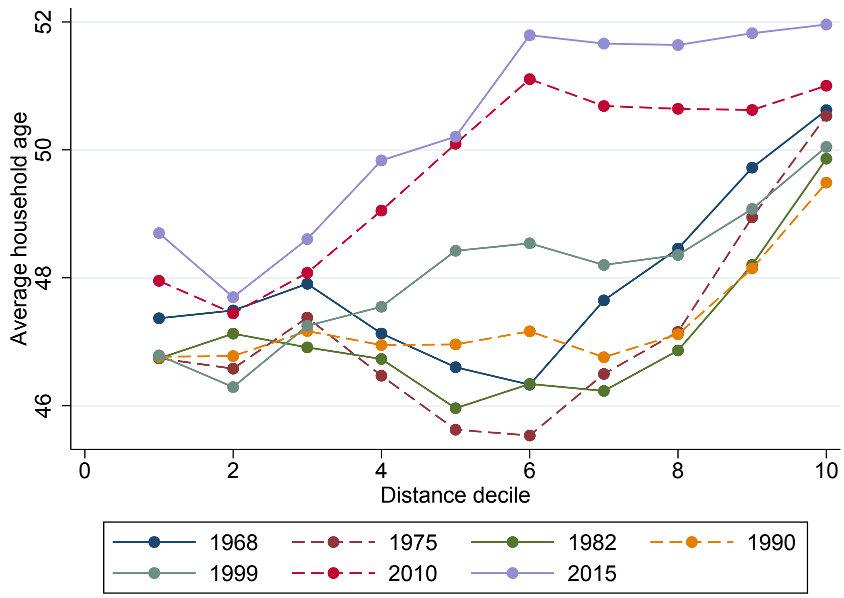


Figure 8: Average age of adults by distance deciles.

Notes: Average age of adults at different distance deciles are obtained by regressing, at the municipality level, the average age of adults on urban area (UA) fixed-effects and dummies by deciles of distance from the center of the UA (Equation 2).

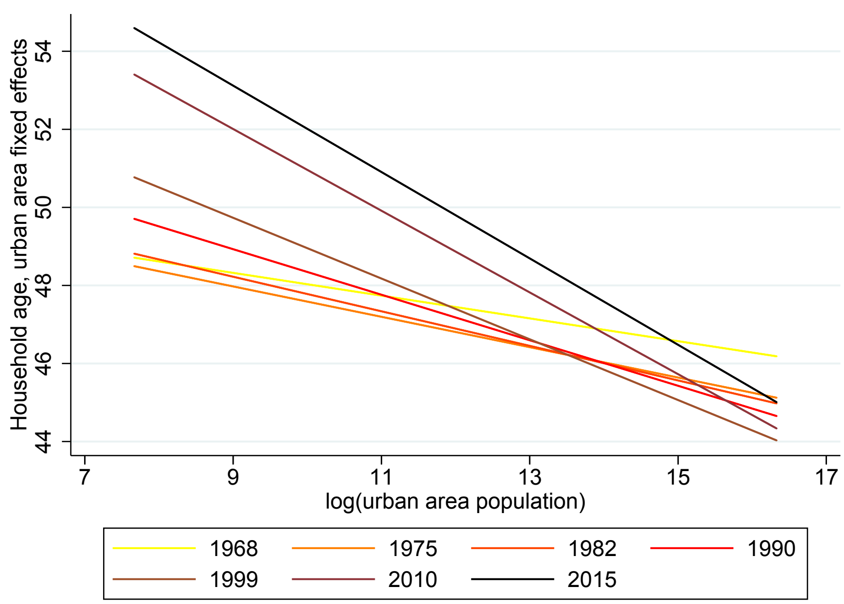


Figure 9: Average age of adults (in the center) and urban area population.

Notes: Each line corresponds to the estimated linear relationship between urban area fixed effects for the average age of adult households (above 18) and log of population of the urban area. Urban area average age of adult households fixed effects (c_k) are obtained by regressing, at the municipality level, the average age of adult households on urban area (UA) dummies and dummies by deciles of distance from the center of the UA (Equation 2).

The age variations across urban locations, within and across urban areas, can be driven by the sorting of younger or older populations (e.g. retirees): young households without children (resp. older households) sorting into more central locations and/or larger cities (resp. more suburban locations and/or smaller cities). To disentangle the importance of young versus older households in explaining the spatial variations in age, we explore these two groups separately. Young (resp. older) adults are within the [18,27[age interval (resp. above 54).

Fact 5. *Within (resp. across) urban areas, young adults sort into central locations (resp. larger cities). Older adults sort into smaller urban areas.*

Using the same empirical methodology, we find that both groups contribute to the age variations across urban locations. Young adults are disproportionately located in the center of urban areas, and particularly so in larger cities. Within urban areas, the number of young per adult is about three times larger in the center than in the furthest away suburbs. Across urban areas, the gap between Paris and the smallest urban areas is of a similar magnitude. For sake of space, Figures summarizing these results are relegated to Appendix B (see Figures B.7 and B.9). For older adults (above 54), the sorting is very sharp across urban areas—older adults are disproportionately residing in smaller urban areas. Within urban areas, the patterns of sorting are more ambiguous: in the earlier years, until the nineties, older adults tend to favor the further away municipalities and, to a lower extent, the most central ones—neglecting suburban locations close to the median in terms distance from the center. In the recent years, older adults tend to spread more equally across suburban locations. More specifically, older household became in the recent years much more present in suburban locations between the fourth and eighth deciles of distance from the center—largely explaining the aging of these locations (see Figures B.8 and B.10 in Appendix B).

Thus, given Facts 2 and 3, elderly compete with families of larger size for the cheaper locations, while the young do not. With falling fertility and rising longevity, the French population aged sharply over the last decades and the share of population above 54 increased significantly, by about 40% since the sixties. As a consequence, competition for locations in cheap areas (smaller cities and suburbs) should have toughened—possibly crowding out families from such locations.

To summarize, we provide suggestive empirical evidence of our housing price mechanism in driving fertility decisions, whereby higher costs of housing might deter families from having more children. First, housing consumption is increasing with family size, tying the price of children to the price of housing. Second, adults in parental age have fewer children in more expensive municipalities—central locations and larger cities. This could be due to the sorting into cheaper locations of households with a preference for larger families or to the decision of reducing family size in the most expensive locations. Both channels require families to react to the costs of housing when deciding on their household size. These broad facts motivate the theory developed in the following Section 3, where location and fertility decisions are jointly determined. While the empirical evidence is not causal, the quantitative nature of the theory will help to disentangle and quantify the mechanisms at play.

3 Model

3.1 Set-up

The economy is populated by overlapping generations of agents living at most until age $\bar{a}$. Time t and age a are discrete. Time t corresponds to different generations of households (see Figure 10). For ease of notations, we abstract from time indices for now. We solve first the model across space focusing on a cross-section of individuals of different ages at a given date t. The dynamics across time is developed in Section 3.3.

Stages of life. Individuals go through four stages of life, children (c), **young** (y), **parents** (p) and old (o). They enter the young stage at age a_y, the parental stage at age a_p and the old stage at age a_o, with $0 < a_y < a_p < a_o \leq \bar{a}$. Children do not make decisions and are sheltered by their parents. After forming a household of size $\mathcal{N}$, young individuals choose where to live, provide labor and consume. When parents, households decide on their number of children n at age a_p, choose their residential location, provide labor and consume. When old, households living without their children decide where to live and receive labor income when active or pension benefits when retired. The retirement age is denoted a_r, with $a_o \leq a_r \leq \bar{a}$. Figure 10 summarizes the different stages of life.

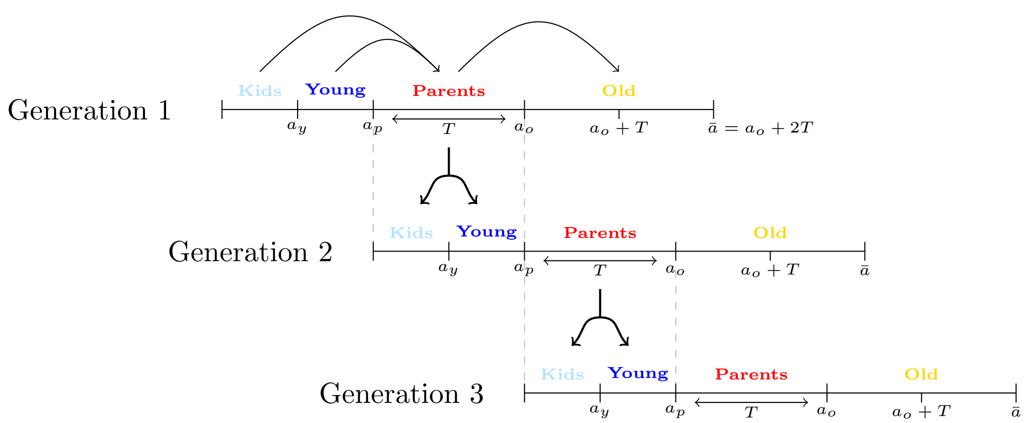


Figure 10: Timeline.

Notes: A each date t, the population is made of a new generation of households in each of the life-stages. The duration T of a period in years (i.e. the duration of a generation) corresponds to the duration of the parental stage. Notice the color coding children (kids), **young** , **parents** and old.

The total number of adult households, L, is composed of the number of young, parent and old households, $L = L_y + L_p + L_o$.

Spatial Structure and Household Income. The economy consists of a fixed number K of cities. A city $k \in \{1, ..., K\}$ is made of a fixed number of locations $\mathcal{L}_k$, where each location in city k is denoted $\ell_k \in \{1, ..., \mathcal{L}_k\}$. $\mathcal{L}$ denotes the entire set of $K \times \mathcal{L}_k$ possible locations.

During working-age, $a_y \leq a < a_r$, a household of age a residing in location ℓ_k of city k earns a labor income

$$y(a, \ell_k) = \theta_k \cdot w(a, \ell_k),$$

where θ_k is a city-specific productivity parameter and $w(a, \ell_k)$ is decreasing with ℓ_k within a city k at all working ages, $w(a, \ell_k) > w(a, \ell_{k'})$ for $k < k'$. Locations in each city are thus ordered by income, from the highest to the lowest. Household income can be interpreted as net of commuting costs. Commuting costs are increasing for further away locations—locations being ordered in terms of commuting distance. This can also reflect the higher productivity of more central locations within the city.[18]

At retirement, $a \geq a_r$, household income is made of retirement benefits $b(a)$ independent on the location. Alternatively, household income at age $a \geq a_y$ can be rewritten,

$$y(a, \ell_k) = \theta_k \cdot w(a, \ell_k) + b(a),$$

where $w(a, \ell_k) = 0$ for $a_r \leq a \leq \bar{a}$ and $b(a) = 0$ for for $a_y \leq a < a_r$. Importantly, contrary to workers, retirees do not benefit through higher wages from living in more productive cities.

Budget constraints. Denote n, $n \in \{0, 1, ..., \bar{n}\}$, the number of children residing in the household. By assumption, $n = 0$ in the young and old stages. At each age a, households residing with n children located in ℓ_k consume $c(a, \ell_k, n)$ out of their income $y(a, \ell_k)$. In each location, households must also finance housing out of their income. Housing space, $h(\mathcal{N}+n)$, is increasing with household size $\mathcal{N} + n$. Households face the following budget constraint,

$$c(a, \ell_k, n) + q_{\ell_k} h(\mathcal{N} + n) = y(a, \ell_k), \tag{3}$$

where q_{ℓ_k} denotes the price of a housing unit in location ℓ_k.

Preferences. At each age a, instantaneous utility, $U(a, \ell_k, n)$ depends on a city amenity A_k, consumption $c(a, \ell_k, n)$ and household specific preferences for location at any age and for fertility at age a_p.

For age $a \neq a_p$, the preference shock for location, denoted ε_{ℓ_k}, is drawn from a type 1 Extreme Value distribution with scale parameter σ. Thus, for age $a \neq a_p$,

$$U(a, \ell_k, n) = A_k + u\left(c(a, \ell_k, n)\right) + \sigma\varepsilon_{\ell_k}.$$

When entering the parental stage, at age $a = a_p$, households face a preference shock for both location ℓ_k and fertility n, ε_{n,ℓ_k}, also drawn from a type 1 Extreme Value distribution with scale parameter σ.[19] Additionally, parents receive utility benefits $v(n) \geq 0$ of having n children, $n \in \{0, 1, ..., \bar{n}\}$,

[18] Household income in the parental age is also net of contribution taxes to finance retirement benefits in old age.

[19] In a structural sense, $\{\varepsilon_{\ell_k}\}_{a \neq a_p}$ and ε_{n,ℓ_k} for $a = a_p$ are information available to each household, but not to the econometrician, and the scale σ can be estimated from data. Note that one could assume different scale parameters σ_a at each age a, $\sigma_a = \sigma$ in our baseline.

with $v(0) = 0$. Hence, for age $a = a_p$,

$$U(a_p, \ell_k, n) = A_k + u\left(c(a_p, \ell_k, n)\right) + v(n) + \sigma\varepsilon_{n,\ell_k}.$$

We turn to the location and fertility choices of households.

3.2 Spatial Equilibrium

We solve for a spatial equilibrium at a given date by deriving the location decisions of each household together with the fertility decision at age $a = a_p$. Decisions are independent from each other at each age under the following additional assumptions,

Assumption 1. *Households can freely change locations at each age and the parental stage starting at age a_p lasts only one period.*

which make sure that location decisions at each age are independent of future housing costs and that fertility decisions at age a_p do not affect location decisions when older. Assumption 1 also guarantees that children stay in the household only one period. Thus, only current housing costs matter for fertility choices. Assumption 1 greatly simplifies the analysis since decisions at each age are the result of a static optimization where each household maximizes its instantaneous utility. Note that under Assumption 1, households only reside with their children at age a_p, implying $n = 0$ for all ages $a \neq a_p$.

Consumer Problem and Optimal Choice. At age $a \neq a_p$, using the budget constraint (Eq. 3) and Assumption 1, the household solves the following discrete choice location problem,

$$\max_{\ell_k \in \mathcal{L}} \{A_k + u\left(y(a, \ell_k) - q_{\ell_k} h(\mathcal{N})\right) + \sigma\varepsilon_{\ell_k}\}.$$

This gives the following choice probabilities for age $a \neq a_p$,

$$P_a(\ell_k) = \frac{\exp\left(\frac{A_k + u\left(y(a,\ell_k) - q_{\ell_k} h(\mathcal{N})\right)}{\sigma}\right)}{\sum_{\ell_s \in \mathcal{L}} \exp\left(\frac{A_s + u\left(y(a,\ell_s) - q_{\ell_s} h(\mathcal{N})\right)}{\sigma}\right)}. \tag{4}$$

Relative to alternative locations, the choice of a given location ℓ_k is favored by the amenity A_k provided by the corresponding city and by the income generated in the location $y(a, \ell_k)$. High housing costs q_{ℓ_k} makes it less likely to choose the location.

At age $a = a_p$, the household chooses fertility n and location ℓ_k. Using the budget constraint (Eq. 3), this leads to the following discrete choice problem,

$$\max_{\ell_k \in \mathcal{L}, n \in \{0,1,\ldots,\bar{n}\}} \{A_k + u\left(y(a, \ell_k) - q_{\ell_k} h(\mathcal{N} + n)\right) + v(n) + \sigma\varepsilon_{n,\ell_k}\}.$$

This leads to the following unconditional choice probability for each combination (ℓ_k, n),

$$P_{a_p}(\ell_k, n) = \frac{\exp\left(\frac{A_k + u\left(y(a_p,\ell_k) - q_{\ell_k} h(\mathcal{N}+n)\right) + v(n)}{\sigma}\right)}{\sum_{\ell_s \in \mathcal{L}} \sum_{m=0}^{\bar{n}} \exp\left(\frac{A_s + u(y(a_p,\ell_s) - q_{\ell_s} h(\mathcal{N}+m)) + v(m)}{\sigma}\right)}. \quad (5)$$

Beside the role of amenity, income and housing costs in guiding the location choice, one should notice that high housing costs reduce fertility in a given location since a larger household size requires more housing space—$h(\mathcal{N} + n)$ increases with n.

The unconditional probabilities of location and child choices are easily obtained,

$$P_{a_p}(\ell_k) = \sum_{n=0}^{\bar{n}} P_{a_p}(\ell_k, n)$$

and

$$P_{a_p}(n) = \sum_{\ell_k \in \mathcal{L}} P_{a_p}(\ell_k, n).$$

The probability of choosing n children, conditionally on residing in location ℓ_k is

$$P_{a_p}(n|\ell_k) = \frac{P_{a_p}(\ell_k, n)}{P_{a_p}(\ell_k)}.$$

From Equations 4 and 5, one can also deduce the average fertility, n_{ℓ_k}, in a given location ℓ_k,

$$n_{\ell_k} = \sum_{m=0}^{\bar{n}} m \times P_{a_p}(\ell_k, m),$$

and the population by age at each location,

$$L_a(\ell_k) = P_a(\ell_k) \cdot L_a,$$

with L_a the number of households of age a.

Housing Market Clearing. Housing supply $H_s(\ell_k)$ in each location ℓ_k is increasing with the housing price with a constant elasticity ρ,

$$H_s(\ell_k) = \delta_k (q_{\ell_k})^{\rho}, \quad (6)$$

where δ_k denotes a city-specific supply-shifter. Housing demand in each location is determined by the housing space, adjusted for household size, that is required by the households of different ages,

$$H_d(\ell_k) = L_p \left(\sum_{n=0}^{\bar{n}} P_{a_p}(\ell_k, n) h(\mathcal{N} + n) \right) + \sum_{a \neq a_p} L_a P_a(\ell_k) h(\mathcal{N}), \tag{7}$$

The first term in Eq. 7 represents the housing demand by parents and the second term the housing demand of households in young and old stages of life. Under Assumption 1 and for realistic life stages, $L_p = L_{a_p}$, $L_y = L_{a_y}$ and $L_o = \sum_{a=a_o}^{\bar{a}} L_a$.

Housing prices q_{ℓ_k} are determined such that housing market clears in all location ℓ_k,

$$H_d(\ell_k) = H_s(\ell_k). \tag{8}$$

The housing market clearing conditions, Equation 8, together with choice probabilities at the different ages, Equations 4 and 5, pin down the spatial equilibrium for a given demographic composition of adult households, $\{L_a\}_{a_y \leq a \leq \bar{a}}$. This determines the price of housing together with the distribution of age-groups and number of children in every location as described in Definition 1.

Definition 1. *For a given set of city/location characteristics, potentially age-specific,* $\{A_k, \theta_k, \delta_k, w(a, \ell_k), b(a)\}_{k \times \ell_k \in \{1,\ldots,K\} \times \{1,\ldots,\mathcal{L}_k\}, a_y \leq a \leq \bar{a}}$ *and a given aggregate demographic composition of adult households,* $\{L_a\}_{a_y \leq a \leq \bar{a}}$*, a static spatial equilibrium is a vector of housing costs,* $\{q_{\ell_k}\}_{\ell_k \in \mathcal{L}}$*, demographic composition,* $\{L_{a,\ell_k}\}_{\ell_k \in \mathcal{L}, a_y \leq a \leq \bar{a}}$*, and average fertility,* $\{n_{\ell_k}\}_{\ell_k \in \mathcal{L}}$*, in each location such that:*

- *Location decisions at each age a and fertility decision at age a_p satisfy the choice probabilities, Equations 4 and 5.*
- *The housing market clears in each and every location $\ell_k \in \mathcal{L}$, Equation 8.*

3.3 Dynamics across time

The dynamics across time of the model is driven by the endogenous evolution of fertility together with the time-variation of exogenous structural parameters (e.g. city-specific amenities and productivities, location-specific household incomes, housing supply-shifters, ...). More specifically, changes in fertility decisions for different cohorts of parents feed into demographics changes, which in turn impact future housing demand and housing prices. The model's dynamics across time is made of a sequence of static spatial equilibria as defined in Section 3.2, where each each date t corresponds to a new generation of households (see Figure 10). The sequence of equilibria depends on the time evolution of exogenous parameters, $\{A_{k,t}, \theta_{k,t}, \delta_{k,t}, w_t, b_t\}_{t \geq 0}$, to accommodate for changes in cities characteristics and household incomes. The demographic composition is given at date $t = 0$ and evolves endogenously in the subsequent periods as detailed below.

Timing. Under Assumption 1, a period lasts the parental stage, with $T = a_o - a_p$ the duration of a period in years (i.e. the duration of a generation). Ages are expressed in years and dates are integer, $t \in \{0, 1, 2, ...\}$. Without loss of generality, for ages above a_p, the duration in between two ages is constant equal to T, $a \in \{a_p, a_p + T, a_p + 2T, ..., \bar{a}\}$. Below age a_p, the duration of the child stage, a_y, and the duration of the young stage, $a_p - a_y$, can be different integers under the following Assumption 2,

Assumption 2. *The child and the young stages last at most one period, $a_y \leq T$ and $a_p - a_y \leq T$.*

which will imply that the population of children and young households will only depend on the fertility of current and past cohorts of parents.

Demographics. Individuals survive with certainty until the parental stage, $a = a_p$. At date t, after the parental stage, at age $a \geq a_p$, households face a certain probability of surviving into the next age, $s_{a,t} \in (0, 1)$, thus until age $\bar{a}$ ($s_{\bar{a}} = 0$ at all dates). From one period to the next, the survival probabilities pin down the future population of the old age groups, $a \geq a_o$ for a given demographic composition of parents and old households. For $a \in \{a_p, ..., \bar{a} - T\}$,

$$L_{a+T,t+1} = s_{a,t} \cdot L_{a,t}. \tag{9}$$

The number of children at each date and the number of individuals entering the young stage are both determined by the endogenous fertility choices made by parental households. It is useful to introduce the average number of children born in a parental household at a given date t by summing across locations,

$$E_t[n] = \sum_{\ell_k \in \mathcal{L}} n_{\ell_k,t} = \sum_{\ell_k \in \mathcal{L}} \sum_{m=0}^{\bar{n}} m \times P_{a_p,t}(\ell_k, m), \tag{10}$$

where the probability $P_{a_p,t}(\ell_k, n)$ at date t is defined in Eq. 5—amenities, $A_{k,t}$, income $y_t(a_p, \ell_k)$ and housing prices, $q_{\ell_k,t}$ depending on date t. Thus, the total number of children of the current (resp. past) cohort of parents is $L_{p,t}E_t[n]$ (resp. $L_{p,t-1}E_{t-1}[n]$)).

Children of different cohorts of parents are distributed proportionately to population of children and young households, according to the respective duration of these life stages. The total number of children adjusted for household size at period t is,

$$L_{c,t} = \frac{a_y}{T} \frac{L_{p,t}}{\mathcal{N}} E_t[n], \tag{11}$$

and the total number of households in the young stage is,

$$L_{y,t} = \left(1 - \frac{a_y}{T}\right) \frac{L_{p,t}}{\mathcal{N}} E_t[n] + \frac{a_p - T}{a_y} L_{c,t-1}, \tag{12}$$

where the first term represents the oldest children of the current cohort of parents, while the second term represents the youngest children of the past cohorts of parents. Lastly, the number of households in the parental stage is,

$$L_{p,t} = \left(1 - \frac{a_p - T}{a_y}\right) L_{c,t-1} + L_{y,t-1}, \tag{13}$$

where the first term represents the oldest children of the previous cohort of parents and the second term, the young households turning into parents.

Dynamics. For a given initial age-distribution of adult households, Equations 9, 11, 12 and 13 summarize the dynamics of the demographic composition across periods, where the average parental fertility, $E_t[n]$, is given by Eqs. 5 and 10 together with the vector of housing prices across locations clearing the market (Eq. 8). The dynamics of the model is a sequence of static spatial equilibria at each date t given the endogenous demographic composition of the population as described in Definition 2.

Definition 2. *For time-varying city/location characteristics, potentially age-specific, $\{A_{k,t}, \theta_{k,t}, \delta_{k,t}, w_t(a, \ell_k), b_t(a)\}_{k \times \ell_k \in \{1,...,K\} \times \{1,...,\mathcal{L}_k\}, a_y \leq a \leq \bar{a}}$ and a given initial aggregate demographic composition of adult households, $\{L_{a,0}\}_{a_y \leq a \leq \bar{a}}$, a sequence of equilibria for $t \geq 0$ is defined recursively such that:*

- *The equilibrium at each date t is a static spatial equilibrium according to Definition 1 for a given distribution by age of the aggregate population, $\{L_{a,t}\}_{a_y \leq a \leq \bar{a}}$.*
- *Aggregate population dynamics by age at each date t, $\{L_{a,t}\}_{a_y \leq a \leq \bar{a}}$, is defined by Equations 9, 11, 12 and 13, where the average parental fertility, $E_t[n]$ is determined in the static equilibrium at date t.*

4 Estimation across French urban areas

Work on full estimation is ongoing. We present several results from the model with the aim to reproduce qualitatively the facts from section 2.

4.1 Simulation Setup

- **Timing.** One period = one generation. $t = 0$ corrresponds to 1950.
  Life-stages: Age $a_y = 18$, $a_p = 27$, $a_o = 54$.
- **Spatial structure.** $K = 5$ cities. $\mathcal{L}_k = 5$ locations in each city.
- **Productivity and amenity.** Cities differ only in their productivity level θ_k and amenity A_k. Positive correlation between productivity and amenity. No change over time for simplicity.
- **Income and commuting costs.** Constant wage within city. Pensions = 80% of average national income in all locations. Commuting costs $\tau_a(\ell_k - 1)$ increasing linearly across locations $\ell_k \in \{1, 2, ..., 5\}$ from center to fringe. Identical for young and parents, **lower for old**.

- **Housing supply.** Constant elasticity of supply of housing $\rho = 2$. No difference in housing supply across locations, $\delta_{\ell_k} = \delta$ constant across locations.
- **Housing space by household size.** Housing space, $h(\mathcal{N} + n) = \underline{h} \cdot (\mathcal{N} + n)^{\alpha}$. $\underline{h}$ set to match aggregate housing spending share. $\alpha = 0.34$ related to scale economies in household size, set to match roughly data of **Fact 1**.
- **Preferences.** Linear consumption. No income effect on fertility.
  Fertility preferences linear in number of children $= \nu n +$ preference shifter for $n = 2$. Set to generate aggregate fertility slightly above 2 at $t = -1$ and reasonable distribution of parental household size.
  Preference shock for location/fertility with scale parameter $\sigma = 0.5$.
- **Mortality and initial demographic composition.** Survival probablity for old in line with data in 1950. Corresponding initial age distribution at $t = -1$.

We generate aggregate demographic changes by affecting fertility preferences and by modifying the longevity parameters, such that the model generates patterns roughly in line with the data:

- **Baby-boom.** Fertility preference shifter $\Delta_t \nu$ in period $t \in \{0, 1, 2\}$, with $\Delta_0 \nu > \Delta_1 \nu > \Delta_2 \nu > 0$.
  Magnitude to roughly match the increase in fertility during the baby-boom in France. Progressive phasing-out.
- **Rising longevity.** Increase in survival probabilities at older ages in line with data. Probablity to survive into old age, above 54 (resp. very old age, above 81) increases from 0.5 to 0.7 (resp. 0.04 to 0.3) between $t = 0$ and today.

4.1.1 Aggregate demographic changes: Population dynamics

In figure 11 we provide an aggregate overview over the dynamics after the 1950's baby boom. As mentioned above, the simulation increases the fertility preference ν for three periods starting in 1950, reverting back to the baseline value. The evolution of aggregate fertility in the central panel of figure 11 is thus partly driven by preference, but after 2004 purely a dynamic consequence of the initial shock. The mechanism is visible in the left panel, where we see that average housing prices in the economy closely track aggregate population, which increases due to higher fertility, as well as increases in longevity of parents and old generations. The right most panel illustrates the sorting effect of the baby boom and aging of the economy, which goes together with spatial sorting. First, we see the aggregate share of old in the economy rise as the baby boomer generation generation (a large cohort born around 1950) starts to enter old age. There is a pure quantity effect (greater population, greater house prices), as well as a displacement effect, as the old compete with parents for cheap housing space. The sorting of parents into more expensive locations reinforces the fall in fertility.

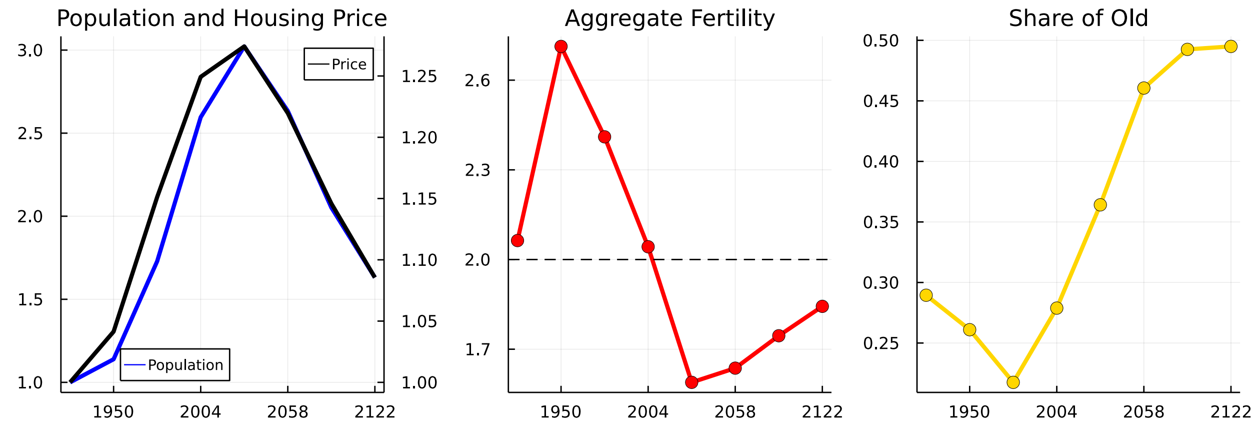
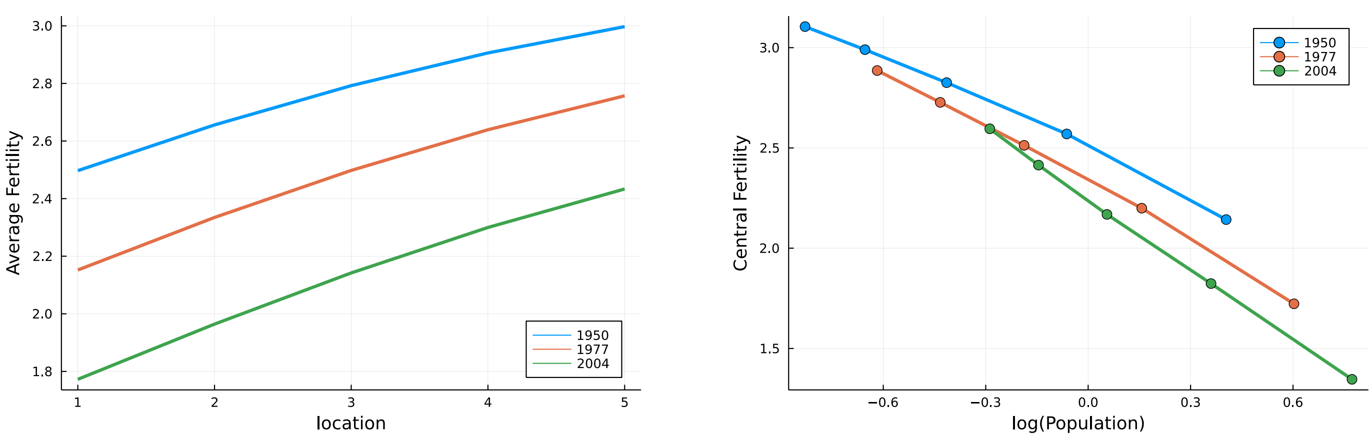


Figure 11: Aggregate Population Dynamics replicating the 1950s Baby Boom and ensuing Baby Bust.

In the following subsections we illustrate that the model is able to replicate the qualitative nature of the stylized facts outlined in section 2, oftentimes matching even the quantitative features fairly well. We put forward one caveat, which is that under the current parameterization the old generation dislikes the center of our big cities too much, resulting in too low a share of old in large cities' centers. We can enrich the model in various dimensions in order to address this issue.

4.1.2 Fertility Across Urban Locations

Figure 12 shows fertility within and across cities for various points in time in the model.



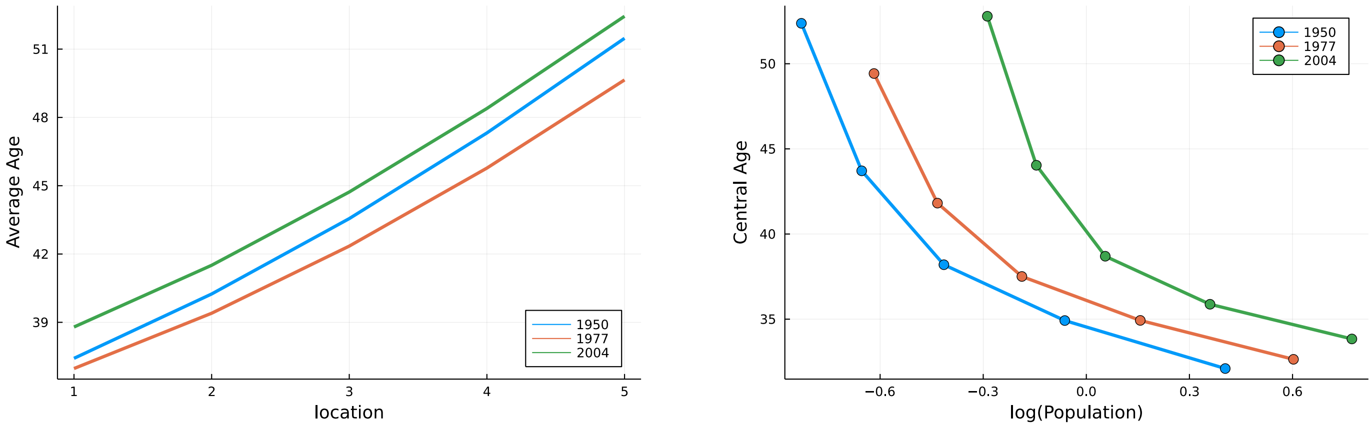
(a) Fact 2: Within city

(b) Fact 3: Across cities

Figure 12: Simulation Results replicating stylized facts number 2 and 3, relating to average fertility within and across urban areas.

4.1.3 Average Age across Urban Locations

Figure 13 shows average age within and across cities for various points in time in the model.



(a) Fact 4: Within city

(b) Fact 4: Across cities

Figure 13: Simulation Results replicating stylized fact 4, relating to average age within and across urban areas.

4.1.4 Spatial sorting by age

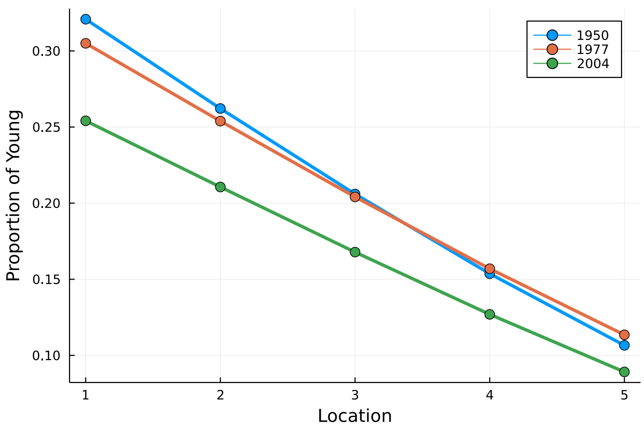
Figure 14 shows sorting by age within city, while 15 shows sorting across centers of urban areas by age groups.

4.1.5 Spatial Distribution of Population

Next we decompose the baby boom by size of cities, comparing the two smallest with the two largest cities by population. In figure 16 we see in the left panel how the model tends to converge to an aggregate fertility rate of 2 (i.e. equal to the replacement rate) in the long run, whereby larger cities converge to a lower level of fertility than smaller ones, their weighted average yielding an economy-wide fertility rate of 2 in the long run. The population distribution is stabilized by migration flows between cities. In the central panel we observe that the baby boom together with increased longevity benefits population growth in smaller cities: parents of large families and old people prefer those locations. Finally, the right most panel illustrates how fertility and suburbanisation co-move. We observe first that the baby boom triggers suburbanisation in large cities, as higher fertility preferences move parents towards their suburbs *as well as* to the centers of smaller cities. Centers of small cities are desirable for parents because of higher income net of housing costs. Second, aging of the economy triggers further suburbanisation, this time more pronounced in the suburbs of smaller cities, where the old locate primarily.

In this context it is also interesting to isolate the labor force from older people, who do no longer need to commute. The associated results in figure 17 indicate in the initial period (until 2004), the relative increase in labor force of small cities is faster, in line with the overall population pattern

(a) Fact 5: Young within city



(b) Fact 5: Old within city

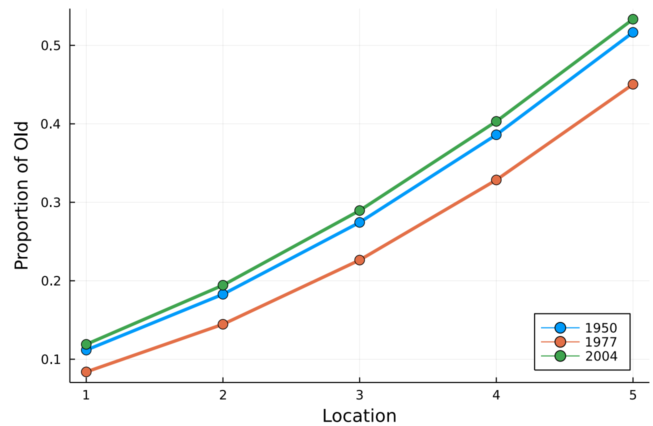
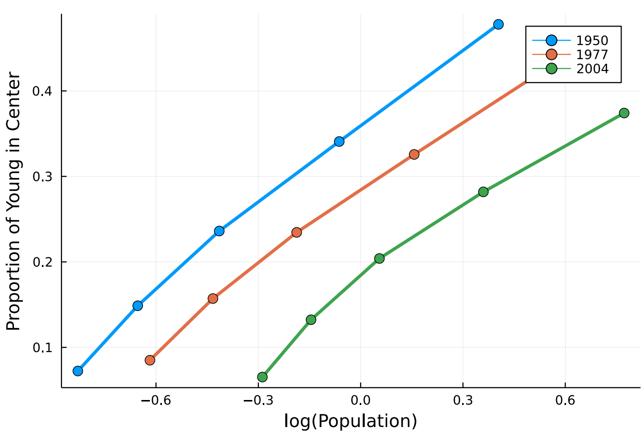


Figure 14: Simulation Results replicating stylized fact 5, relating to sorting by age within urban areas.

(a) Fact 5: Young across cities



(b) Fact 5: Old across cities

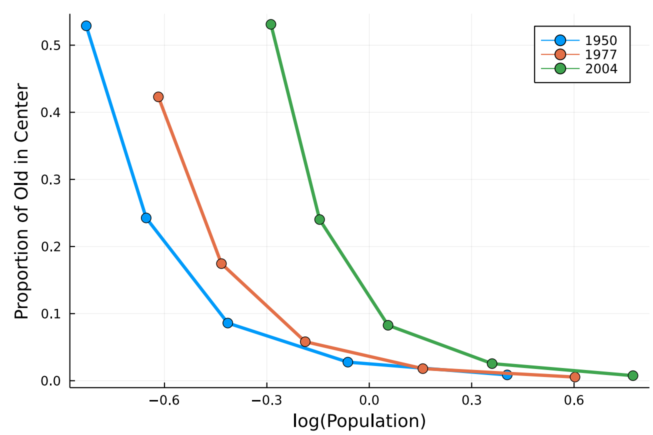


Figure 15: Simulation Results replicating stylized fact 5, relating to sorting by age across centers of urban areas.

observed in figure 16. However, as the effects of again kick in, we see a reversal of this trend, hence, growth of workers is significantly larger in bigger cities, as smaller towns are increasingly occupied by older people. The central panel of figure 17 makes this point also within cities, i.e. we see that the share of workers in the center of both small and large cities initially drops during the baby boom years, but then there is a move back into the center of both small and large cities, for the same reason as above: congestion in the housing market caused by older people at the fringe (particular of small cities).

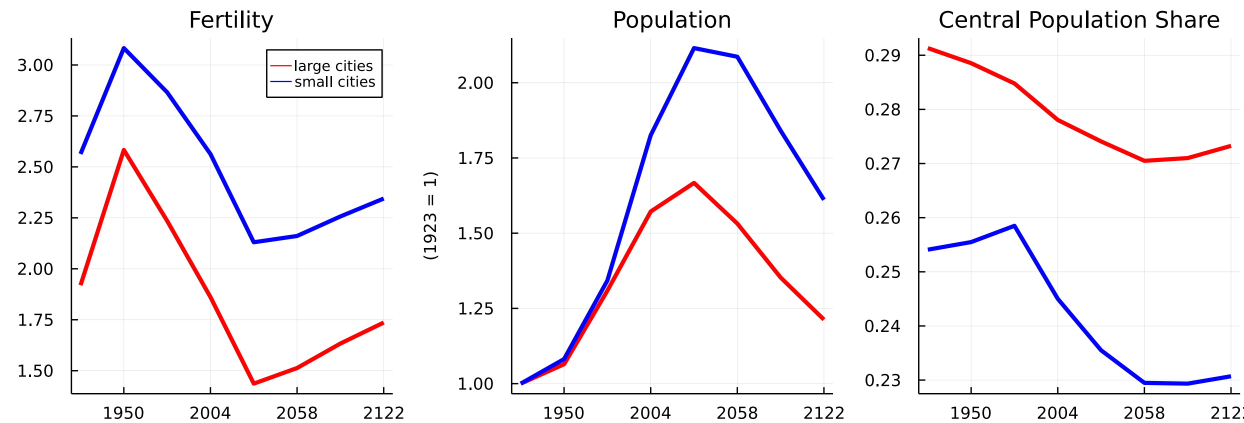


Figure 16: Spatial Population distribution after baby boom episode.

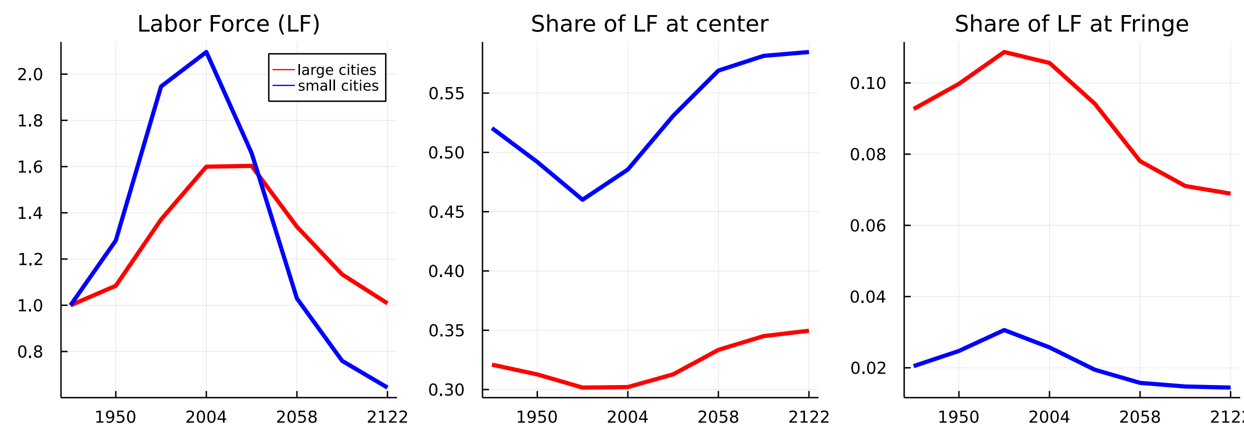


Figure 17: Spatial labor force distribution after baby boom episode.

4.2 Counterfactuals

We perform a series of counterfactual exercises with the calibrated model. First, in order to generate intuition and explain the model mechanism further, we undo the baby boom and the increase in longevity, respectively. Then we move on to policy relevant experiments where we simulate the effects of a decrease in commuting cost, most notably through the wide diffusion of car ownership (but we can draw similar conclusions thinking about remote work during the COVID pandemic, for example), and of changes in housing supply regulations on the dyamics of aggregate fertility.

4.2.1 The Role of the Baby Boom

In this subsection we turn off the shock to fertility preferences, which is responsible to trigger the baby boom and thus observed fertility dynamics. In figure 18, we observe that the boom and bust periods are absent in the counterfactual (left panel), and that population growth happens much later and to a lesser degree. Given that fertility is falling over time, we conclude that population growth is solely a result of increasing longevity. In the right panel we observe that, because aggregate population is lower, and hence since housing costs are lower, the share of the population living in the center is larger than in the baseline.

In terms of sorting behaviour of the working population, figure 19 shows in the left panel that the strong growth of labor force in small cities is undone by the absence of the baby boom, and that aging means the workers' share in the center is increasing over time in both small and larger cities.

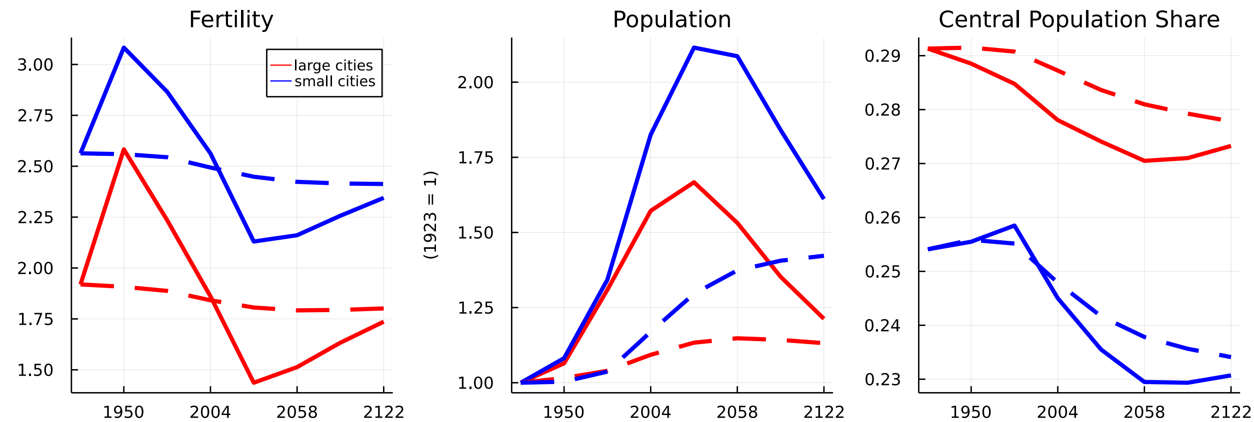


Figure 18: Aggregate implications of undoing the baby boom. The dashed line illustrates the counterfactual.

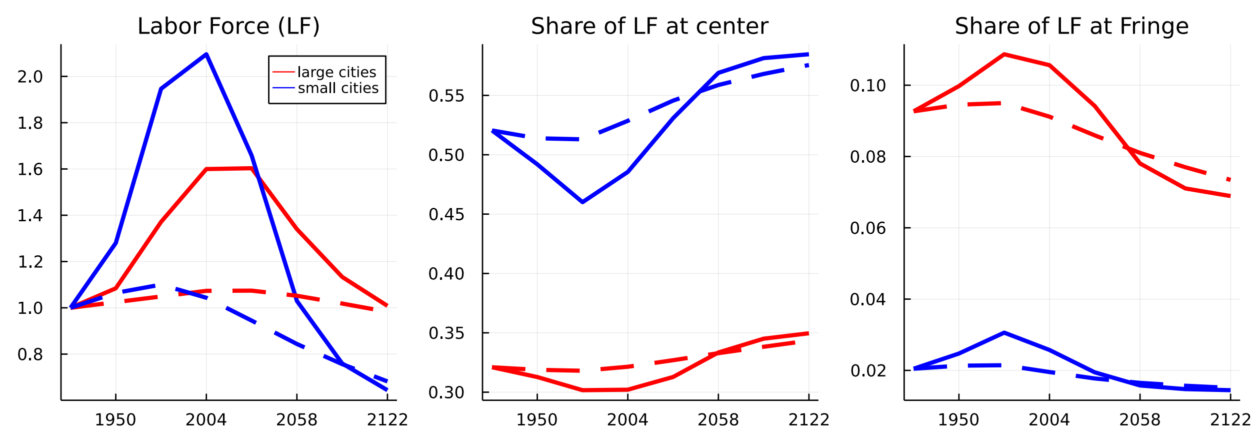


Figure 19: Implications of Labor Force sorting of undoing the baby boom. The dashed line illustrates the counterfactual.

4.2.2 Increasing Longevity

Recall that the baseline model features increasing survival probabilities for parents into old age and from old to very old age in line with the data. We want to investigate here the effects of shutting down this channel on aggregate dynamics. Notice that we keep the baby boom active for this experiment. The first order effect here is to reduce aging in the economy. Looking at figure 20, we recognize again the characteristic shape of aggregate fertility in the left panel, displaying baby boom and ensuing baby bust. Notice that the fall fertility during the bust is less pronounced in the counterfactual (dashed line), driven by the fact that fewer old people are present, reducing housing costs, hence allowing higher fertility. The smaller fall in fertility, however, is not enough to make up for the low survival of parents into old age, and we see in the central panel that population growth is much attenuated with respect to the baseline. In the right panel, finally, we see that while increased again led to strong forces pushing the population away from city centres in the baseline, the absence of a large number of old people means that the central share falls less and actually increases towards the end of the experiment.

In terms of sorting behavior of workers, figure 21 shows that initially worker growth is stronger in the small cities, but eventually increases even in the big cities, absent great numbers of old. The central and right panels of the figure highlight this further by showing that workers now sort increasingly towards the fringe of the city, where they can enjoy cheaper housing, while this part of the city was previously occupied by older people.

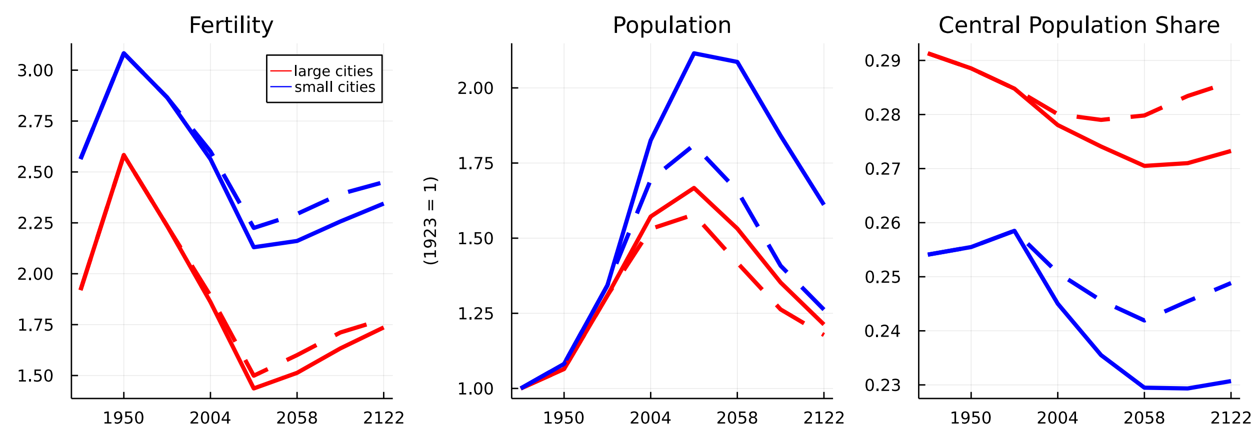


Figure 20: Aggregate implications of undoing increases in longevity. The dashed line illustrates the counterfactual.

4.2.3 Decrease in Commuting Cost

Here we focus on the strong increase in automobile usage for commuting in the 1960s and it's effect on city structure and fertility. The share of people using automobiles for their commute increase from about 10% in the 1950s to about 70% 30 years later in France. We implement this change in the model with a drop in commuting costs at date $t = 1$ (corresponding to 1950), $\tau_t = \tau - \Delta_t \tau$, with $\Delta_t \tau > 0$ for $t \geq 1$ and 0 otherwise. The decrease in commuting costs corresponds to better

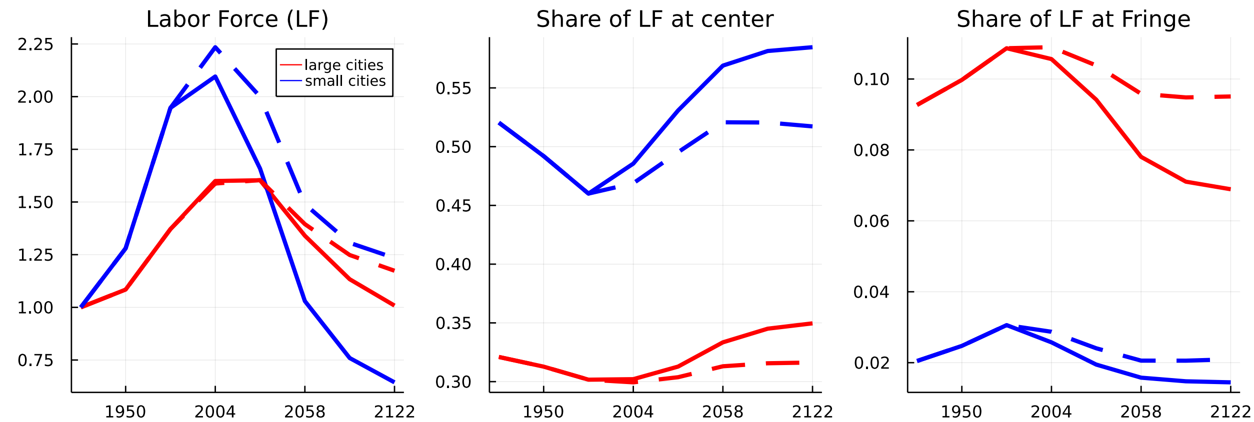


Figure 21: Implications of Labor Force sorting of undoing increases in longevity. The dashed line illustrates the counterfactual.

commuting technologies in general in the 1960s-1970s and is not limited to automobiles alone.

Figure 22 makes the point that faster commutes trigger higher fertility, which is visible in both left and central panel. It is worth pointing out that the effect on fertility is not very large because the general equilibrium is at play here: lower commuting costs allow to commute from further away, reducing pressure on the housing market, reducing housing costs, and hence increasing fertility. This increase, however, is dampened by population growth which now endogenously increases faster, and thus drives housing costs higher again. The right panel emphasizes the tight link between suburbanisation and fertility. A pronounced fall in the central population share, particularly in the large cities, goes hand in hand with an increase of fertility in the aggregate.

This latter point is illustrated further in figure 23, where the left panel confronts sorting of the labor force across large and small cities in baseline and low commuting cost counterfactual. We see stronger growth large city work force, and weaker growth in smaller cities, particularly in the later periods. The central and right panels show that suburbanization happens in both small and large cities, but that the center of smaller cities empties at a faster rate.

4.2.4 Housing Supply Regulation

In this counterfactual excercise we want to highlight the role of housing regulation for aggregate fertility dynamics. The effects on fertility are relatively straightforward: more restrictive housing supply means higher prices, and leads to lower fertility and population growth. Contrarily, the effects on the extent of suburbanization are a priori ambiguous in this setting: on the one hand, tighter supply increases housing costs, leading to lower fertility, hence relocating more people towards the centre of cities; on the other hand, however, the fall in fertility makes aging of the economy even more salient, thereby increasing suburbanization. In terms of implementation, we reduce housing supply in all locations by 20%, starting in the late 1990s, i.e. at model date $t \geq 2$, $\delta_t = \delta - \Delta_t \delta$, with $\Delta_t \delta = \Delta \delta > 0$, for $t \geq 2$ and 0 otherwise. The experiment aims at partly mimicking the recent

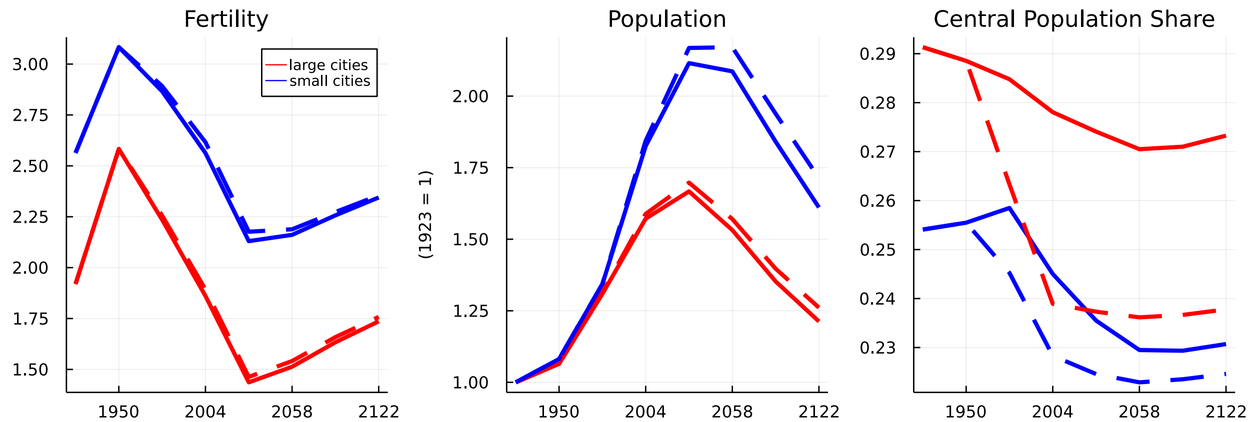


Figure 22: Counterfactual decrease in commuting costs. The dashed line illustrates the counterfactual.

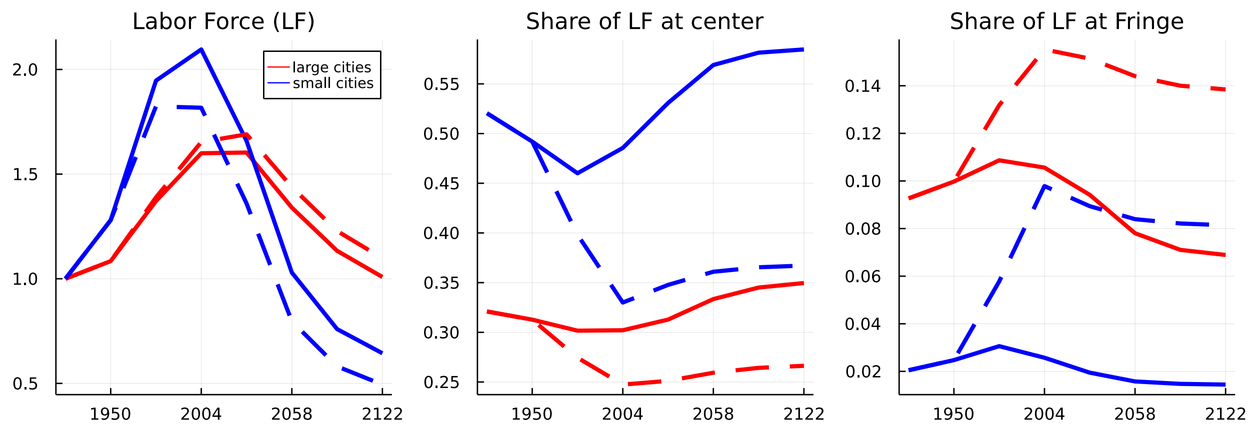


Figure 23: Labor force sorting after a counterfactual decrease in commuting costs. The dashed line illustrates the counterfactual.

rise in housing prices.

Figure 24 shows in the left panel that fertility falls as prices are increased. Population growth is lower as compared to the baseline, and there is a smaller push to the suburbs. Interestingly, and as illustrated in figure 25, while there are overall fewer workers as a result of lower population in general, there is a slight increase in the share of the labor force in the center of cities. Because those workers have fewer children, they sort closer to the center of cities.

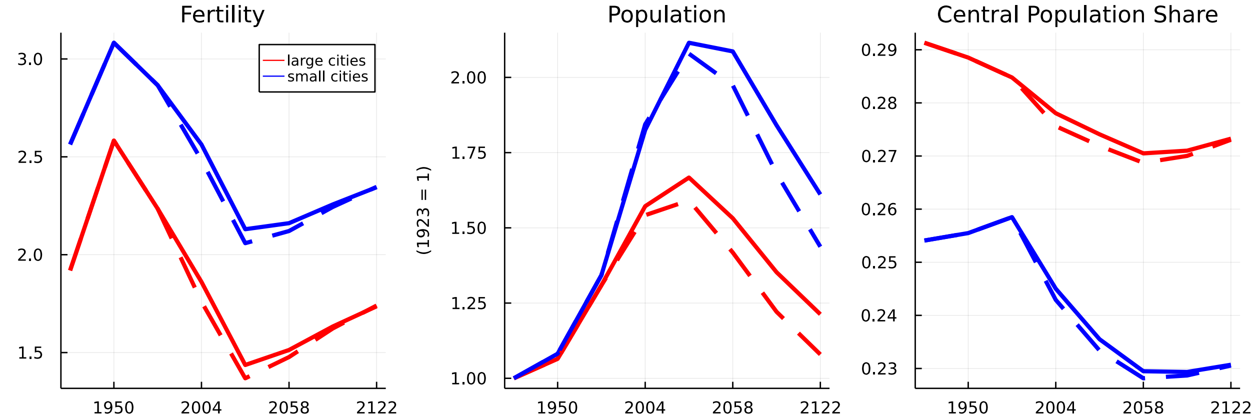


Figure 24: Counterfactual with stricter housing supply regulation. The dashed line illustrates the counterfactual.

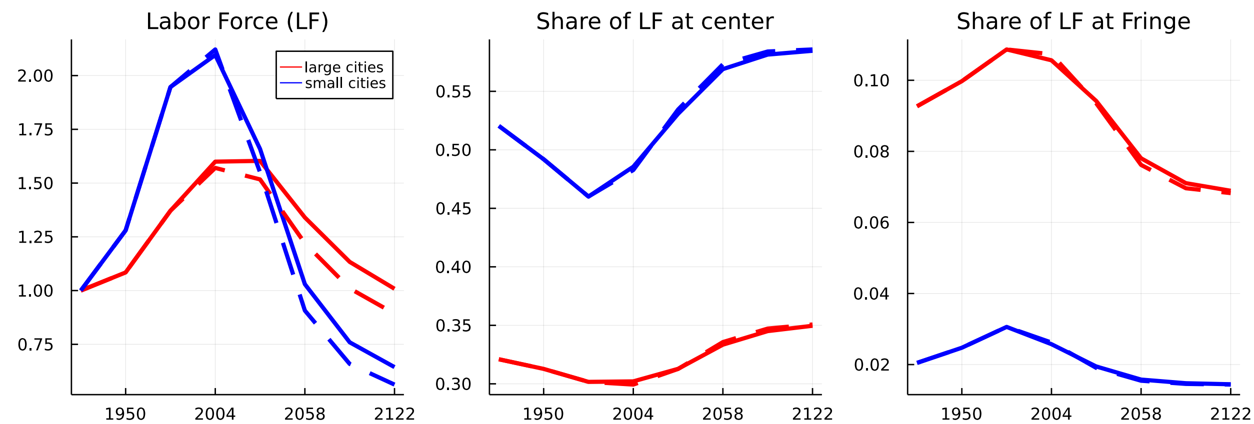


Figure 25: Labor force sorting with counterfactually stricter housing supply regulation. The dashed line illustrates the counterfactual.

4.3 Identification

to be written

4.4 Estimation and Results

to be written

4.5 Counterfactuals

to be written

5 Conclusion

to be written

References

Baum-Snow, Nathaniel, "Did highways cause suburbanization?," *The quarterly journal of economics*, 2007, *122* (2), 775–805.

_ **and Ronni Pavan**, "Understanding the City Size Wage Gap," *The Review of Economic Studies*, 08 2011, *79* (1), 88–127.

Becker, Gary S, "An Economic Analysis of Fertility. Bureau of Economic Research. Demographic and Economic Changes in Developed Countries," 1960.

_ **and H Gregg Lewis**, "On the Interaction between the Quantity and Quality of Children," *Journal of political Economy*, 1973, *81* (2, Part 2), S279–S288.

Black, Dan A, Natalia Kolesnikova, Seth G Sanders, and Lowell J Taylor, "Are children "normal"?," *The review of economics and statistics*, 2013, *95* (1), 21–33.

Bolhuis, Marijn A and Judd NL Cramer, "The Millennial Boom, the Baby Bust, and the Housing Market," *arXiv preprint arXiv:2003.11565*, 2020.

Caldwell, John C, "Demographic theory: A long view," in "Demographic transition theory," Springer, 2006, pp. 301–320.

Chapelle, Guillaume, Laurent Gobillon, and Benjamin Vignolles, "Public housing development and segregation: SRU law in France," Technical Report, Centre for Economic Policy Research 2022.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon, "Spatial wage disparities: Sorting matters!," *Journal of urban economics*, 2008, *63* (2), 723–742.

_ **,** _ **, and** _ , "The costs of agglomeration: House and land prices in French cities," *The Review of Economic Studies*, 2018, *86* (4), 1556–1589.

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

_ **, Cecile Gaubert, Jessie Handbury, and Erik Hurst**, "Income growth and the distributional effects of urban spatial sorting," Technical Report, National Bureau of Economic Research 2019.

Davis, Kingsley, "Reproductive institutions and the pressure for population," *The Sociological Review*, 1937, *29* (3), 289–306.

Dettling, Lisa J and Melissa S Kearney, "House prices and birth rates: The impact of the real estate market on the decision to have a baby," *Journal of Public Economics*, 2014, *110*, 82–100.

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

Doepke, Matthias, Anne Hannusch, Fabian Kindermann, and Michèle Tertilt, "The economics of fertility: A new era," Technical Report, National Bureau of Economic Research 2022.

Easterlin, Richard A, "The American Baby Boom in Historical Perspective," *The American Economic Review*, 1961, *51* (5), 869–911.

_ , "Population change and farm settlement in the northern United States," *The Journal of Economic History*, 1976, *36* (1), 45–75.

Eckstein, Zvi, Steven Stern, and Kenneth I Wolpin, "Fertility choice, land, and the Malthusian hypothesis," *International Economic Review*, 1988, pp. 353–361.

Eeckhout, Jan, Roberto Pinheiro, and Kurt Schmidheiny, "Spatial sorting," *Journal of Political Economy*, 2014, *122* (3), 554–620.

Glaeser, Edward L and David C Mare, "VCities and SkillsV," *Journal of Labor Economics*, 2001, *19* (2), 316.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst, "Endogenous gentrification and housing price dynamics," *Journal of Public Economics*, 2013, *100*, 45–60.

Hiller, Norbert and Oliver W Lerbs, "Aging and urban house prices," *Regional Science and Urban Economics*, 2016, *60*, 276–291.

Hotz, V Joseph, Jacob Alex Klerman, and Robert J Willis, "The economics of fertility in developed countries," *Handbook of population and family economics*, 1997, *1*, 275–347.

Jones, Charles I, "The end of economic growth? Unintended consequences of a declining population," Technical Report, National Bureau of Economic Research 2020.

Jones, Larry E, Alice Schoonbroodt, and Michele Tertilt, "Fertility theories: can they explain the negative fertility-income relationship?," in University of John Shoven John Shoven, ed., *Demography and the Economy*, University of Chicago Press, 2008.

Knoll, Katharina, Moritz Schularick, and Thomas Steger, "No price like home: Global house prices, 1870-2012," *American Economic Review*, 2017, *107* (2), 331–53.

Komissarova, Kristina, "Location Choices over the Life Cycle: The Role of Relocation for Retirement," 2022.

la Croix, David De and Paula E Gobbi, "Population density, fertility, and demographic convergence in developing countries," *Journal of Development Economics*, 2017, *127*, 13–24.

Lee, Ronald D, "Population dynamics of humans and other animals," *Demography*, 1987, pp. 443–465.

Lino, M, K Kuczynski, N Rodriguez, and T Schap, "Expenditures on children by families, 2015. Miscellaneous Publication No. 1528-2015," *US Department of Agriculture, Center for Nutrition Policy and Promotion*, 2017.

Lovenheim, Michael F and Kevin J Mumford, "Do family wealth shocks affect fertility choices? Evidence from the housing market," *Review of Economics and Statistics*, 2013, *95* (2), 464–475.

Lutz, Wolfgang, Maria Rita Testa, and Dustin J Penn, "Population density is a key factor in declining human fertility," *Population and Environment*, 2006, *28* (2), 69–81.

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

Mills, L Scott, *Conservation of wildlife populations: demography, genetics, and management*, John Wiley & Sons, 2012.

Moreno-Maldonado, Ana and Clara Santamaria, "Delayed Childbearing and Urban Revival," *mimeo*, 2022.

Murphy, Kevin M, Curtis Simon, and Robert Tamura, "Fertility decline, baby boom, and economic growth," *Journal of Human Capital*, 2008, *2* (3), 262–302.

Notestein, Frank W, "Population-The long view.," *Food for the World.*, 1945, pp. 36–57.

Notestein, F.W., "Economic problems of population change," in "Proceedings of the Eighth International Conference of Agricultural Economists" 1953.

OECD, "House Prices," 2022.

Redding, Stephen J, "Suburbanization in the United States 1970-2010," Technical Report, National Bureau of Economic Research 2021.

– **and Esteban Rossi-Hansberg**, "Quantitative spatial economics," *Annual Review of Economics*, 2017, *9*, 21–58.

Roca, Jorge De La and Diego Puga, "Learning by working in big cities," *The Review of Economic Studies*, 2017, *84* (1), 106–142.

Sato, Yasuhiro, "Economic geography, fertility and migration," *Journal of Urban Economics*, 2007, *61* (2), 372–387.

Sibly, Richard M and Jim Hone, "Population growth rate and its determinants: an overview," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 2002, *357* (1425), 1153–1170.

Simon, Curtis J and Robert Tamura, "Do higher rents discourage fertility? Evidence from US cities, 1940–2000," *Regional Science and Urban Economics*, 2009, *39* (1), 33–42.

Sinclair, Anthony RE, "Mammal population regulation, keystone processes and ecosystem dynamics," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 2003, *358* (1438), 1729–1740.

Sinclair, ARE, "Population regulation in animals. Ecological Concepts," *Black well Scientific, Oxford*, 1989, pp. 197–241.

Thompson, Warren, "Population," *American Journal of Sociology*, May 1929, *34*, 959–975.

Thompson, Warren S, *Population problems 4th Edn.*, McGraw-Hill Book Company, New York, 1916.

United Nations, "World population prospects 2019," Technical Report, United Nations Population Division 2019.

Willis, Robert J, "A new approach to the economic theory of fertility behavior," *Journal of political Economy*, 1973, *81* (2, Part 2), S14–S64.

A Data

This Appendix gives more details on our data sources.

A.1 Census data

Censuses over the 1968-1999 period are exhaustive but the whole information of census forms was digitized only for a restricted sample of every census. Restricted samples were obtained by selecting dwellings randomly according to the following sampling rates: 1968 ($1/4^{th}$), 1975 ($1/5^{th}$), 1982 ($1/4^{th}$) and 1999 ($1/4^{th}$). Censuses data in 2010 and 2015 are constructed from 5 annual census surveys covering respectively the 2008-2012 and 2013-2017 periods. Every annual census surveys $1/5^{th}$ of the population in every municipality with more than 10,000 inhabitants and $1/5^{th}$ of whole municipalities with less than 10,000 inhabitants. Restricted samples include all sampled households in municipalities with more than 10,000 inhabitants and 25% of sampled households in municipalities with less than 10,000 inhabitants.[20] The SAPHIR dataset consists in all harmonized restricted samples over the 1968-2015 period. For ordinary dwellings, weights are at the dwelling or address level whatever the census, and we use them at the household level.

A.2 Housing price data

Notary databases over the 2000-2012 period contain information on transactions of second-hand dwellings for the Paris region (BIEN database) and other regions (PERVAL database) reported at the end of every even year. Interval brackets for the construction year are: before 1850, 1850-1913, 1914-1947, 1948-1959, 1960-1980, 1981-1991 and after 1991. Information on construction year is missing for around 30% of observations. Missing is considered as a category when introducing dummies for construction year brackets.

A specific procedure is used to impute floor areas which are missing for 25.7% of transactions Chapelle et al. (2022). Dwellings with missing values are attributed the average floor area of dwellings in FILOCOM data (an exhaustive panel of dwellings) located in the same cadastral section, which are involved in a transaction during the same year, which are of the same type (single-family house or flat) and have the same number of rooms. Accuracy is evaluated by imputing the floor area for dwellings for which it is not missing and by comparing the imputed values with the observed ones. It is found that the average absolute error is around 5%, and the R^2 of the regression of the observed floor area on the imputed one is around 0.75. The imputation is more accurate for flats (resp. 2% and 0.83) than for single-family houses (resp. 15% and 0.51). After imputation, the proportion of dwellings with missing floor area decreases to 5.1%, and observations with remaining missing values are dropped. In the final sample used to compute housing price indices, there are around 560 thousands transactions per year (290 thousands for flats and 270 thousands for single-family houses).

[20] More details on the sampling schemes and weights are given at the address: https://www.insee.fr/fr/statistiques/2414232#documentation-sommaire.

A.3 Housing consumption data

[Provide details on Enquête National Logement (ENL).]

B Empirical Facts

B.1 Introductory Facts

Figure B.1 plots for 2015 the measure of fertility in the center (children per adult in parental age in the densest municipality of the urban areas) by bins of population in the urban area (UA). Data are averaged in each bin across urban areas (UA's population weights). Results are almost identical if one takes the main historical municipality of the urban area instead of the densest as central location.

Figure B.2 plots for 2015 a measure of fertility (children per adult in parental age) in the Parisian urban area by bins of distance from the center of Paris (bins of 4kms). Data are averaged in each bin across municipalities (population weights). The plot censors observations above 40 kms as fertility barely varies at higher level of distance.

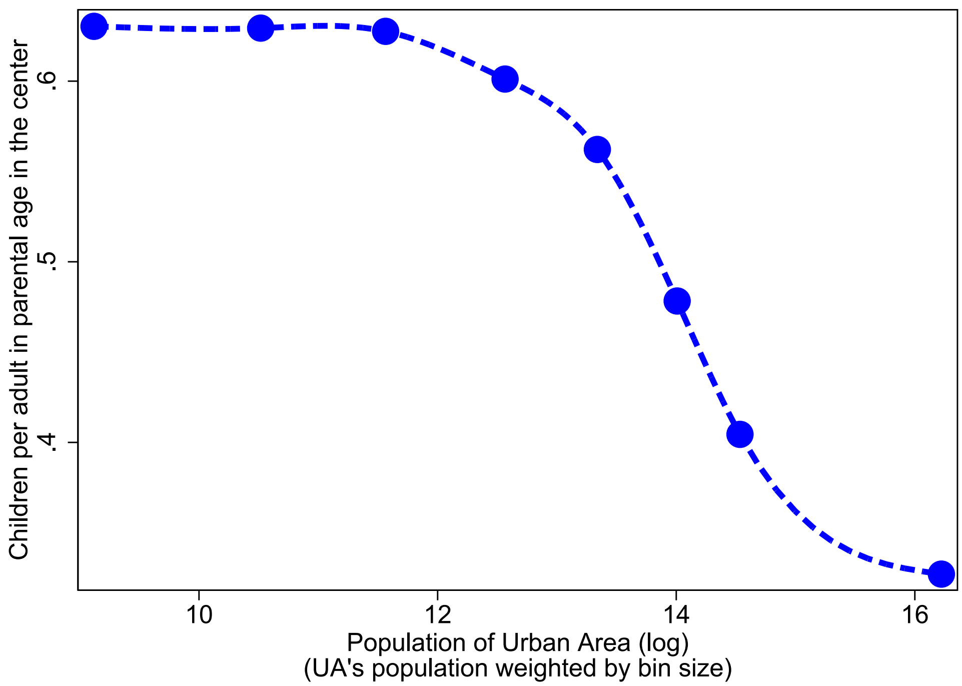


Figure B.1: Central fertility and population of the urban area in 2015.

Notes: The figure shows the number of children (age 1-17) per adult in parental age (age 27-53) in the center of urban areas as a function population of urban areas, binned into 8 size categories of population (in 1000s): ([0,10[, [10,30], [30, 100[, [100, 200[, [200, 500[, [500, 900[, [9, 2000[, and above 2000. The population of the urban area and children per adult in parental age in each bin are defined as the UAs population weighted average of the municipalities in this bin. The central location is most dense municipality within the urban area. *Source*: French Census 2015.

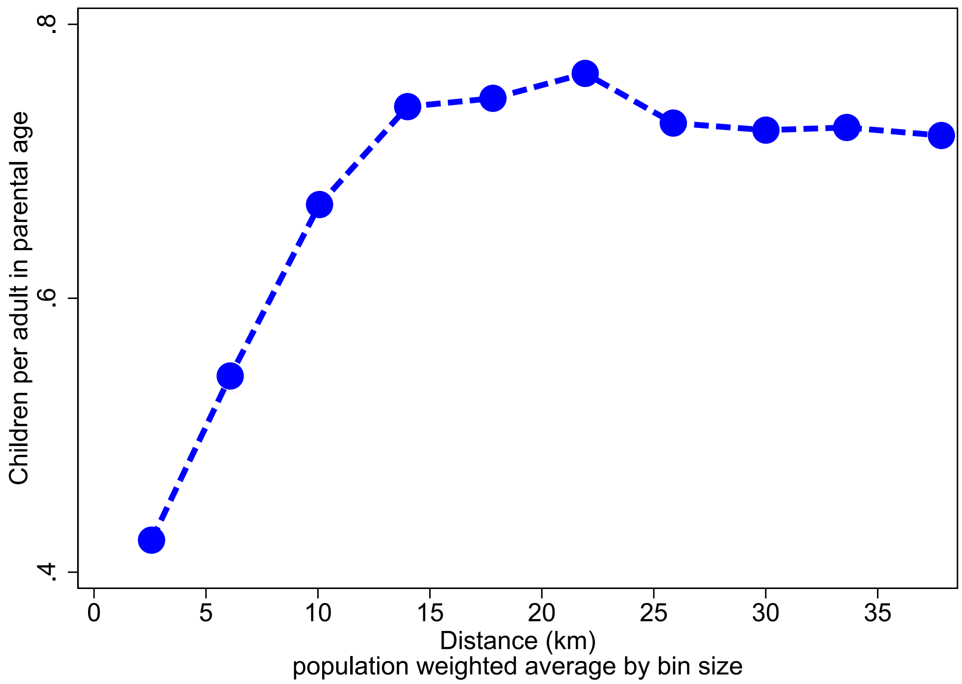


Figure B.2: Fertility within Paris and distance from the center in 2015.

Notes: The figure shows the number of children (age 1-17) per adult in parental age (age 27-53) as a function of distance from the center, binned into ten 4kms wide bins of distance (up to 40kms). The distance and children per adult in parental age in each bin are defined as the municipality population weighted average of the municipalities in this bin. *Source*: French Census 2015.

B.2 Section 2.2. Housing consumption and demographics

[TO BE WRITTEN]

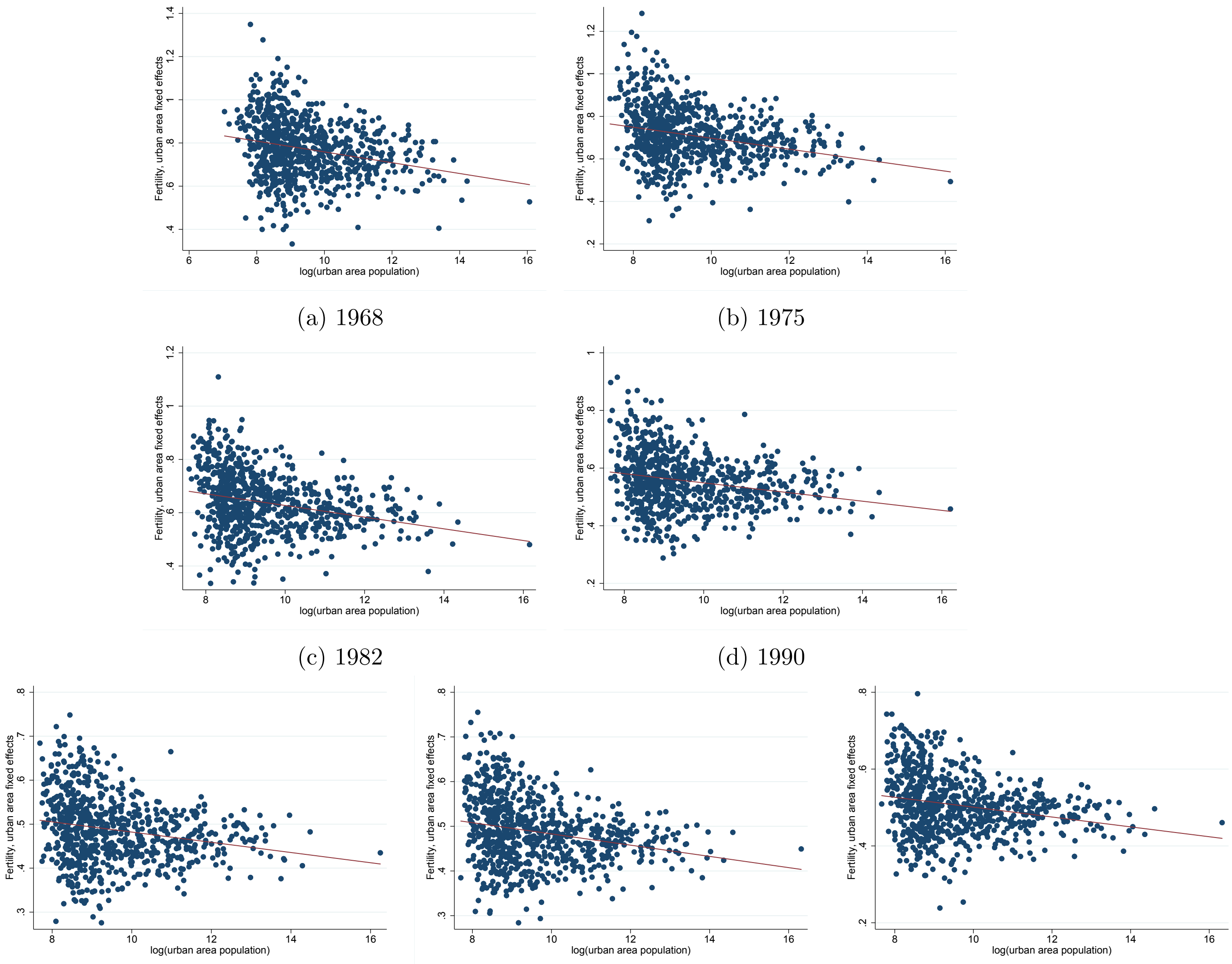
B.3 Section 2.3. Urban spatial sorting across demographics

Fertility across urban locations. The following scatter plots (Figure B.3) show the fertility in the center of urban areas (UA fixed-effects of Equation 2) as a function of the population in the urban area (in log) for each wave of the Census.

Spatial sorting by age.

Average age. The following scatter plots (Figure B.3) show average age of adults in the center of urban areas (UA fixed-effects of Equation 2) as a function of the population in the urban area (in log) for each wave of the Census.

Young versus older. The following scatter plots (Figures B.5 and B.6) show the ratios of young (18-26) to adults (18+) ratio and the old (54+) to adults ratio in the center of urban areas (UA fixed-effects of Equation 2) as a function of the population in the urban area (in log) for each wave of the Census. Figures B.7 and B.8 summarize the results across urban areas by Census waves for



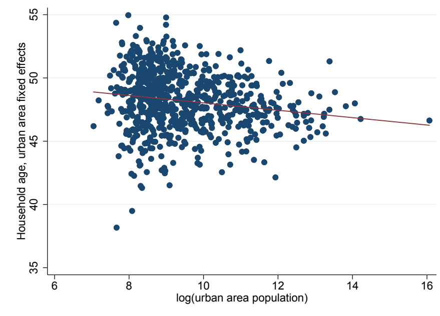
(a) 1968 (b) 1975

(c) 1982 (d) 1990

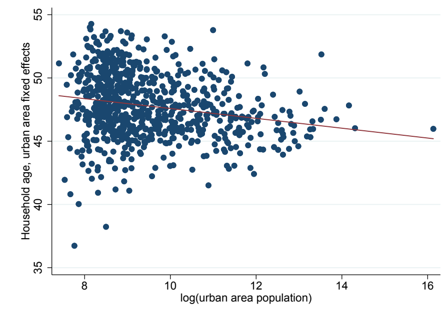
(e) 1999 (f) 2010 (g) 2015

Figure B.3: Children per adult in parental age (in the center) and urban area population.

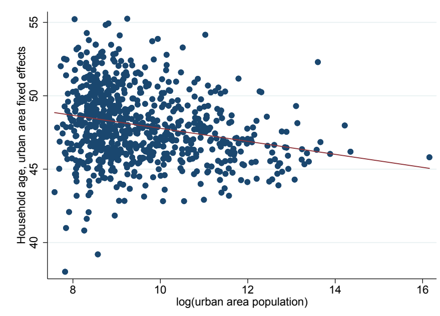
Notes: Each panel plots the relationship between urban area children per adult fixed effects (c_k) and log of population of the urban area. Urban area children per adult fixed effects (c_k) are obtained by regressing, at the municipality level, children per adults in parental age on urban area (UA) dummies and dummies by deciles of distance from the center of the UA (Equation 2).



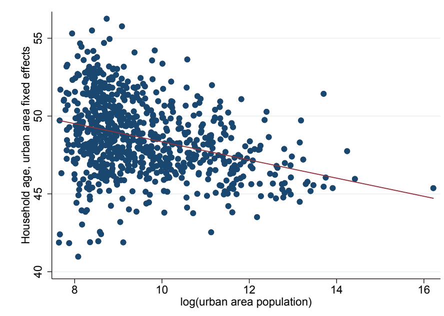
(a) 1968



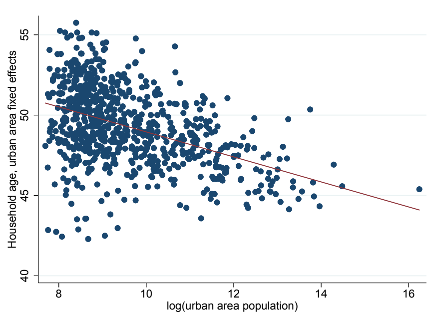
(b) 1975



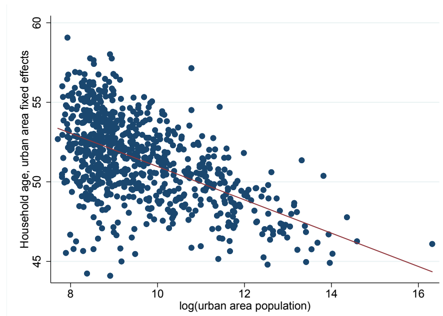
(c) 1982



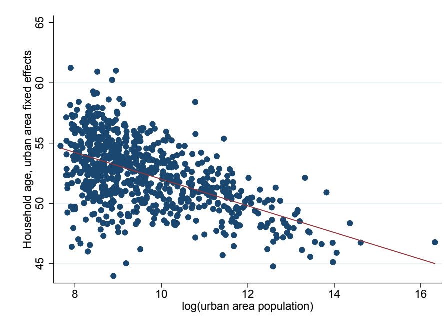
(d) 1990



(e) 1999



(f) 2010



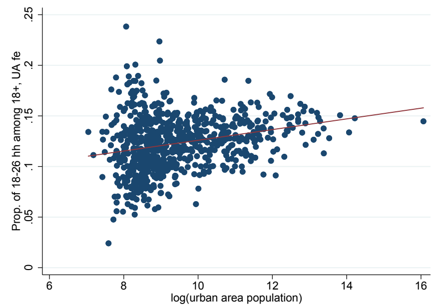
(g) 2015

Figure B.4: Average age of adults (in the center) and urban area population.

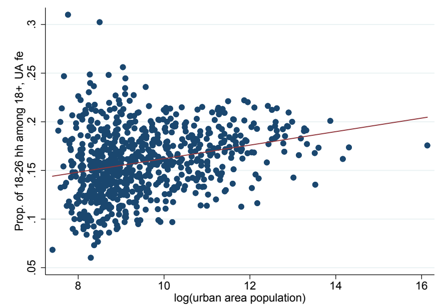
Notes: Each panel plots the relationship between average adult age fixed effects (c_k) and log of population of the urban area. Urban area average adult age fixed effects (c_k) are obtained by regressing, at the municipality level, the A avergae age of adults on urban area (UA) dummies and dummies by deciles of distance from the center of the UA (Equation 2).

the young and old to adults ratios. Figure B.9 (resp. Figure B.10) show how the young to adults ratio (resp. the old to adults ratio) evolve within cities by deciles of distance from the center.

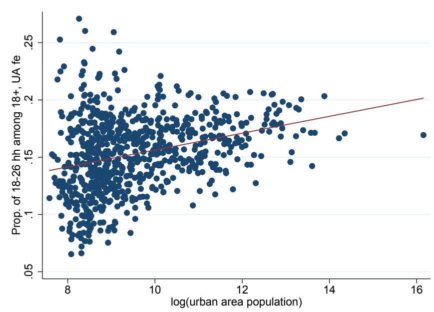
Sensitivity analysis. [TO BE WRITTEN]



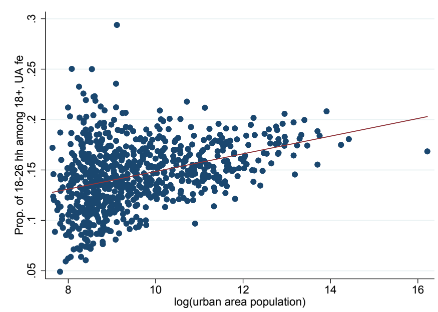
(a) 1968



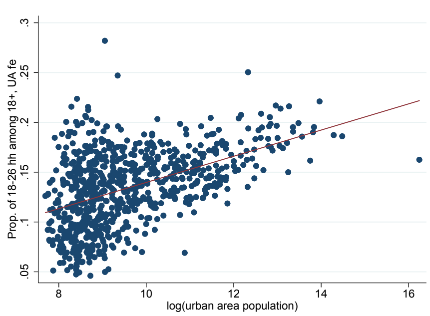
(b) 1975



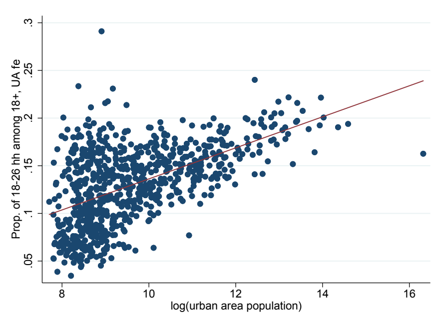
(c) 1982



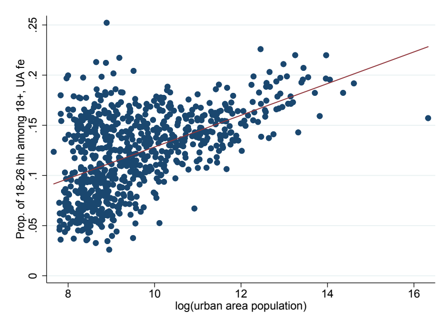
(d) 1990



(e) 1999



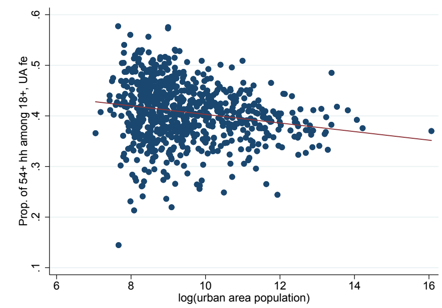
(f) 2010



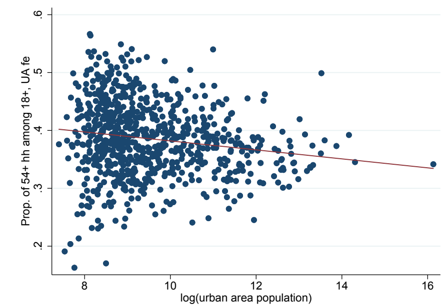
(g) 2015

Figure B.5: Young (18-26) per adult (18+) (in the center) and urban area population.

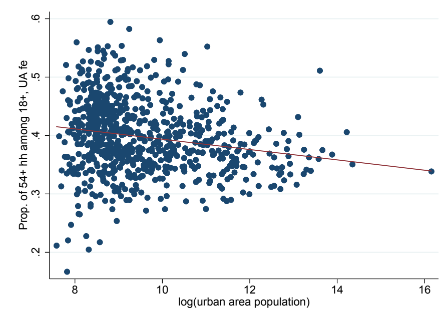
Notes: Each panel plots the relationship between urban area Young (18-26) per adult (18+) fixed effects (c_k) and log of population of the urban area. Urban area children per adult fixed effects (c_k) are obtained by regressing, at the municipality level, the young to adults ratio on urban area (UA) dummies and dummies by deciles of distance from the center of the UA (Equation 2).



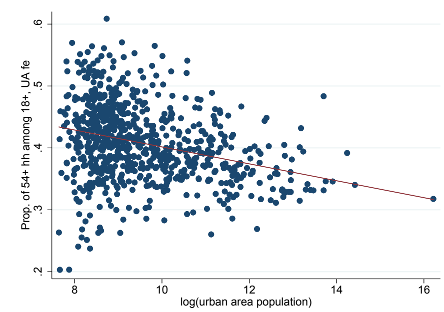
(a) 1968



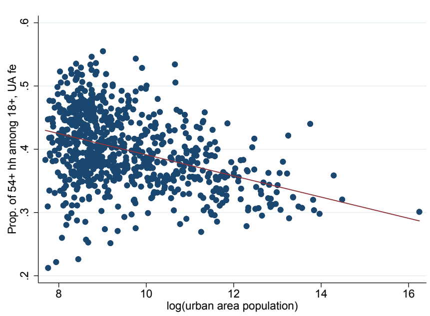
(b) 1975



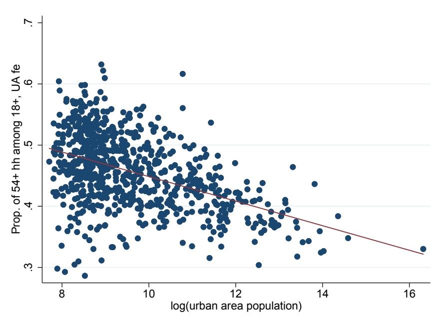
(c) 1982



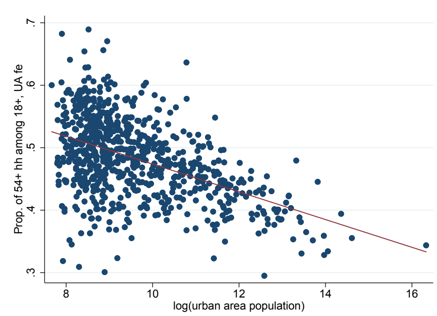
(d) 1990



(e) 1999



(f) 2010



(g) 2015

Figure B.6: Old (54+) per adult (18+) (in the center) and urban area population.

Notes: Each panel plots the relationship between urban area Old (54+) per adult (18+) fixed effects (c_k) and log of population of the urban area. Urban area children per adult fixed effects (c_k) are obtained by regressing, at the municipality level, the old to adults ratio on urban area (UA) dummies and dummies by deciles of distance from the center of the UA (Equation 2).

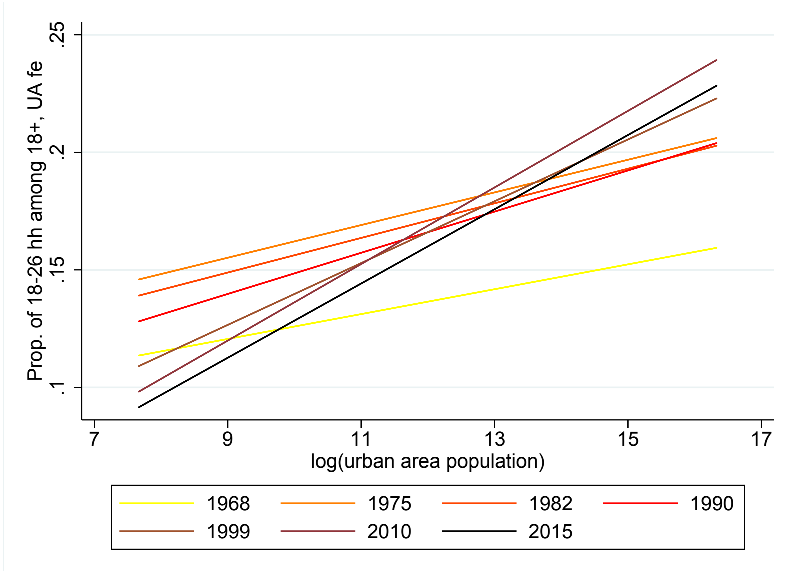


Figure B.7: Young (18-26) per adult (18+) and urban area population.

Notes: Each line corresponds to the estimated linear relationship between urban area fixed effects for the number of young (18-26) per adult (18+) and log of population of the urban area. Urban area number of young (18-26) per adult (18+) fixed effects (c_k) are obtained by regressing, at the municipality level, the number of young (18-26) per adult (18+) on urban area (UA) dummies and dummies by deciles of distance from the center of the UA (Equation 2).

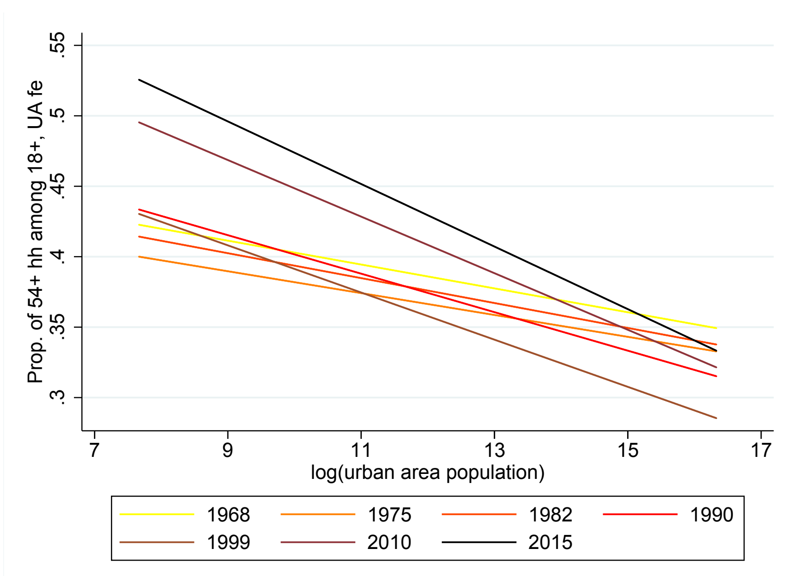


Figure B.8: Older (age 54+) per adult (18+) and urban area population.

Notes: Each line corresponds to the estimated linear relationship between urban area fixed effects for the number of older adults (age 54+) per adult (18+) and log of population of the urban area. Urban area older (age 54+) per adult (18+) fixed effects (c_k) are obtained by regressing, at the municipality level, the number of older adults (age 54+) per adult (18+) on urban area (UA) dummies and dummies by deciles of distance from the center of the UA (Equation 2).

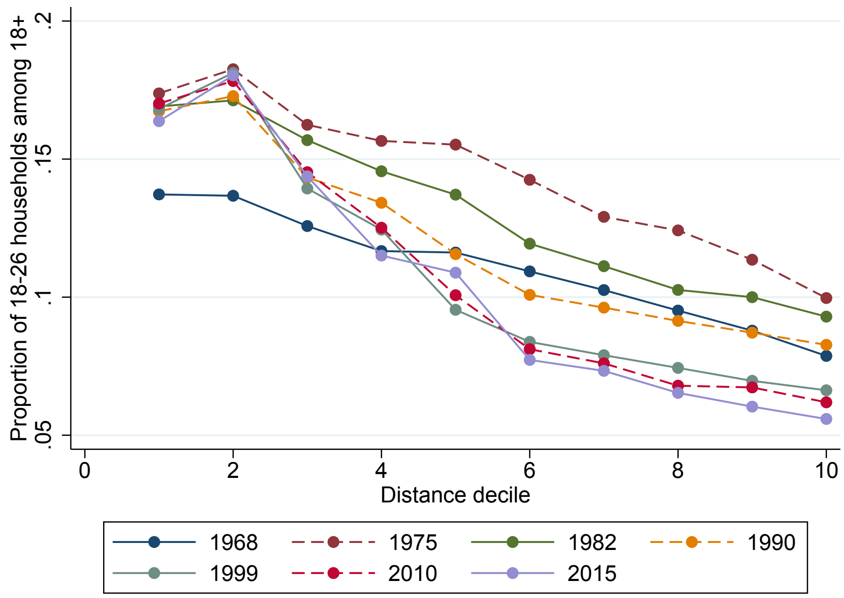


Figure B.9: Young (age 18-26) per adult (above 18) by distance deciles.

Notes: Number of young (18-26) per adult (18+) at different distance deciles are obtained by regressing, at the municipality level, the number of young (18-26) per adult (18+) on urban area (UA) fixed-effects and dummies by deciles of distance from the center of the UA (Equation 2).

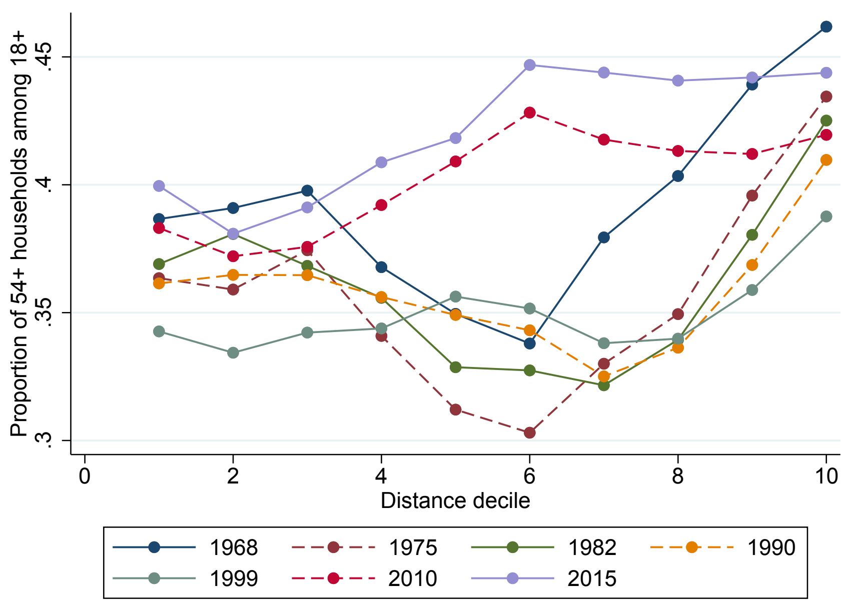


Figure B.10: Older (age 54+) per adult (18+) by distance deciles.

Notes: Older (age 54+) per adult (18+) at different distance deciles are obtained by regressing, at the municipality level, the number older adults (age 54+) per adult (18+) on urban area (UA) fixed-effects and dummies by deciles of distance from the center of the UA (Equation 2).