Delayed Childbearing and Gentrification

Ana Moreno-Maldonado Clara Santamaria

Click here for the latest version

Abstract

This paper highlights the role of delayed childbearing as an important driver of gentrification. While downtowns provide shorter commuting times and more consumption amenities, limited housing space and schools' worse quality considerably reduce the value of this location choice when children are born. Therefore, our argument in this paper is that, as technological advances enable women to postpone maternity, the life period in which individuals benefit the most from living downtown is extended. Consequently, downtowns become more attractive overall and gentrification takes place. We exploit exogenous variation in the cost of postponing childbearing to obtain causal estimates of the impact of delayed maternity on gentrification. Exogeneity comes from spatial variation in mandated insurance of Assisted Reproductive Technology (ART). We find that a higher availability of ART in the state increases income downtown by 5.4% relative to the suburbs. Moreover, consistently with the delay in fertility and subsequent relocation to the suburbs, the percentage of women living in the city center increases for the age group 25 to 30, while it decreases for those between 30 and 35.

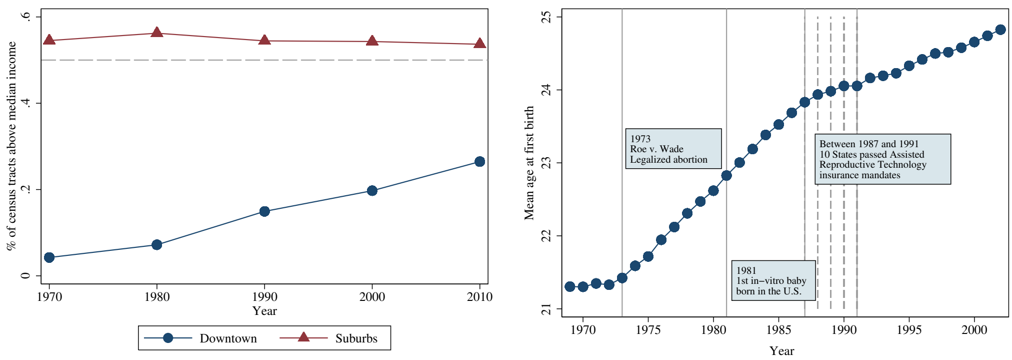
1 Introduction

Over the last decades, the increasing gentrification of downtown neighborhoods in many developed countries has attracted the attention of researchers and policy-makers.[1] While the influx of high-income individuals to low-income central neighborhoods may unfold a range of benefits to incumbent residents, such as amenity improvements or crime reduction, it certainly raises important concerns. Among

[1]In line with most of the economic literature, we define gentrification as the influx of households with higher socioeconomic status to central city neighborhoods, that often results in the displacement of earlier and less affluent residents (Baum-Snow and Hartley (2019)). This definition accounts for several important dimensions, being usually characterized by increased numbers of prosperous and educated individuals living in downtown urban locations, the renovation and refurbishment of houses, and rising rents and house prices.

them, one that stands out is the parallel increase of housing costs in the neighborhood, which may effectively displace low-income residents and thus become an important source of inequality for at least two reasons. First, given that displaced residents face higher traveling time to consumption amenities, displacement affects consumption inequality within the city. Second, and depending on the spatial distribution of jobs, displacement may erode their labour market opportunities by substantially increasing commuting times to high-wage occupations.

Although the consequences of gentrification have drawn considerable attention in public policy debates, the identification of the underlying factors behind this growing phenomenon remains controversial, while their understanding is still fairly limited. Recent studies outlined some potential determinants of gentrification, such as the rising valuation of downtown amenities (Couture and Handbury (2017), Couture et al. (2019)), the reduction in downtown crime rates (Ellen et al., 2019), and the increased opportunity cost of long commuting times (Edlund et al. (2015), Su (2018)). Nonetheless, in this paper we take a different perspective of this issue by focusing on the remarkable cultural and socioeconomic transformations that the role of women has undergone in society during recent decades, and which have modified traditional lifestyles and family structures. In particular, we realise that fertility decisions are inextricably linked to residential location choices and propose delayed childbearing as a novel driver of the gentrification process.



(a) Gentrification (b) Age at first birth

Figure 1: Delayed fertility and gentrification

We first notice that the decades of rapid gentrification are precisely those in which a steady increase in the age of first-time mothers has taken place. Figure 1 shows that from 1970 to 2010 the probability that a census tract in the city center exhibited an average income above the median income in the city went from 5 percent to 25 percent (panel 1a). During the same period, the average age of first mother went from just over 21 to close to 25 years old (panel 1b). Although, in principle, the causal relationship between these two trends, if any, could go in either direction, our goal in this paper is to identify a plausible causal effect of the delay in fertility on gentrification. We do so by exploiting exogenous policy changes at the state level that decreased the cost of postponing childbearing, and document that the implementation of such policies has important effects on the demographic composition and gentrification of US cities.

Moreover, we show that, over our period of study (1970-2010), families are much more likely to locate in the suburbs of the city than other household types. We argue that families' location patterns are driven by the proximity to better schools and to larger housing size in the suburbs as compared to downtown locations. In contrast, downtowns offer a higher density of consumption amenities, such as restaurants, bars, cinemas and theatres, which are typically more valued among young individuals, especially by those enjoying higher income and living in households without children. Therefore, as technological advances enable women to postpone childbearing, the life period in which downtown amenity consumption is highest is extended, enhancing the incentives to locate in the center among this group of people.

Despite the fact that delayed parenthood is common to all developed countries, postponement can be very costly given that fertility decays sharply with age. In this context, technological advances in Assisted Reproductive Techniques (ART) offer some insurance against the risk of infertility associated to late childbearing. However, these treatments are very expensive and access to them remains quite limited. Our empirical strategy exploits state variation in the price of ART induced by infertility insurance mandates. In the late 80's, several US states enacted ART insurance mandates, which in practice implied a substantial reduction of the price of ART treatments that couples faced. This resulted in a large rise in the access rate to ART and in an increase in the average age at first birth in those states (Hamilton and McManus (2012), Buckles (2005), Schmidt (2007)). Therefore, this policy provides a nice scenario to assess the impact of delayed parenthood on the gentrification of downtown neighborhoods. Admittedly, postponed maternity is a much broader phenomenon which it is certainly not limited to the states that enacted infertility mandates. In this sense, we believe that our results are more general than simply measuring the impact of the mandates on gentrification, since they are also useful to understand the interaction between demographic change and neighborhood development.

Using a difference-in-difference-in-differences (triple difference) approach, we find that the existence of a state mandate to cover ART leads to gentrification. Downtown income relative to the suburbs increases by 3.5% more in treated cities than in cities that belong to the control group, and downtown neighborhoods are 7.3 percentage points (p.p.) more likely to be above the median city income. Moreover, the larger average income of residents in the city center goes in parallel with a demographic change that is consistent with gentrification. Specifically, the share of college graduates in downtown neighborhoods belonging to treated cities increases by almost 3 p.p. both relative to the suburbs and the non-treated cities. In addition, the age distribution of women also changed in the expected direction, with an increase of 2 p.p. of women between the ages of 25 and 29 and a subsequent decrease of 2 p.p. among those between 30 and 35 years old. The age distribution of men reacts similarly but lagged by a few years, likely due to male partners being slightly older. We argue that these changes in the age composition downtown are fully consistent with couples postponing childbearing and moving to the suburbs.

In order to assess the overall contribution of our proposed channel to gentrification, we estimate a dynamic model with endogenous fertility and within city location. This approach allows us to quantify how much of the observed gentrification can be explained by delayed childbearing. That is, our model captures the full impact of postponing maternity, taking into consideration a broader set of reasons for this trend, such as changes in the career cost of children, instead of focusing only on the availability of ART. However, we plan to exploit the impact of infertility insurane mandates to estimate the elasticity of fertility decisions, which is key for our estimation.

In our model, we consider a city with two types of locations, downtown and suburbs, that differ in amenities, income, and housing supply. We restrict our attention to couples that are heterogeneous in age and skill, and have idiosyncratic preferences for locations and children. Couples decide jointly whether to have kids and when and where to live. While young couples are always successful when trying to have children, mature couples face some infertility risk. Therefore, couples with a very high taste for children choose to have children as young while couples with a more intermediate draw may postpone. In particular, our results show that an increase in the probability of having a sucessful birth when mature or an increase in the income penalty from having children early in their lifetime result in a higher share of young couples postponing maternity and locating in the center, where amenities for couples with no children are larger than in the suburbs.

Related Literature. First and foremost, our paper relates to a growing literature that analyzes the causes of downtown gentrification. Baum-Snow and Hartley (2019) points out that the propensity of young and high-income individuals to live in the city center is largely driven by two factors: (i) divergent preferences towards downtown amenities between different racial groups, and (ii) the rising suburban concentration of labor market opportunities for low-education workers. Likewise, Couture and Handbury (2017) also emphasize the role of amenity valuations, arguing that increases in gentrification in the 2000-2010 period can be explained by a growing taste for downtown amenities among college graduates. On the other hand, Edlund et al. (2015) argue that longer hours worked among high-skilled workers have increased their distaste for commuting, which in turn has pushed up house and rental prices in the city center. Similarly, Su (2018) examines the growing importance of long work hours in well-paid downtown-located jobs as an exogenous factor driving the demand for central locations by high-skilled workers. Couture et al. (2019) evaluate the impact of top-income growth and its associated rise in income inequality on the location choices of rich households. In order to quantify the welfare consequences of urban gentrification, they introduce idiosyncratic preferences shocks and endogenous amenities to a spatial model of urban sorting. Recent work by Almagro and Dominguez-Iino (2019), Curci and Yusaf (2020) and Hoelzlein (2019) also study how endogenous amenities reinforce sorting by income within cities. We contribute to this literature by outlining a novel important channel that leads to gentrification: delayed parenthood.

Moreover, our work speaks to the literature on women's timing of family formation. Goldin and Katz (2002) and Bailey (2006) examine the impact of the availability of the birth control pill on birth, marriage timing, and female labour supply. Goldin and Katz (2002) show that greater access to the pill reduces the likelihood of marrying before age 23 and therefore increases the likelihood of women being employed in professional and high-skilled occupations. In a similar vein, Bailey (2006) pinpoints that reducing the age at which it becomes legal to access to the pill reduces the likelihood of a first birth before age 22 and increases labor supply on both the intensive and extensive margins. Postponing childbearing may benefit women for several reasons. Caucutt et al. (2002) show that fertility delay is related to changes in marriage and labour markets. Thus, high-skilled women delay marriage and fertility in order to obtain a better match, even in the absence of returns to labour market experience. Moreover, when labour market experience is taken into account, fertility is delayed even further. Using biological fertility shocks to instrument for motherhood delay, Miller (2011) finds that postponing motherhood has a statistically significant and positive impact on earnings and career paths, particularly for the highly educated women. Our contribution to this strand of the literature is therefore to highlight a new set of consequences of delayed maternity, those related to neighborhood development and within-city inequality.

Several studies have confirmed the effectiveness of infertility insurance mandates on increasing ART utilization. Hamilton and McManus (2012) find that mandates to cover ART lead to a substantial increase in the usage of these technologies in the market. Moreover, they show that variations in the insurance regulations of states are largely due to different general political preferences rather than to unobserved preferences for ART. Similarly, Jain et al. (2002) find that states with required coverage for In Vitro Fertilization (IVF) - the most effective and most widely used form of ART - have the highest rates of IVF utilization. While these works provide suggestive evidence on how the mandates have increased IVF usage, they do not control for unobservable differences in patients or clinics that may be state-specific. This gap is filled by Bitler and Schmidt (2006), who use a difference-in-differences approach to show that the sizable increase in the use of infertility treatments as a result of the ART mandates is mainly concentrated among highly- educated older women, with no significant impact on other socioeconomic groups.

In addition, there is another stream of literature which has estimated the causal impact of infertility insurance mandates on several outcomes that are relevant to our framework. First, Schmidt (2007) finds a significant increase on first birth rates for women over 35. Consistent with this, Machado and Sanz-de Galdeano (2015) report a positive association between the mandates and women's mean age at first birth. However, they show that fertility rates over women's reproductive lives are unaffected by these mandates. Second, Buckles (2005) encounters that mandates that cover IVF are associated with an increase in labour force participation and earnings for women under 35 and a reduction in participation for older women. Similarly, Abramowitz (2017) points to an increase in women's age at marriage and at first birth after the enactment of these mandates, though only for college graduate women. Third, Kroeger and La Mattina (2017) find that such mandates led to a rise the probability that women hold a professional college degree and work in professional occupations. Our contribution to this literature is to uncover another consequence that had not been previously considered, namely, their effect on the spatial distribution of income within cities.

The rest of the paper is organized as follows. Section 2 presents the data. Section 3 provides the background about the implementation of the infertility treatment mandates that we exploit for the causal identification. Section 4 introduces the empirical specification and discusses the identification assumptions. Section 5 presents the main results on the causal effect of mandated ART treatments on gentrification and discusses the main mechanisms. Section 6 includes the structural model about the effect of fertility timing on within city location choices. Sections 7 and 8 present the main quantitative results from our estimated model and the counterfactual exercise, respectively. Finally, Section 9 concludes.

2 Data Sources and Definitions

Our analysis is conducted at the census tract level, defined as small geographical units encompassing between 2,500 and 8,000 people, which provides a good approximation for our definition of neighborhoods. We combine decennial Census data and the American Community Survey (ACS) 2008-2012, downloaded from the National Historical Geographic Information System (NHGIS), and construct constant 2010 census tract boundaries using the Longitudinal Tract Data Base (LTDB).

Cities. Our definition of city is the Core-Based Statistical Areas (CBSA) constructed by the Census Bureau. Given that gentrification is a big city phenomenon (Hwang and Lin (2016)), we restrict our sample to neighborhoods located in metropolitan areas with more than 1 million inhabitants.[2] The sample size includes 82,129 census tract- census year observations (51,469 in the treated group and 30,660 in the control group).

Downtown and suburbs. In line with the previous literature, we normalize distance to the city center using the cumulative share of the metropolitan population who lives in the nearest locations. That is, we consider rings of population around the city center such that a given share of population is included; e.g. a distance equal to 0.3 includes the area of the city including the 30 percent of population that is the closest to the city center. In particular, we use data from Lee and Lin (2018) to locate the geographical center of each CBSA and define the city center as the area within 0.1 distance from it. The main advantage of this definition is its flexibility as compared to geographical distances, since it adjusts for the fact that downtowns are generally more extensive in larger metropolitan areas. Similarly, we define the suburbs as the area of the city that contains the 50 percent of population that lives the furthest away from the city center.

Gentrification. Several measures of gentrification are used in this paper. First, we use the probability that the average income in a specific census tract is above the median income in the city. This measure provides a good metric to describe gentrification processes, as it captures the income in that particular area relative to the median income in the entire city. However, it potentially misses changes in income at the tails of the distribution. This pitfall is overcome by our second measure, the (log) median income in the census tract. Third, we use the percentage

[2] In the Appendix, we relax this restriction by replicating our main results for a sample which includes all cities that have more than 100,000 inhabitants, and obtain similar estimates.

of college graduates in the neighborhood, as gentrification is mostly driven by high-skilled individuals (Couture and Handbury, 2017).

3 Infertility insurance mandates

While several studies have found positive effects of delaying childbearing on women's lifetime earnings (Buckles 2008; Caucutt et al. 2002; Miller 2011; Wilde et al. 2010), it is well known that fertility decays sharply with age (Menken et al. 1986; van Noord-Zaadstra et al. 1991). In particular, the probability of having a successful conception within one year after starting to try to conceive is 75% for women at age 30, while it declines to 66% and 44% at ages 35 and 40 years, respectively (Leridon, 2004). Thus, by enabling women to postpone childbearing, ARTs may relax the career-family trade-off that women often face.

However, ART treatments (specially IVF) are very expensive. According to Hamilton and McManus (2012), one cycle of IVF entails an out-of-pocket cost of \$10,000 to \$15,000 to the patient and it is common to attempt multiple cycles of treatment. Moreover, it is rare that insurers cover these costs unless required by law.

Starting at the end of the 80's, several US states enacted mandates to enhance ART access. In practice, infertility insurance mandates amount to a significant reduction of the price borne by the patient, which is expected to increase utilization by making it affordable to a broader segment of the female population. Although several studies have shown that access to ART remained mostly limited to high-skilled white women (Bitler and Schmidt 2006; Hamilton and McManus 2012), they also notice that these mandates may affect younger women's decisions without necessarily increasing their own utilization of ART afterwards. The reason is that infertility insurance mandates affect the expected value of delaying childbearing: by lowering the cost of ART treatments, they reduce the risk associated with infertility at older ages. On top of that, they may have increased awareness about the availability of IVF and consequently changed women's misconceptions about its effectiveness. Lastly, increased IVF usage may have reduced the stigma associated to marrying and having children at an older age for the whole population of women.

Table 1: States with mandated infertility insurance

State	Date enacted	Mandate to cover	Mandate to offer	IVF coverage	Type of insurers	Treated
Arkansas	1987	X		X	HMOs excluded	
California	1989		X		All	
Connecticut	1989	X		X	All	X
Hawaii	1987	X		X	All	X
Illinois	1991	X		X	All	
Louisiana	2001	X			All	
Maryland	1985	X		X	All	X
Massachusetts	1987	X		X	All	X
Montana	1987	X			HMOs only	
New Jersey	2001	X		X	All	
New York	1990	X			HMOs excluded	
Ohio	1991	X		X	HMOs only	
Rhode Island	1989	X		X	All	X
Texas	1987		X	X	All	
West Virginia	1977	X			HMOs only	

Notes: This table summarizes the main features of acts mandating infertility insurance in all states that ever passed a mandate of this type. HMOs refers to Health Maintenance Organizations. The column treated displays which states we consider as part of the treatment.

Table 1 lists all states that have enacted mandates affecting the insurance of ART procedures over the five decades covering our census samples (1970-2010) and summarises their main features. There are several sources of heterogeneity across state mandates. First, while most states require insurers to cover ARTs treatments in every available insurance policy, mandates in California and Texas only require insurers to offer infertility treatments. In addition, not all mandates include IVF treatments nor affect every type of insurance provider. In particular, some mandates exclude health maintenance organizations (HMOs) while others only target HMOs.

As shown by Hamilton and McManus (2012), this heterogeneity is very relevant. These authors document that "universal mandates" (those requiring all insurers to cover ART) lead to a substantial increase in IVF utilization while other types of insurance mandates have a smaller effect. Consistent with this, studies focusing on the impact of the mandates on different outcomes (see, inter alia, Kroeger and La Mattina (2017), Machado and Sanz-de Galdeano (2015) or Schmidt (2007)) have found larger effects in states with universal mandates. Therefore, we only include in the treatment group those states that enacted mandates to cover IVF treatment and that applied to all insurers. This means that our group of treated states includes the following ones: Connecticut, Hawaii, Illinois, Maryland, Massachusetts, New Jersey, and Rhode Island.

In addition, we eliminate variation in the timing in which the mandates were enacted by pooling together all states that passed reforms between 1980 and 1990, which excludes New Jersey and Illinois from the treated group.[3] As a result of this rule, we are left with five treated states which are listed for convenience in the last column in Table 1.

Lastly, there are some US metropolitan areas which belong to several states, such as Boston. In these cases, we consider that a city is treated if at least some part of the metropolitan area belongs to a state in our treated group. The rationale for this choice is that we think it is likely that residents in parts of the metropolitan area belonging to other states were also affected by the policy, as metropolitan statistical areas have a high degree of economic and social integration. Regarding the control group, it is composed of all states that never enacted any kind of infertility insurance mandates or did it after the 90's.[4]

Finally, our identification strategy requires that assignment to the treatment is exogenous, that is, that the enactment of the mandates in some states did not respond to a greater demand for infertility insurance by the population. As mentioned earlier, these concerns have been addressed in Bitler and Schmidt (2012) and Hamilton and McManus (2012). Both studies show that state differences in the enactment were due to the electorate's view toward mandates in general. More concretely, Hamilton and McManus (2012) show that states that enacted mandates regarding other health issues (such as colorectal cancer screenings, Medicaid funding of abortions, and mental health parity) also adopted regulations for IVF. Thus, it is reasonable to think that state adoption of infertility insurance mandates was due to residents' preferences regarding government intervention in healthcare markets as opposed to a larger demand for infertility insurance on its own. In addition, these authors found no pre-mandates differences across mandate and non-mandate states in ART intensity, measured by the number of clinics in the state or the number of treatments per 10,000 women aged 25-44 years.

[3] In our analysis, we use census data because it allows us to identify neighborhoods' location. However, since these data are only available every ten years, we include 1970 and 1980 in the pre-treatment period, and consider 1990, 2000, and 2010, as being part of the post-treament period.

[4] That is, we drop from our sample states that had a reform but are not in the treated group (all states included in Table 1 that are not marked as treated in the last column).

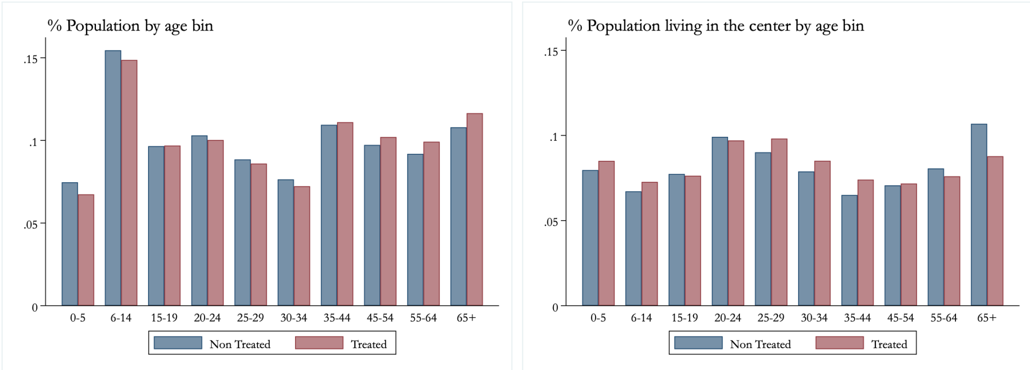
Table 2: Summary Statistics before Mandates

	Non Treated	Treated	Std. Diff
Avg City Population	2,766,916	6,798,305	-0.99
	(372,144)	(1,900,509)	
Avg City Household Income	21,431	22,332	-0.35
	(460)	(908)	
Avg City Housing Value	55,047	58,506	-0.27
	(2,325)	(4,335)	
% College Graduate in the City	10	10	-0.11
	(1)	(1)	
Downtown Household Income	15,298	14,658	0.29
	(582)	(753)	
Downtown Housing Value	41,113	40,810	0.02
	(3,872)	(7,272)	
Downtown % College Graduate	9	7	0.38
	(1)	(2)	
Number of observations	51,469	30,660	.
	(.)	(.)	

Notes: This table displays city averages regarding some relevant characteristics in treated and non-treated states in 1980. Standard errors are in parenthesis. The last column shows standardized mean differences for each reported variable.

Comparing treated and non-treated States Table 2 summarizes the main descriptive statistics of cities in treated and non-treated states before the mandates were introduced. As can be observed, average city size was considerably larger in treated states. Although both groups of states include a similar number of large cities, non-treated states host a greater number of small cities, which results on a large difference in means. It is also noteworthy that, while cities in treated states were richer both in terms of household income and average housing value, their city centers were poorer than those in the control group. Given that differences in city income and population across groups were substantial at the time of the reform, we control for these characteristics in all our regression specifications.

Lastly, Figure 2 shows that the average age distribution before the reform was quite similar across treated and non-treated states. This can be seen in Panel 2a, which displays the percentage of population by age bin in treated and non-treated cities prior to the mandates. In addition, both groups of states exhibited a similar spatial distribution of individuals for a given age group. Panel 2b in Figure 2 documents the absence of statistically significant differences in the percentage of individuals that live downtown within each age group in 1980 between treated and non-treated cities.



(a) City

(b) Downtown

Figure 2: Age Distribution before Mandates

Notes: This figure displays the age distribution of treated and non-treated cities in 1980, before mandates were enacted. Panel 2a displays the percentage of population in treated and non-treated cities by age bin. Panel 2b shows the percentage of individuals that locate downtown within each age bin.

4 Econometric specification

In this section, we explore the impact of adopting insurance mandates on standard measures of gentrification, i.e., changes in income and the percentage of college graduates in downtown neighborhoods. As already mentioned, we are interested in understanding whether the lower cost of postponing childbirth has an effect on the faster income growth at the city center and on the location patterns of high-skilled individuals. If the cost of living downtown increases with the presence of small kids, delaying childbirth would allow couples to live downtown for a longer period of time. Therefore, postponing both childbirth and moving to the suburbs would lead to a higher presence in downtown areas of young married couples on a more advanced stage in their professional career paths, therefore implying a higher average income of the residents in the city center. This demographic group not only increases median income at the city center but their preferences also contribute to endogenously increasing the supply of amenities, making room for further waves of gentrification. Moreover, as the group of people who prefer to live downtown expands, the demand for downtown housing rises, leading to higher sorting of individuals on income, since more wealthy households can afford higher rents.

4.1 Triple-Difference Specification

In order to estimate the causal effect of insurance mandates on gentrification, we employ a triple difference specification. The first difference is taken between the pre-treatment and the post-treatment period. As explained in section 3, we consider that the post-treatment period starts in 1990 for all treated states. The second difference is taken between treated and non-treated states. It captures how different was the change in the variable of interest in census tracts that were treated versus those that were not treated between the pre-treatment and the post-treatment periods. The third difference is taken between being part of the city center or of the suburbs. Hence, this triple difference captures how different was the change in the outcome variable of interest: (i) before and after treatment date, (ii) between the city center, and (iii), between the suburbs in the treated states compared to non-treated states. The general form of the regression we run is:

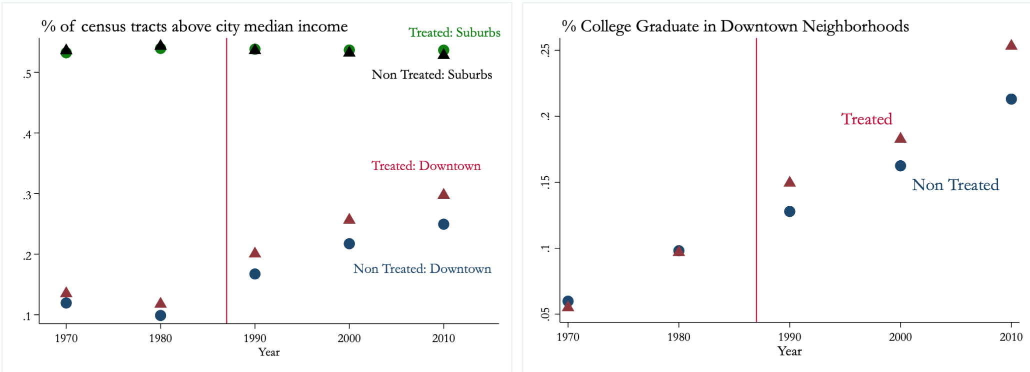
$$\begin{aligned} y_{i,t} = \alpha &+ \beta_1 Treated_i + \beta_2 Post_t + \beta_3 Center_i + \beta_4 Treated_i \times Post_t \\ &+ \beta_5 Treated_i \times Center_i + \beta_6 Post_t \times Center_i \\ &+ \beta_7 Treated_i \times Post_t \times Center_i + \beta_8 X_{i,t} \\ &+ \phi_{State(i)} + \delta_{State(i)} \times t + \psi_{CitySize(i)} + \gamma_{CitySize(i)} \times t + \epsilon_{i,t}, \end{aligned} \tag{1}$$

where $y_{i,t}$ is the outcome of interest for a census tract i at time period t. There are three indicator variables: $Treated$, which takes value one for those states which enacted an insurance mandate between 1980 and 1990; $Post$, which takes value one for periods after 1980, both for treated and non-treated states; and $Center$, which takes value one if the census tract is within the radius around the city center which contains 10 percent of the population of the city, $X_{i,t}$ includes controls, which vary at the census tract and year level. Finally, we include state and city size fixed effects (ϕ and ψ), as well as state and city size time trends (δ and γ).

Recall that throughout the analysis we only keep observations in cities with more than one million inhabitants, since gentrification is a large-city phenomenon. Moreover, we control for city size and time trends by city-size category to ensure that the results are not driven by a different presence of large vs small cities in the treated states.

4.2 Parallel trends

Our identification strategy relies on the existence of parallel trends in the outcomes of interest before mandates were introduced. Figure 3 displays the evolution of income by area of the city and the location patterns of college-graduates within the city in treated and non-treated cities over our period of study. Panel 3a displays the percentage of census tract whose income is above median income in the city both downtown and in the suburbs. Panel 3b shows the percentage of college graduates in downtown neighborhoods. Both panels support the existence of parallel trends before the dates of the mandates which, together with the exogenous policy enactments, allows us to interpret the estimates in a causal fashion.



(a) Income

(b) College Graduates

Figure 3: Parallel Trends

Notes: The left panel of this figure displays the percentage of census tract in each area of the city with income above median city income in treated and non-treated cities over time. The right panel of this figure displays the percentage of the city college-graduates that live in the center of the city. The red line signals the time in which infertility insurance mandates were enacted.

4.3 Confounding drivers of gentrification

In this section, we investigate whether our treatment may be correlated with other known drivers of gentrification. Even though, as we show in the previous section, treated and non-treated states were on similar gentrification trends at the time of the treatment, it is possible that they differed in some underlying characteristic that would have led to divergent gentrification even in the absence of treatment. Recent studies have pointed out to changes in the spatial distribution of jobs and to growing income as the main drivers of gentrification (Couture and Handbury (2017), Couture et al. (2019), Edlund et al. (2015), Su (2018)).

Notice that these trends are widespread across all US cities and, hence, should not affect our estimation. In principle, there are no reasons to think that cities that are located in states that enacted infertility insurance mandates experienced larger changes in the spatial distribution of jobs or greater income growth, except as a consequence of the policy itself. The confounding factor that is most worrying is the presence of long-hour occupations since this could also be related to a greater incentive for women to insure against infertility treatments. Thus, we explore this last factor in more detail.

Moreover, we include the (log) median income of the city and the share of jobs within 3 miles distance from the census tract as additional controls ($X_{i,t}$) in our regression to ensure that our estimates are not driven by alternative explanations in the literature. Further, we control for the share of college graduates in the city, which is very related to gentrification and could have been increasing faster in treated cities for reasons unrelated to the policy. An important concern when controlling for these variables is that the policy itself may have affected them directly, leading to problems of endogeneity. For instance, postponing the arrival of children is associated with positive effects on female wages, which in turn could affect average income in the city. While this is likely to have occurred, we believe these effects are second order and do not pose a challenge to our estimates. Some support for this conjecture is provided by the results reported in Table 13 in Appendix A, where is is shown that excluding these controls from our regressions hardly changes our main findings.

Presence of long-hour occupations. Recent studies have pointed to the rise in the returns to working long hours as an important driver of gentrification (Edlund et al. (2015), Su (2018)). Long-hour occupations create incentives to have a shorter commute, driving demand for downtown locations. The potential threat arises because women in long-hour occupations have a higher incentive to postpone maternity. Therefore, states with a larger fraction of employment in long-hour occupations may have also been more likely to implement insurance mandates.

To test for this possibility, we employ the measure developed by Cortés and Pan (2019) to classify occupations into terciles according to returns to long-hours, both in 1980 and in 2010. Once we have divided occupations into high, middle and low returns to long-hours, we check whether their presence differed across treated and non-treated states. In Table 3, we report the share of male and female employment in occupations with high, medium, and low returns to working long hours by city. We show that the occupational composition of treated and non-treated states in 1980 is not statistically significant. The difference is not significant regardless of whether we compute returns to working long hours in 1980 or in 2010. Therefore, we conclude that the occupational composition of states cannot explain the diverging gentrification.

Table 3: Employment composition by returns to working long hours

		Share of employment					
		Returns in 1980			Returns in 2010		
		Control	Treated	Diff.	Control	Treated	Diff.
		(1)	(2)	(3)	(4)	(5)	(6)
	High Ret.	0.17 ***	0.18***	0.01	0.62***	0.62***	0.00
		(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)
Male employment	Middle Ret.	0.64***	0.65***	0.01	0.20***	0.23***	0.02
		(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)
	Low Ret.	0.20***	0.17***	-0.02	0.20***	0.17***	-0.02
		(0.03)	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)
	High Ret.	0.18***	0.17***	-0.00	0.39***	0.40***	0.00
		(0.03)	(0.03)	(0.05)	(0.05)	(0.04)	(0.07)
Female employment	Middle Ret.	0.42***	0.43***	0.01	0.07***	0.08***	0.00
		(0.03)	(0.02)	(0.04)	(0.02)	(0.02)	(0.03)
	Low Ret.	0.40***	0.39***	-0.01	0.40***	0.39***	-0.01
		(0.05)	(0.02)	(0.06)	(0.05)	(0.02)	(0.06)

Notes: This table reports the share of employment in occupations by terciles of the return to working long hours. Columns (1) to (3) rank occupations based on the return to working long hours in 1980 and Columns (4) to (6). Returns to working long hours are calculated as the elasticity of log weekly earnings with respect to log hours with controls using the US Census, following the specification in Goldin (2014).

Table 4: Top-5 occupations by returns to working long hours

High returns in 1980	High returns in 2010
Business and Financial Operations	Business and Financial Operations
Economists and other Social Science Occs	Economists and other Social Science Occs
Arts, Design, Entertainment, Sports, and Media	Executive, Administrative, and Managerial Occs
Technicians, Paralegals and Pilots	Lawyers
Police and Detectives	Skilled Sales Occs

Notes: This table reports the top-5 occupations by the return to working long hours in 1980 and 2010. The consistent occupational classification is taken from Dorn (2009). The returns to working long hours are calculated as the elasticity of log weekly earnings with respect to log hours with controls using the US Census, following the specification in Goldin (2014).

It is important to notice that the occupations that reward more long-hours have changed over time. In Table 4 we present the top 5 occupations by returns to working long-hours. The top two occupations have not changed: "Business and Financial Occupations" as well as "Economists and Social Science Occupations" were and continue to be the ones in which working long hours pays off the most. However, in 1980 the next three occupations are different from the ones in 2010. For example, occupations related to arts, design, and entertainment used to be associated with long-hours in 1980 while in 2010 managers and lawyers have higher returns to long-hours. For our instrument to be valid, the treatment must be uncorrelated with alternative drivers of gentrification, that is, occupations that will become long-hour in 2010 and will drive gentrification.

5 Main results

5.1 Infertility insurance mandates and gentrification

We start by analyzing the effect of the insurance policy mandate on income at the city center. As explained in section 2, two complementary outcome variables are being used in this respect: (i) the probability that a census tract's income is above the median household income in the city and; (ii) the (logged) average income in the census tract. The first two columns in Table 5 display the results of running the triple differenced specification in equation 1 for each of the two above-mentioned variables. In what follows, we discuss the main findings for each row in Table 5 which correspond to the different interactions at play.

Table 5: The effect of infertility insurance mandates on gentrification

	Prob. above median	Log median income	% College Graduate
	(1)	(2)	(3)
Center	-0.457***	-0.637***	0.00128
	(0.0165)	(0.0131)	(0.00378)
Treated × Post	0.682***	0.378***	0.0612**
	(0.128)	(0.101)	(0.0293)
Center × Treated	-0.100***	-0.0189	-0.00376
	(0.0211)	(0.0167)	(0.00483)
Center × Post	0.105***	0.0834***	0.0157***
	(0.0196)	(0.0156)	(0.00450)
Treated × Center × Post	0.0733***	0.0348*	0.0289***
	(0.0260)	(0.0207)	(0.00597)
Observations	65733	65733	65733
Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
City Size FE	Yes	Yes	Yes
State Trends	Yes	Yes	Yes
City Size Trends	Yes	Yes	Yes

Notes: This table displays the impact of infertility insurance mandates on several measures of gentrification: (1) the probability that a census tract's income is above median income in the city; (2) the census tract's (log) median income; and (3) the percentage of college graduates in a census tract. Controls include: city's population, city's (log) median income, the share of college graduates in the city, and the share of jobs within 3 miles distance from the census tract. This table reports only selected coefficients, the full specification can be found in equation 1. Standard errors in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

The first row shows that census tracts that are located in the city center had lower income levels than the suburbs *before* the treatment took place, both for treated and non-treated cities. Tracts in the city center are 45.7 percentage points (p.p.) less likely to be above the median income and have on average 63.7 percent lower income.

The second row documents that census tracts outside the city center in treated states also became more affluent *after* treatment. Following the policy implementation, median income in these tracts was 68 percent higher while the probability of being above the median in the city increased in 38 p.p., which implies that the policy also led to favourable effects on income in the suburbs. This effect is in line with the findings in Kroeger and La Mattina (2017), who show that infertility mandates increased the percentage of women that entered professional occupations in treated states as compared to non-treated states, irrespective of whether they lived downtown or in the suburbs. Moreover, it is consistent with the literature that documents positive effects in women's wages from postponing maternity (Caucutt et al. (2002), Goldin and Katz (2002), Miller (2011), among others).

The third row reveals that census tracts in the city center (relative to the suburbs) in treated states were poorer than their analogous counterparts in non-treated states. Nevertheless, we do not find that this difference is important for our results.

In the fourth row we see that income in the city center also increased after the treatment in non-treated states. Downtown locations became 10.5 percentage points more likely to be above the median income of the city, and on average this meant a 8.3 percent rise in median income. This is most likely due to a general trend towards gentrification to different degrees everywhere. First, notice that the trend to postpone childbearing is broader than the delay induced by this policy and certainly not restricted to states that enacted infertility insurance mandates (as was illustrated in Figure 1). Therefore, the observed widespread growth in downtown income is fully compatible with the mechanism we propose in this paper. In addition, we think that our mechanism is compatible with other drivers of gentrification such as the changing spatial distribution of labor market opportunities or the rising importance of long working hours in high-skilled jobs.

The fifth row implies that center tracts in treated states had an even larger increase in income during the post treatment period compared to the increase in income experienced in non-treated states. In particular, center tracts' income became 7.3 percentage points more likely to be above the median income in the city, beyond the 68.2 percentage point increase of center tracts in non-treated states. This implied an average increase in downtown income of 3.5 percent on top of the 8.3 percent increase in non-treated states.

Next, we switch attention to another commonly used measure of gentrification: the percentage of college graduates that locate downtown. As reported in the third column in Table 5, the fraction of college graduates in a neighborhood: (i) was not significantly different in each area of the cities; (ii) increased 6 p.p. in treated cities after the mandates were enacted; (iii) was not different in the center of treated cities; and, (iv) had a positive trend downtown everywhere. More crucial to our analysis, college graduates were 2.9 p.p. more likely to locate downtown after the mandates in treated cities as compared to the same difference in non-treated cities. Therefore, the results regarding college graduates' location patterns are fully in line with the observed changes in income downtown, as expected.

To summarize our findings, the effect of the policy on income at the city center and the college graduates' location patterns is statistically significant and sizable in magnitude. As discussed in subsection 4.1, treated and non-treated states were on parallel trends leading up to the treatment year. Moreover, it seems reasonable to assume that state differences in the enactment of mandates were politically motivated, instead of being due to some underlying variable that could have increased income at the city centre faster than in the suburbs. For this reason, we believe there are sufficiently strong grounds for interpreting the estimated coefficients in a causal fashion.

5.2 Discussion of the Mechanisms

In this section, we first proceed to analyze the variation in the demographic composition of the city center supports our preferred mechanism whereby the effect of the policy on gentrification could result from women delaying having kids and staying downtown rather than moving to the suburbs. Next, we document the higher propensity of families to locate in the suburbs of the city with respect to non-families and highlight some suburban characteristics that make these areas more suitable for families.

One could think of two main reasons explaining why families disproportionately locate in the suburbs. First, the housing stock downtown is not ideal for the children. For instance, houses are too small and usually lack outdoor space. Second, school quality is known to be worse in central neighborhoods. Indeed, in Appendix A we show that houses tend to have a larger number of rooms in the suburbs than downtown. In addition, we include some illustrative evidence about differences in school quality between the suburbs and the center in some cities. Notice that, even if children could attend schools in any location in the city, there are clear advantages of attending nearby schools, as many extracurricular activities are closely linked to the school and other children are also likely to live close by.

5.2.1 Changes in the demographic composition

We claim that infertility insurance mandates extended the life period in which individuals benefit the most from locating downtown, fueling the process of gentrification. Therefore, we should observe a change in the demographic composition of central neighborhoods towards slightly older couples.

In order to capture this change, we restrict our attention to individuals with ages between 20 and 44 and examine their location choices. We focus on couples in childbearing age because these are the ones for which the timing of family formation influences their residential choices. Therefore, we run again equation 1 where this time the dependent variable is the percentage of individuals in each age bin of the census tract population who are in childbearing age (20-44). Figure 4 plots the coefficients of the triple interaction term, which displays the impact of the policy for each of our 4 age bins (20-24, 25-29, 30-34, and 35-39), both for males and females. That is, this figure displays the impact of ART mandates in the age distribution downtown relative to the suburbs in treated cities and compared to the same difference in non-treated cities.

Consistent with the idea that postponing childbearing allows couples to reside in the city centre until later stages of their lifetimes, we find that the policy leads to around a 2 percentage points increase in the proportion of adults aged between 20 and 30 living downtown, while in parallel the percentage of older adults goes down. Interestingly, the proportion of women postponing maternity is a bit younger than men. This is consistent with women delaying having kids until the early thirties and moving out of the city centre with their partners and kids afterwards. Since, male partners tend to be a little older, the effect is delayed for men. In line with our mechanism, the proportions of men and women aged between 35 and 40 living downtown decrease by around 2 percentage points.

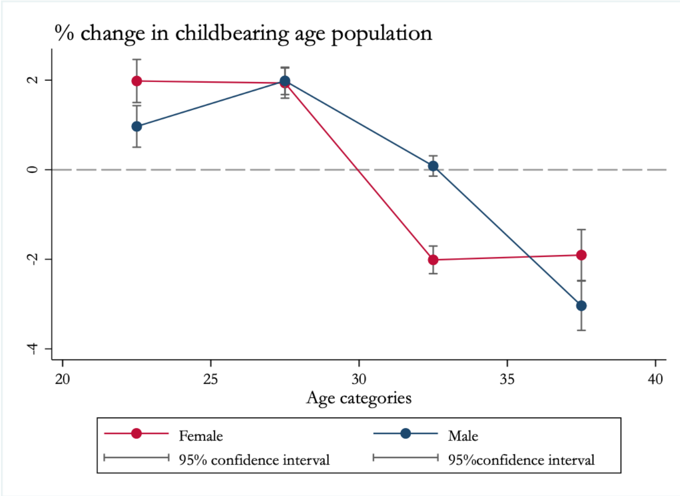


Figure 4: Age composition males vs females

Notes: This figure shows the change in the age distribution of female and male individuals in downtown neighborhoods as compared to the suburbs in treated cities vs non-treated cities. That is, it plots the coefficients of the triple difference of running equation 1 for the percentage of females or males in four different age categories.

5.2.2 The location patterns of families

Table 6 shows the result of running an analogous triple difference regression to the one considered earlier, this time using the percentage of families in a given neighborhood as the dependent variable. The only change with respect to equation 1 is that we now place the focus of the subsequent discussion on the suburbs (rather than the center) as the reference location category. Both specifications are equivalent, we only choose the one presented in Table 6 for ease of exposition.

The first row in Table 6 shows that the percentage of families in a suburban neighborhood is 22.7 p.p. larger than downtown, while the third row reflects the substantial decline in fertility and family formation after the 80's. Notice that the estimated coefficients on the indicator variable for the treatment and its interaction with the post-treatment period and with the suburbs dummy are all non significant. This indicates that there were no differences between treated and non-treated cities in the overall percentage of families before and after the mandates nor in their location patterns.

Turning our attention to the last two rows of Table 6, we observe that while the percentage of families in suburban neighborhoods decreased by almost 5 p.p. in non-treated cities after the 80's, this decline was limited to about 1 p.p. in treated cities. The decreasing trend in the percentage of families that live in the suburbs could be due to the end of suburbanization, meaning that downtowns are

Table 6: Families' location patterns

	% Families in the census tract
	(1)
Suburbs	0.227***
	(0.00321)
Treated	-0.00202
	(0.00653)
Post	-0.212***
	(0.0224)
Treated × Post	-0.00285
	(0.00798)
Suburbs × Treated	-0.00714
	(0.00465)
Suburbs × Post	-0.0485***
	(0.00367)
Treated × Suburbs × Post	0.0321***
	(0.00560)
Observations	73387
Controls	Yes
Year FE	Yes
State FE	Yes
City Size FE	Yes
State Trends	Yes
City Size Trends	Yes

Notes: This table shows the effect of the policy in the location patterns of families. The dependent variable is the percentage of family households in a census tract. Controls include: city's population and city's (log) median income, and the share of jobs within 3 miles distance from the census tract. This table reports only selected coefficients, the full specification can be found in equation 1, replacing the dummy for center for the complementary dummy for suburbs. Standard errors in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

increasingly providing some of the living facilities that families used to find in the suburbs. Moreover, it could be that at least some families are changing their location patterns as downtowns become more attractive. Lastly, the fact that this change is considerably lower in treated cities points towards some difficulties to locate in thriving downtowns among families in these locations, as treated cities experienced gentrification to a larger extent.

Therefore, the evidence reported in Table 6 illustrates several patterns that are relevant for our analysis. First, it shows that families are much more likely to locate in the suburbs than non-family households over the last decades. Second, it confirms a severe fall in family formation, which we consider to be a key element in explaining the gentrification process. Lastly, it highlights that the shift in the location patterns of families towards downtown neighborhoods is more contained in cities experiencing a higher degree of gentrification.

5.3 Delayed childbearing and gentrification: an IV approach

So far, we have shown that differences in infertility insurance mandates across US states had a significant effect on the gentrification of downtown neighborhoods. Moreover, we have provided evidence consistent with the postponement of maternity mediating the effect of these policies on the spatial distribution of income. However, the rise in the age at first-time mother is a much more general trend that is not solely related to ART mandates. As illustrated in Figure 1 above, starting in the late 70's, the age at first birth has gone from just over 21 to close to 25 years old in 2010. Over the same period, the probability that a census tract in the city center had an average income above the median income in the city went from 5% to 25%. To relate both trends in a causal fashion, in this section we use the mandates as an instrumental variable for the average age at first birth in the city when estimating the impact of the latter variable on gentrification. In addition, we provide some preliminary quantitative assessment of the overall effect that a delayed age at first birth could have on the relative income growth of city centers and suburbs over the sample period under consideration.

The choice of an IV approach in this exercise is dictated by the following reasoning. As pointed out above, downtown neighborhoods tend to be wealthier in cities in which women have their first kid at an older age. However, the direction of causality is unclear. In particular, it could be the case that as gentrification gets stronger (because central areas of the city become more attractive due to shorter commuting times or increased density of amenities), women reacted by postponing having children and moving to the suburbs, leading in this way to reverse causality. Thus, in order to estimate the causal effect of age at first birth on gentrification, we use the ART policy enactment to instrument the average age at first birth in a city, on the basis that the approval of these policies across different states is unrelated to the specific preferences of their populations about delayed fertility treatments (see discussion on this issue in Section 3 above). The identifying assumption in this empirical strategy is that the mandates affected gentrification only by affecting the age at first birth, but not directly.

Our specification requires that we run our regressions at the city level instead of using census tract as done in our previous analysis. We obtain the age of first-time mothers at the county level from the National Center for Health Statistics (NCHS) Natality Birth Data. We then construct a measure of gentrification at the city level by dividing the average income in central counties by the average income in suburban counties. Therefore, an increasing income ratio will be indicative of gentrification happening in the city.

$$Gentrification_{i,t} \equiv \frac{\sum_{j \in Downtown} MeanIncome_j}{N_{downtown}} / \frac{\sum_{j \in Suburbs} MeanIncome_j}{N_{suburbs}} \tag{2}$$

Since the instrument we employ is a binary instrument, we use the Wald estimator, also known as the grouping estimator. The estimator is implemented through the following three steps. First, we regress our measure of gentrification city i at time t on the instrument and controls:

$$\begin{aligned} Gentrification_{i,t} = & \alpha_0 + \alpha_1 Treated_i + \alpha_2 Treated_i \times Post_t + CitySize_i \\ & + \mu_t + \phi_{State(i)} + \delta_{State(i)} \times t + \epsilon_{i,t}, \end{aligned} \tag{3}$$

where μ and ϕ are time and state fixed effects, and δ are state trends and we also control for city size. Next, we run a similar regression for the average age of mothers at their first birth in city i at time t:

$$\begin{aligned} AgeFirstBirth_{i,t} = & \beta_0 + \beta_1 Treated_i + \beta_2 Treated_i \times Post_t + CitySize_i \\ & + \mu_t + \phi_{State(i)} + \delta_{State(i)} \times t + \nu_{i,t}, \end{aligned} \tag{4}$$

where the time, state, and state trends are denoted with the same symbols as before for comparability. Finally, we combine both estimates together to obtain the Wald estimator, which captures the effect of age at first birth on gentrification, instrumented with the insurance mandate.

$$\widehat{W} = \frac{\hat{\alpha}_2}{\hat{\beta}_2}, \qquad SE_{\widehat{W}} = \frac{\hat{\alpha}_2}{\hat{\beta}_2} \sqrt{\left(\frac{SE_{\hat{\alpha}_2}}{\hat{\alpha}_2}\right)^2 + \left(\frac{SE_{\hat{\beta}_2}}{\hat{\beta}_2}\right)^2} \tag{5}$$

The results of the first and second step are included in Table 7. We find that the policy increased the age at first birth by 0.62 years, and the ratio of downtown to suburb income goes up by 1.5 percentage points. For reference, the average ratio of downtown to suburb income in our sample is 57 percent with a standard deviation of 15 percentage points. Moreover, the average mean age at first birth is 24 years with a standard deviation of 2 years. If the exclusion restriction holds, this implies that for each year of increase in the average age at first birth, cities should expect the ratio of income downtown to income in the suburbs to increase by 2.4 percentage points.

The magnitude of the estimated effect points towards a potentially large economic significance. A back-of-the-envelope calculation tells us that the increase in the age at first birth from 23.27 in 1980 to 24.7 in 2000 could explain an increase in the income ratio of downtown to suburb of 5.8 percentage points. The average increase in the ratio from 1980 to 2000 was 3.7 percentage points. Of course, there are many other mechanisms working at the same time and we do not claim that all of the increase in age at first birth is exogenous, nor that it is the sole driver of gentrification. However, these results are suggestive of the potential central role that a delay in the age at first birth may have played in explaining gentrification.

Table 7: Causal effect of Age at 1st Birth on Gentrification

	First step	
	Gentrification	Age at 1st birth
	(1)	(2)
Treated × Post	0.015***	0.618***
	(0.0021)	(0.0255)
Observations	72737	72737
Year FE	Yes	Yes
State FE	Yes	Yes
State time trends	Yes	Yes
City Size FE	Yes	Yes
	Second step	
	Gentrification	
	(1)	
Age at 1st birth	0.024***	
(IV: Treated × Post)	(0.0035)	

Notes: This table displays the impact of delayed maternity on gentrification using a Wald estimator. The top panel displays the results of the first step regressions while the bottom panel displays the result of the second step regression. Standard errors in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

6 A model of fertility timing and location choice

In this section, we propose a model of endogenous fertility timing and location choice. The goal of the model is to quantify how much of the observed gentrification can be explained by the delay in the age at first birth. In the previous section, we provided causal estimates for the effect of the policy change. However, the delay in childbearing has been more widespread than as the result of the policy. Therefore, we plan to employ the causal estimate to quantify the model and perform a counterfactual in which we change the incentives to delay maternity in a way that replicates the changes in the data and observe its impact on gentrification.

6.1 Model setup

Geography. The geography in this economy consists on a set of locations indexed by $l = \{1, ..., N\}$. There are two types of locations: downtown (d) and suburbs (s). There is one downtown location and $N_s = N - 1$ suburb locations. The suburb locations are identical except for an idiosyncratic amenity that agents derive from living in a particular location. The suburb locations differ from the downtown location in three dimensions: an amenity, income, and housing supply.

We assume that there is no cost to move across locations and that amenities are local, they can only be consumed by residing in the location. Moreover, we assume that income is location dependent, which can be interpreted as the optimal income that an agent can receive given that she lives in a given location and commutes to the best possible jobb. Thus, the commuting cost would then be incorporated in the amenity of a location. There is free trade of the final good which is used as the numeraire. The housing supply in each location is fixed and owned by absentee landlords.

Households. The economy is inhabited by a mass of households indexed by i. Households are composed by a couple and live for three periods. The first period they are young (y), then mature (m), and finally old (o), we let $a \in \{y, m, o\}$ index age. There are two types of households depending on their skill, z: high-skilled and low-skilled. Households choose where to live, and whether and when to have children.

They derive an idiosyncratic utility from residing in location l. The vector of idiosyncratic preferences, $\varepsilon^i = \left\{\varepsilon^i_{l,a}\right\}$, is distributed as a Gumbel with scale parameter β_l which is independently distributed across types and time periods. Moreover, if households choose to have children, they derive an idiosyncratic utility ε^i_k which is distributed as a Gumbel with scale parameter β_k. This idiosyncratic amenity is enjoyed only in the period in which the children are born and it captures the present value of the idiosyncratic amenity associated to the presence of children.

Children. Children remain in their household for two periods. We define a kid state variable k which captures the presence and age of kids. Let $k = 0$ if there are no kids present, $k = 1$ if the household had kids that period, and $k = 2$ if the household had kids the previous period. Conditional on wanting children, there is a probability that the household will be successful. For simplicity, we assume that the probability is equal to one when young, $\rho_y = 1$, zero when old $\rho_o = 0$, and a number between zero and one when mature $0 < \rho_m < 1$. Both amenities, $\delta(a, k, z, l)$, and income, $I(a, k, z, l)$, depend on the presence and age of kids, k, in addition to age, a, skill z, and location of residence l. This flexibility on the amenities is meant to capture location-specific amenities related to kids, for example, the availability of high-quality schools or the proximity to parks. The flexibility on income can capture child penalty effects and how they vary with the timing of kids. This penalty could be related to changing jobs in order to gain more flexible hours or a shorter commute.

Preferences. An agent i at age a with kids aged k, skill z, and living in location l derive a period utility from the consumption of final output, c, and housing according to:

$$\begin{aligned} U^i(c, h; a, k, z, l) =& c^{1-\alpha} h^{\alpha} + \delta(a, k, z, l) + \varepsilon_l^i + \varepsilon_k^i D_{k=1} \\ \text{subject to: } & c + p_l h = I(a, k, z, l), \end{aligned}$$

where $D_{k=1}$ is a dummy taking value one on the period the household has kids, and p_l is the housing price in location l. Agents apply discount factor ϕ to future periods. They are myopic when forming expectations about income, prices and amenities.

Timing. When couples are born, they draw an idiosyncratic preference for children ε_k^i which stays constant for their lifetime. Each period they first observe whether they had kids the previous period. If they do not have kids yet, they can decide whether to try to have kids that period. They then observe if they are successful in having kids. Once they have discovered their kid state this period, they draw an idiosyncratic preference for location, ε_l^i and choose where to live, they consume and produce. The key timing assumption is that agents observe their location preference only after having made the decision on whether to have children. Otherwise, it is possible that the preference for location affects the fertility decision. We believe it is a reasonable assumption since the preference for location are more likely to change quickly and unexpectedly while the childbearing decisions are more permanent. Therefore, we are capturing that at least some of the taste shocks related to location are realized only after making decisions on whether and when to have children.

6.2 Definition and characterization of equilibrium

Location choice. A household i of age a, kids aged k, skill z, and an idiosyncratic preference vector ε^i chooses the optimal location in order to solve the following problem:

$$\tilde{v}\left(a,k,z,\varepsilon^i\right)=\max_l\left\{x\left(a,k,z,l\right)+\delta\left(a,k,z,l\right)+\varepsilon^i_l\right\}+\varepsilon^i_k D_{k=1},$$

where $x\left(a,k,z,l\right)=\alpha^{\alpha}\left(1-\alpha\right)^{1-\alpha}\frac{I(a,k,z,l)}{p_l^{\alpha}}$ is the observed component of the indirect utility from living in location l, which is common to all individuals of the same demographic group.

Given the assumption that ε^i_l is distributed as a Gumbel with scale parameter β_l, we can obtain the fraction of households that will choose to live in the city center in a given period:

$$\pi\left(c|a,k,z\right)=\frac{exp\left\{\frac{x(a,k,z,c)+\delta(a,k,z,c)}{\beta_l}\right\}}{exp\left\{\frac{x(a,k,z,c)+\delta(a,k,z,c)}{\beta_l}\right\}+N_s exp\left\{\frac{x(a,k,z,s)+\delta(a,k,z,s)}{\beta_l}\right\}}, \tag{6}$$

where N_s is the number of locations of type suburb. Since all locations in the suburbs are identical, we employ $d=1$ for the location index of the downtown location and $s\in\{2,...,N_s+1\}$ for the index of any of the suburb location.

We can now define the expected optimal utility for an agent of age a, kids of age k and skill z as:

$$v^*\left(a,k,z\right)=E_{\varepsilon_l}\tilde{v}\left(a,k,z,\varepsilon_l\right)+\varepsilon^i_k D_{k=1},$$

where the expectation is taken over the idiosyncratic preference for location.

Fertility choice. Households can choose whether to have children, and whether to have them as young, or postpone and try to have them as mature. Notice that this decision is taken at the beginning of the household's life and there is no reason to deviate later on. In order to make this decision, household i compares the discounted present utility from the three possible outcomes, not taking into account the idiosyncratic preference. Let's define the utility from having kids as young, v^*_{ky}, as mature v_{km}^*, and from no kids v_{nk}^* as:

$$\begin{aligned}
v_{ky}^*(z) &= v^*(y, k=1, z) + \phi v^*(m, k=2, z) + \phi^2 v^*(o, k=0, z),\\
v_{km}^*(z) &= v^*(y, k=0, z) + \phi v^*(m, k=1, z) + \phi^2 v^*(o, k=2, z),\\
v_{nk}^*(z) &= v^*(y, k=0, z) + \phi v^*(m, k=0, z) + \phi^2 v^*(o, k=0, z).
\end{aligned}$$

Remember that households cannot simply choose to have children as mature, since there is a probability they will not be successful. Households choose whether and when to try to have children. The idiosyncratic utility derived from each possible action is given by:

$$\begin{aligned}
V\left(\text{kids young}|z, \varepsilon_k^i\right) &= v_{ky}^*(z) + \varepsilon_k^i,\\
V\left(\text{try kids mature}|z, \varepsilon_k^i\right) &= \tilde{\rho}_m\left(v_{km}^*(z) + \phi\epsilon_k^i\right) + (1-\tilde{\rho}_m)\, v_{nk}^*(z)\\
V\left(\text{no kids}|z, \varepsilon_k^i\right) &= v_{nk}^*(z),
\end{aligned}$$

where $\tilde{\rho_m}$ is the perceived probability of being successful, that is, the probability that households take into account when making the decision. We distinguish it from the actual probability to allow for households beliefs to differ from the truth. The optimal choice depends on the idiosyncratic preference for children ε_k^i. We can define three thresholds per skill type z as the three levels of idiosyncratic preference that makes individuals indifferent among each pair of choices:

$$\begin{aligned}
\bar{\epsilon}_{ky,km}(z) &= \frac{\tilde{\rho}_m v_{km}^*(z) + (1-\tilde{\rho}_m)\, v_n^*(z) - v_{ky}^*(z)}{1-\phi\tilde{\rho}_m},\\
\bar{\epsilon}_{ky,nk}(z) &= v_{nk}^*(z) - v_{ky}^*(z),\\
\bar{\epsilon}_{km,nk}(z) &= \frac{v_{nk}^*(z) - v_{km}^*(z)}{\phi}.
\end{aligned}$$

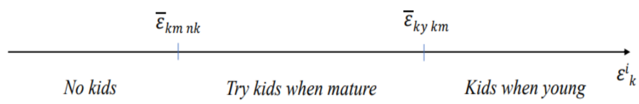
The thresholds are such that for households of skill z it holds that:

$$\begin{aligned}
\varepsilon_k^i > \bar{\epsilon}_{ky,km}(z) &\implies V\left(\text{kids young}|z, \varepsilon_k^i\right) > V\left(\text{try kids mature}|z, \varepsilon_k^i\right)\\
\varepsilon_k^i > \bar{\epsilon}_{ky,nk}(z) &\implies V\left(\text{kids young}|z, \varepsilon_k^i\right) > V\left(\text{no kids}|z, \varepsilon_k^i\right)\\
\varepsilon_k^i > \bar{\epsilon}_{km,nk}(z) &\implies V\left(\text{try kids mature}|z, \varepsilon_k^i\right) > V\left(\text{no kids}|z, \varepsilon_k^i\right)
\end{aligned}$$

We focus on equilibria in which at least some people of both skill types have kids when mature. For this to be the case, it must be that there exists a region of idiosyncratic preferences for which having kids as mature dominates. This will happen whenever the following condition holds:

$$\bar{\epsilon}_{km,nk}(z) < \bar{\epsilon}_{ky,km}(z).$$

Whenever this is the case, there will be only two active thresholds, $\bar{\epsilon}_{km,nk}(z)$ and $\bar{\epsilon}_{ky,km}(z)$, and the optimal decision will be:



Now we can compute the fraction of people of each skill z who choose each of the three options:

$$\begin{aligned} \pi_k(\text{kids young}; (z)) =& 1 - F_k(\bar{\epsilon}_{ky,km}(z)) \\ \pi_k(\text{try kids mature}; (z)) =& F_k(\bar{\epsilon}_{ky,km}(z)) - F_k(\bar{\epsilon}_{km,nk}(z)) \\ \pi_k(\text{no kids}; (z)) =& F_k(\bar{\epsilon}_{km,nk}), \end{aligned} \tag{7}$$

where F_k is the distribution of the idiosyncratic preference for children and it follows a Gumbel distribution with scale parameter β_k.

Housing market. Housing supply is fixed at H_l in each location. All the suburb locations will have an identical housing supply, and may differ from the housing supply in the downtown location. The housing price is such that the housing market will clear in each location. Namely,

$$H_l = \sum_a \sum_k \sum_z \frac{\alpha I(a,k,z,l)}{p_l}, \forall l \tag{8}$$

Decentralized spatial equilibrium An equilibrium for this economy is a distribution of households across locations $\pi(l|a,k,z)$, across fertility choices $\pi_k(\text{kids young}; z)$, $\pi_k(\text{try kids mature}; z)$, $\pi_k(\text{no kids}; z)$ and a housing price for each location p_l such that:

1. Households optimally choose location (Eq. 6).

2. Fertility choices are optimal (Eq. 7).

3. Housing prices p_l are such that all local housing markets clear (Eq. 8).

7 Model Quantification

7.1 Data and Definitions

The quantification of the model employs Census individual-level data: for the years 1980, 1990, and 2000. The geographic unit in the census data vary in each year. We employ the smallest unit available in each year. For 1980 we use county groups, and Public Use Microdata Areas for the remaining years. We select only couples in our data and treat each household as an individual agent in the model. Households are assigned to groups according to their age, fertility and location. For each group we compute the population, and the average income net of racial and nativity disparities and of city fixed effects.

The age of a household is assigned based solely on the age of the female. Households between the ages fo 20 and 30 are classified as young, between 30 and 40 as mature, and above 40 as old. This classification is meant to capture three distinct faces in the fertility phases. There are three fertility states: no kids, young kids if the household had them in the current age bin, or old kids if the household had them when in the previous age bin. Moreover, we assign households to the high-skill group if at least one of the household members completed a bachelor degree, and to the low-skill group otherwise.

Finally, we assign households to either the suburbs or downtown depending on the geographic unit where they live according to the following procedure. First, we establish the point location of a city center as in Fee & Hartley (2013). This measure is also used by Lee and Lin (2018) . For 268 MSAs, they identify the CBD using 1982 Census of Retail Trade for the central city of the MSA. For the remaining 117 MSAs, the center is found by geocoding the MSA's central city found using ArcGIS's 10.0 North American Geocoding Service. Second, we calculate the distance of each census tract to this point city center. We are grateful to Lee and Lin (2018) for making this measure readily available. Third, using this measure of distance to the center we classify census tracts in a city as downtown or suburbs for the year 2000. In particular, we classify as downtown all census tracts which are closest to the center and include 10 percent of the population. We then take the census tracts which are furthest from the center and include 50 percent of the city population and classify them as suburbs. This results in a polygon for downtown and a polygon for suburbs for every city in our sample which is kept fixed. Fourth, we follow Couture and Handbury (2017) and classify a PUMA (or county group) as downtown if at least 60% of the PUMA's population belongs to census tracts classified as downtown. Finally, we select only cities for which we can accurately identify the downtown in all of our sample years (1980-2000). Following Couture and Handbury (2017), we consider that we can identify the center in cities for which, at least 60% of the population in the center lives in a PUMA (or county group) that is classified as downtown. Table 9 includes the cities that are included in this sample.

Table 8: Sample of cities

	Pop. Suburbs	Pop. Center
Albany-Schenectady-Troy,NY	499194	104683
Atlanta,GA	3353039	528514
Baltimore,MD	1645214	651428
GrandRapids,MI	785981	198126
Hartford-Bristol-Middleton,CT	585653	123090
Miami-Hialeah,FL	1741699	380025
Minneapolis-St.Paul,MN	2475685	380610
SanFrancisco-Oakland-Vallejo,CA	3866281	779549
Seattle-Everett,WA	1773182	325867
Tampa-St.Petersburg-Clearwater,FL	2102771	284010
Youngstown-Warren,OH/PA	378682	106592

Notes: This table includes the cities that are used to estimate the model. These are the cities for which we can accurately identify the downtown in all of the years used in estimation. Following Couture and Handbury (2017), we consider that we can identify the center in cities for which, at least 60% of the population in the center lives in a PUMA (or county group in 1980) that is classified as downtown.

7.2 Estimation

Calibrated parameters There is a set of parameters that we quanitfy following the literature.

Table 9: Externally estimated parameters

Parameter	**Definition**	**Value**
α	Cobb-Douglas weight for housing	0.24
ϕ	Discount Factor	0.96
ρ_m	Prob. have kids when mature	0.66
β_l	Gumbel parameter idiosyncratic taste for location	0.3
β_k	Gumbel parameter idiosyncratic taste for children	0.3

Notes: This table includes the parameters of the model that are estimated externally.

Estimation of downtown amenities. The only parameters left to estimate are the amenities for each age, kid's age, skill, and location, $\delta(a,k,z,l)$. First, we normalize the suburbs amenity for couples with no kids: $\delta(a,k=0,z,s)=0$. For given β_l and β_k, the fraction of each group that lives in the city center allows us to estimate the difference in amenities between the center and the suburbs. Recall the probability of choosing center is:

$$\pi_c\left(a,k,z\right)=\frac{exp\left\{\frac{x(a,k,z,d)+\delta(a,k,z,d)}{\beta}\right\}}{exp\left\{\frac{x(a,k,z,d)+\delta(a,k,z,d)}{\beta}\right\}+N_s exp\left\{\frac{x(a,k,z,s)+\delta(a,k,z,s)}{\beta}\right\}}$$

Let denote $\Delta_l\left(a,k,z\right)$ the amenity downtown relative to the suburbs:

$$\Delta_l\left(a,k,z\right)=\delta(a,k,z,d)-\delta(a,k,z,s),\forall a,k,z.$$

Multiplying and dividing by $exp\left\{\frac{1}{\beta}\left(-\delta\left(a,k,z,s\right)\right)\right\}$:

$$\pi_c\left(a,k,z\right)=\frac{exp\left\{\frac{1}{\beta}\left(x\left(a,k,z,d\right)+\Delta\left(a,k,z\right)\right)\right\}}{exp\left\{\frac{1}{\beta}\left(x\left(a,k,z,d\right)+\Delta\left(a,k,z\right)\right)\right\}+N_s exp\left\{\frac{1}{\beta}\left(x\left(a,k,z,s\right)\right)\right\}}$$

Taking logs and re-arranging:

$$\Delta_l\left(a,k,z\right)=\beta_l\left[log\left(\pi_c\left(a,k,z\right)N_s\right)-log\left(1-\pi_c\left(a,k,z\right)\right)\right]+x\left(a,k,z,s\right)-x\left(a,k,z,d\right)$$

For a given β_l and N_s locations we can solve for these amenities $\Delta_l\left(a,k,z\right)$ since in the data we observe the fraction of each group that chooses to live downtown, $\pi_c\left(a,k,z\right)$. Table 10 presents the estimates of these amenities for 1980 and 2000.

Table 10: Amenity at center relative to suburbs, $\Delta_l\left(a,k,s\right)$

			1980			
		Low-skill			High-skill	
	Young	Mature	Old	Young	Mature	Old
No kids	0.3364	0.4867	0.3407	0.2857	0.3346	0.2724
Young kids	0.2623	0.2792	.	0.1767	-0.0529	.
Old kids	.	0.2372	0.3043	.	0.1348	0.2467
			2000			
		Low-skill			High-skill	
	Young	Mature	Old	Young	Mature	Old
No kids	1.4240	1.6166	1.5134	1.4635	1.7166	1.6507
Young kids	1.3246	1.5460	.	1.5988	1.3937	.
Old kids	.	1.6074	1.6982	.	1.8336	1.5600

Estimation of children amenities. From the previous step we can only compare the amenity of living downtown versus the suburbs but we cannot compare the amenity of, for example, living downtown with and without children. Therefore, we estimate next the amenity of having children using the observed fraction of people of a given skill that have kids as young. The first step is to back out the thresholds $\bar{\varepsilon}_{ky,km}$ and $\bar{\varepsilon}_{km,nk}$ from the observed fertility timing choices.

Recall that the fraction of people that have kids as young π_{ky}, as mature, π_{km}, and that do not to have kids π_{nk}, are given by:

$$\begin{aligned}
\pi_{ky}(z) &= 1 - F\left(\bar{\varepsilon}_{ky,km}(z)\right),\\
\pi_{km}(z) &= \rho_m(F\left(\bar{\varepsilon}_{ky,km}(z)\right) - F\left(\bar{\varepsilon}_{km,nk}(z)\right)),\\
\pi_{nk}(z) &= F\left(\bar{\varepsilon}_{km,nk}(z)\right) + (1-\rho_m)\left(F\left(\bar{\varepsilon}_{ky,km}(z)\right) - F\left(\bar{\varepsilon}_{km,nk}(z)\right)\right),
\end{aligned}$$

where F is the distribution of idiosyncratic preferences for children and it is assumed to be a Gumbel distribution with scale parameter β_k Thus, there are two thresholds and two equations[5]. We should be able to back out the thresholds from these conditions. For the observed choices $\pi_{ky}(z)$, $\pi_{km}(z)$, and $\pi_{nk}(z)$ and probability of having a successful pregnancy as mature, ρ_m, it is straightforward to back out the thresholds $\bar{\varepsilon}_{ky,km}(z)$, and $\bar{\varepsilon}_{km,nk}(z)$, as a function of skill, z, from these system of equations.

Next, we write the amenities to be estimated as a function of the thresholds. Recall the thresholds are given by:

$$\begin{aligned}
\bar{\varepsilon}_{ky,km}(z) =& \frac{\rho_m v^*_{km}(z) + (1-\rho_m)\, v^*_n(z) - v^*_{ky}(z)}{1-\phi\rho_m},\\
\bar{\varepsilon}_{ky,nk}(z) =& v^*_{nk}(z) - v^*_{ky}(z),\\
\bar{\varepsilon}_{km,nk}(z) =& \frac{v^*_{nk}(z) - v^*_{km}(z)}{\phi},
\end{aligned}$$

where,

$$\begin{aligned}
v^*_{ky}(z) &= v^*(y, k=1, z) + \phi v^*(m, k=2, z) + \phi^2 v^*(o, k=0, z)\\
v^*_{km}(z) &= v^*(y, k=0, z) + \phi v^*(m, k=1, z) + \phi^2 v^*(o, k=2, z)\\
v^*_{nk}(z) &= v^*(y, k=0, z) + \phi v^*(m, k=0, z) + \phi^2 v^*(o, k=0, z)
\end{aligned}$$

The first step is to characterize the expected period utility, $v^*(y, k, z)$ for a given age, age of kids, and skill before the agents know their location idiosyncratic choice.

[5] Notice that one equation is colinear, since $1 = \pi_{ky} + \pi_{km} + \pi nk$

The distributional assumption on the idiosyncratic preference for location implies that the optimal utility $v^*(a,k,z)$ is distributed Gumbel with parameters $\mu = \Phi_{a,k,z}$ and β_l. From the properties of the Gumbel we know the expectation is $\mu + \beta_l\gamma = \Phi_{a,k,z} + \beta\gamma$, where $\gamma \approx 0.58$ is the Euler-Mascheroni constant. Therefore, we can write:

$$\begin{aligned} v^*(a,k,z) &= E_{\varepsilon_l} \max_l \left\{ v(a,l,k) + \varepsilon_l^i \right\} \\ &= \beta log \left(\sum_l exp \left\{ \frac{x(a,k,z,l) + \delta(a,k,z,l)}{\beta} \right\} \right) + \beta_l \gamma \\ &= \delta(a,k,z,s) + \tilde{\Phi}_{a,k,z} + \beta\gamma \end{aligned}$$

where $\tilde{\Phi}_{a,k,z} = \beta log \left(exp \left\{ \frac{1}{\beta} \left(x(a,k,z,d) + \Delta_l(a,k,z) \right) \right\} + N_s exp \left\{ \frac{1}{\beta} \left(x(a,k,z,s) \right) \right\} \right)$ can be treated as data at this point since it is a combination of the calibrated parameters β_l and N_s, the previous estimation step, $\Delta_l(a,k,z)$, and the income and housing price data $x(a,k,z,l) = \Gamma \frac{I(a,k,z,l)}{p_l^\alpha}$.

Now we can re-write the expected utility of each fertility timing choice as:

$$\begin{aligned} v_{ky}^*(z) &= \delta(y,1,z,s) + \tilde{\Phi}_{y,1,z} + \phi \left(\delta(m,2,z,s) + \tilde{\Phi}_{m,2,z} \right) + \phi^2 \left(\delta(o,0,z,s) + \tilde{\Phi}_{o,0,z} \right) + \Psi, \\ v_{km}^*(z) &= \delta(y,0,z,s) + \tilde{\Phi}_{y,0,z} + \phi \left(\delta(m,1,z,s) + \tilde{\Phi}_{m,1,z} \right) + \phi^2 \left(\delta(o,2,z,s) + \tilde{\Phi}_{o,2,z} \right) + \Psi, \\ v_{nk}^*(z) &= \delta(y,0,z,s) + \tilde{\Phi}_{y,0,z} + \phi \left(\delta(m,0,z,s) + \tilde{\Phi}_{m,0,z} \right) + \phi^2 \left(\delta(o,0,z,s) + \tilde{\Phi}_{o,0,z} \right) + \Psi, \end{aligned}$$

where $\Psi = (1 + \phi + \phi^2)\beta\gamma$.

Now, we can substitute them into the equation for the thresholds. First, we employ the threshold of indifference between having no kids and having kids as mature $\bar{\varepsilon}_{km,nk}$:

$$\begin{aligned} \bar{\varepsilon}_{km,nk}(z) =& \frac{1}{\phi} \left(v_{nk}^*(z) - v_{km}^*(z) \right) \\ =& \delta(m,0,z,s) - \delta(m,1,z,s) + \tilde{\Phi}_{m,0,z} - \tilde{\Phi}_{m,1,z} \\ &+ \phi \left(\delta(o,0,z,s) - \delta(o,2,z,s) + \tilde{\Phi}_{o,0,z} - \tilde{\Phi}_{o,2,z} \right) \end{aligned}$$

If we normalize utility to the utility of having no kids in the suburbs, that is,

$\delta(a,0,z,s)=0$,[6] for each age and skill group, then:

$$
\begin{aligned}
\bar{\varepsilon}_{km,nk}(z) =& \frac{1}{\phi}\left(v^{*}_{nk}(z)-v^{*}_{km}(z)\right) \\
=& \tilde{\Phi}_{m,0,z}-\tilde{\Phi}_{m,1,z}+\phi\left(\tilde{\Phi}_{o,0,z}-\tilde{\Phi}_{o,2,z}\right)-\delta(m,1,z,s)-\phi\delta(o,2,z,s).
\end{aligned}
$$

Let $\delta^{LT}_{k}(m,z,s)=\delta(m,1,z,s)+\phi\delta(o,2,z,s)$ denote the lifetime utility of having kids in the suburbs when young.

$$
\delta^{LT}_{k}(m,z,s) = \tilde{\Phi}_{m,0,z}-\tilde{\Phi}_{m,1,z}+\phi\left(\tilde{\Phi}_{o,0,z}-\tilde{\Phi}_{o,2,z}\right)-\bar{\varepsilon}_{km,nk}(z).
$$

Therefore, the threshold of indifference between having no kids and having kids as mature allows us to estimate the lifetime utility of having kids in the suburbs, relative to no having kids in the suburbs, which we normalized to zero for every age and skill group.

Next, from the threshold of indifference between having kids as young and having kids as mature:

$$
\begin{aligned}
(1-\phi\rho_m)\,\bar{\varepsilon}_{ky,km}(z) &= \left(\rho_m v^{*}_{km}(z)+(1-\rho_m)\,v^{*}_{nk}(z)-v^{*}_{ky}(z)\right) \\
&= \phi\rho_m\delta^{LT}_{k}(m,z,s)-\delta^{LT}_{k}(y,z,s)+\tilde{\Phi}_{y,0,z}-\tilde{\Phi}_{y,1,z} \\
&+\phi\left(\rho_m\left(\tilde{\Phi}_{m,1,z}\right)+(1-\rho_m)\,\tilde{\Phi}_{m,0,z}-\tilde{\Phi}_{m,2,z}\right) \\
&+\phi^2\rho_m\left(\tilde{\Phi}_{o,2,z}-\tilde{\Phi}_{o,0,z}\right)
\end{aligned}
$$

where $\delta^{LT}_{k}(y,z,s)=\delta(y,1,z,s)+\phi\delta(m,2,z,s)$.

Let's define operator $\Delta_{k,k'}\tilde{\Phi}_{a,z}=\tilde{\Phi}_{a,k,z}-\tilde{\Phi}_{a,k',z}$ to re-write the expression as:

$$
\begin{aligned}
(1-\phi\rho_m)\,\bar{\varepsilon}_{ky,km}(z) &= \phi\rho_m\delta^{LT}_{k}(m,z,s)-\delta^{LT}_{k}(y,z,s)-\Delta_{1,0}\tilde{\Phi}_{y,z} \\
&+\phi\left(\rho_m\Delta_{1,0}\tilde{\Phi}_{m,z}-\Delta_{2,0}\tilde{\Phi}_{m,z}\right)+\phi^2\rho_m\left(\Delta_{2,0}\tilde{\Phi}_{o,z}\right)
\end{aligned}
$$

Then we can solve for the lifetime utility of having children as young in the suburbs relative to the lifetime utility of not having kids and living in the suburbs:

First, let's use the operator to re-write the expresion for the lifetime utility of

[6] This assumption is necessary because it is not possible to compare utility for different age and skill groups.

having kids in the suburbs as mature:

$$\bar{\varepsilon}_{km,nk} = \Delta_{0,1}\tilde{\Phi}_m + \phi\Delta_{0,2}\tilde{\Phi}_o - \delta_k^{LT}(m,s),$$

and use this to re-write the lifetime utility of having kids as young in the suburbs relative to living in the suburbs with no kids:

$$\delta_k^{LT}(y,s,z) = \Delta_{0,1}\tilde{\Phi}_{y,z} + \phi\Delta_{0,2}\tilde{\Phi}_{m,z} - \phi\rho_m\bar{\varepsilon}_{km,nk}(z) - (1-\phi\rho_m)\bar{\varepsilon}_{ky,km}(z).$$

To sum up, we have shown that we can use the observed fraction of households who have kids as young, kids as mature, and no kids, in order to identify the lifetime utility of having kids in the suburbs for each age and skill group. Table 11 includes the estimated amenities for each year and skill group. In the previous step, we had also identified the amenity of living downtown relative to the suburbs for each age, kids' age, and skill group. Notice that we can combine both to obtain the lifetime amenity of having kids in the city center relative to living in the suburbs with no kids.

Table 11: Lifetime amenity of kids relative to no-kids in the suburbs, δ_s^{LT}

1980			
Low-skill		High-skill	
Young	Mature	Young	Mature
0.3550	0.3053	0.1113	0.0507
2000			
Low-skill		High-skill	
Young	Mature	Young	Mature
0.4764	-0.0187	-0.0491	-1.6208

Estimation of β_k. We have estimated the model under the assumption of knowing the value for the scale parameters of the Gumbel distributions of both location idiosyncratic amenities, β_l and children idiosyncratic amenities, β_k. We can exploit the causal estimates obtained from the insurance mandates in the previous section to estimate these key parameters. This is work in progress.

8 Counterfactual exercise

In this section, we perform a counterfactual exercises to quantify how much of the observed gentrification can be generated by the observed changed in fertility (and fertility timing). Recall that the causal estimate could only capture the effect of the policy on gentrification. The goal of the model is to quantify the impact of the large changes in fertility in the recent decades and not only those that resulted from the policy. For instance, we would like to capture the effect of medical improvements, better information on the possible risks of delaying maternity, and ways to reduce them, but also changing norms and stigma associated to having kids as mature.

In the main counterfactual, we explore the effect of a change in the perceived probability of being able to have children when mature, conditionally on wanting to have children. The first step is to quantify this perceived probability. The medical literature has estimated that for women around 35 years old (the middle point of the mature age bin) the probability of getting pregnant naturally is 66 percent. Thus, we assume that in 2000 households have the correct information and the perceived probability is the actual one. In order to get an approximation for the perceived probability in 1980, we assume that households do not have accurate information and that they roughly estimate the probability based on the observed rate of pregnancy for mature households who did not have children as young. In the data, only 22 percent of households who did not have children as young, have children as mature in 1980. We take this number as the perceived probability.

Results from the counterfactual are preliminary and not included in this draft. Please, contact us if you're interested.

9 Conclusions

In the US, forming a family and having kids is associated with couples moving to the suburbs, where housing is larger and schools are better. However, more and more, young couples are choosing to postpone both fertility and the move to the suburbs. This has been made possible by medical advances in infertility treatments that allow couples to delay childbirth into the 30s without much risk. As couples stay downtown longer, precisely at a time when their incomes are growing fast, they become gentrifiers of their downtown neighborhoods by increasing the demand for amenities such as bars, movie theaters, and restaurants.

This paper provides causal evidence of the importance of delaying fertility on gentrification by exploiting state-level variation in the enactment of policies that essentially decreased the cost of delaying maternity. We find that these policies had a direct and statistically significant effect on the income growth of downtown vs. the suburbs which took place in parallel with a demographic change in the city center consistent with postponing the arrival of children and suburban life.

Recent work on gentrification has highlighted the responsiveness of downtown amenities to changes on the demographic composition of surrounding neighborhoods, which reinforces this process (Almagro and Dominguez-Iino (2019), Curci and Yusaf (2020)). We consider that this is an important avenue to explore in our context, as it could lead to additional incentives to postpone the arrival of children, amplifying the initial effect of the policy. The idea is that, as some couples decide to rely on the eventual utilization of ART treatments and extend the life period in which they live downtown, these areas become more attractive due to the endogenous response of amenities. This, in turn, induce more couples to postpone childbearing and moving to the suburbs, fuelling further waves of gentrification. Hence, we believe that a dynamic model of fertility and within-city location would be very useful to account for the general equilibrium effects of delaying maternity, improving our understanding of this matter. Furthermore, such a model would allow us to learn about the effects of delayed parenthood and subsequent gentrification on welfare and inequality, which are central from a policy perspective.

Another promising area for future research that is related to the findings in this paper concerns the spatial distribution of female labour force participation. On the one hand, postponing maternity is associated to increases in wages and to less costly career interruptions, which should raise female labour supply. On the other hand, gentrification affects the location choices of individuals and their commuting times to work, which are known to discourage substantially the labour force participation of women. A general equilibrium model that incorporated both channels would be useful to understand which effect prevails and, more generally, to evaluate the impact on gentrification on labour supply.

References

Abramowitz, Joelle, "Assisted Reproductive Technology and Women's Timing of Marriage and Childbearing," *Journal of Family and Economic Issues*, 2017, *38* (1), 100–117.

Almagro, Milena and Tomas Dominguez-Iino, "Location Sorting and Endogenous Amenities: Evidence from Amsterdam," 2019.

Bailey, Martha J, "More power to the pill: The impact of contraceptive freedom on women's life cycle labor supply," *The quarterly journal of economics*, 2006, *121* (1), 289–320.

Baum-Snow, Nathaniel and Daniel Hartley, "Accounting for central neighborhood change, 1980-2010," *Journal of Urban Economics*, 2019, p. 103228.

Bitler, Marianne and Lucie Schmidt, "Health disparities and infertility: impacts of state-level insurance mandates," *Fertility and sterility*, 2006, *85* (4), 858–865.

Bitler, Marianne P and Lucie Schmidt, "Utilization of infertility treatments: the effects of insurance mandates," *Demography*, 2012, *49* (1), 125–149.

Buckles, Kasey, "Stopping the Biological Clock: Infertility Treatments and the Career% Family Tradeoff, oBU Dissertation," 2005.

_ , "Understanding the returns to delayed childbearing for working women," *American Economic Review*, 2008, *98* (2), 403–07.

Caucutt, Elizabeth M, Nezih Guner, and John Knowles, "Why do women wait? Matching, wage inequality, and the incentives for fertility delay," *Review of Economic Dynamics*, 2002, *5* (4), 815–855.

Cortés, Patricia and Jessica Pan, "When Time Binds: Substitutes for Household Production, Returns to Working Long Hours, and the Skilled Gender Wage Gap," *Journal of Labor Economics*, 2019, *37* (2), 351–398.

Couture, Victor and Jessie Handbury, "Urban revival in America, 2000 to 2010," Technical Report, National Bureau of Economic Research 2017.

_ **, 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.

Curci, Federico and Hasin Yusaf, "Gentrifying cities, amenities and income segregation: evidence from San Francisco," Technical Report 2020.

Dorn, David, "Essays on inequality, spatial interaction, and the demand for skills." PhD dissertation, University of St. Gallen 2009.

Edlund, Lena, Cecilia Machado, and Maria Micaela Sviatschi, "Gentrification and the Rising Returns to Skill," Technical Report, National Bureau of Economic Research 2015.

Ellen, Ingrid Gould, Keren Mertens Horn, and Davin Reed, "Has falling crime invited gentrification?," *Journal of Housing Economics*, 2019, *46*, 101636.

Goldin, Claudia, "A Grand Gender Convergence: Its Last Chapter," *American Economic Review*, April 2014, *104* (4), 1091–1119.

— **and Lawrence F Katz**, "The power of the pill: Oral contraceptives and women's career and marriage decisions," *Journal of political Economy*, 2002, *110* (4), 730–770.

Hamilton, Barton H and Brian McManus, "The effects of insurance mandates on choices and outcomes in infertility treatment markets," *Health economics*, 2012, *21* (8), 994–1016.

Hoelzlein, Mattias, "Two-Sided Sorting and Spatial Inequality in Cities," 2019.

Hwang, Jackelyn and Jeffrey Lin, "What have we learned about the causes of recent gentrification?," *Cityscape*, 2016, *18* (3), 9–26.

Jain, Tarun, Bernard L Harlow, and Mark D Hornstein, "Insurance coverage and outcomes of in vitro fertilization," *New England Journal of Medicine*, 2002, *347* (9), 661–666.

Kroeger, Sarah and Giulia La Mattina, "Assisted reproductive technology and women's choice to pursue professional careers," *Journal of Population Economics*, 2017, *30* (3), 723–769.

Lee, Sanghoon and Jeffrey Lin, "Natural amenities, neighbourhood dynamics, and persistence in the spatial distribution of income," *The Review of Economic Studies*, 2018, *85* (1), 663–694.

Leridon, Henri, "Can assisted reproduction technology compensate for the natural decline in fertility with age? A model assessment," *Human reproduction*, 2004, *19* (7), 1548–1553.

Machado, Matilde P and Anna Sanz de Galdeano, "Coverage of infertility treatment and fertility outcomes," *SERIEs*, 2015, *6* (4), 407–439.

Menken, Jane, James Trussell, and Ulla Larsen, "Age and infertility," *Science*, 1986, *233* (4771), 1389–1394.

Miller, Amalia R, "The effects of motherhood timing on career path," *Journal of population economics*, 2011, *24* (3), 1071–1100.

Schmidt, Lucie, "Effects of infertility insurance mandates on fertility," *Journal of health economics*, 2007, *26* (3), 431–446.

Su, Yichen, "The rising value of time and the origin of urban gentrification," *Available at SSRN 3216013*, 2018.

van Noord-Zaadstra, Boukje M, Caspar W Looman, Hans Alsbach, J Dik Habbema, Egbert R te Velde, and Jan Karbaat, "Delaying childbearing: effect of age on fecundity and outcome of pregnancy.," *Bmj*, 1991, *302* (6789), 1361–1365.

Wilde, Elizabeth Ty, Lily Batchelder, and David T Ellwood, "The mommy track divides: The impact of childbearing on wages of women of differing skill levels," Technical Report, National Bureau of Economic Research 2010.

A Appendix

Impact of infertility insurance mandates on gentrification

The following evidence shows that restricting our attention to cities with at least one million inhabitants barely changes our results. Table 12 displays the results of running the exact same specification than that in Table 5 but including all cities with at least 100,000 inhabitants instead. The results are very similar.

Table 12: The effect of infertility insurance mandates on gentrification

	Prob. above median	Log median income	% College Graduate
	(1)	(2)	(3)
Center	-0.418***	-0.620***	-0.00413*
	(0.0102)	(0.00773)	(0.00217)
Treated × Post	0.0326*	0.00210	0.0233***
	(0.0198)	(0.0151)	(0.00424)
Center × Treated	-0.142***	-0.0322***	-0.000950
	(0.0164)	(0.0125)	(0.00351)
Center × Post	0.0449***	0.0357***	0.0116***
	(0.0131)	(0.00994)	(0.00279)
Treated × Center × Post	0.123***	0.0433***	0.0175***
	(0.0211)	(0.0160)	(0.00451)
Observations	104608	104608	104608
Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
City Size FE	Yes	Yes	Yes
State Trends	Yes	Yes	Yes
City Size Trends	Yes	Yes	Yes

Notes: This table displays the impact of infertility insurance mandates on several measures of gentrification: (1) the probability that a census tract's income is above median income in the city; (2) the census tract's (log) median income; and (3) the percentage of college graduates in a census tract. Controls include: city's population, city's (log) median income, the share of college graduates in the city, and the share of jobs within 3 miles distance from the census tract. This table reports only selected coefficients, the full specification can be found in equation 1. Standard errors in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

Table 13 shows that our results are robust to the exclusion of the control variables list in the main specification: (log) income of the city, the share of jobs within 3 miles distance from the neighborhood, (log) population of the city, and the share of college graduates in the city.

Table 13: The effect of infertility insurance mandates on gentrification

	Prob. above median	Log median income	% College Graduate
	(1)	(2)	(3)
Center	-0.423***	-0.609***	0.000222
	(0.0137)	(0.0105)	(0.00302)
Treated × Post	-0.00283	0.120***	-0.00457
	(0.0223)	(0.0172)	(0.00492)
Center × Treated	-0.140***	-0.0498***	-0.0195***
	(0.0223)	(0.0171)	(0.00490)
Center × Post	0.0624***	0.0329**	0.00800**
	(0.0167)	(0.0128)	(0.00367)
Treated × Center × Post	0.0479*	-0.0423**	-0.0000706
	(0.0278)	(0.0213)	(0.00612)
Observations	70289	70289	70289
Controls	No	No	No
Year FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
City Size FE	Yes	Yes	Yes
State Trends	Yes	Yes	Yes
City Size Trends	Yes	Yes	Yes

Notes: This table displays the impact of infertility insurance mandates on several measures of gentrification: (1) the probability that a census tract's income is above median income in the city; (2) the census tract's (log) median income; and (3) the percentage of college graduates in a census tract. This table reports only selected coefficients, the full specification can be found in equation 1. Standard errors in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

The location patterns of families

As we argued in the main text, one reason why couples may decide to move out of the city center whenever they have children is that the characteristics of the housing stock may not be ideal for children. For instance, houses downtown may be smaller and lack outdoor space. To provide some evidence of this channel, Figure 5 shows the distribution of houses by number of rooms in the suburbs and downtown. As can be inspected, houses are larger in the suburbs than downtown and thus more suitable for family life.

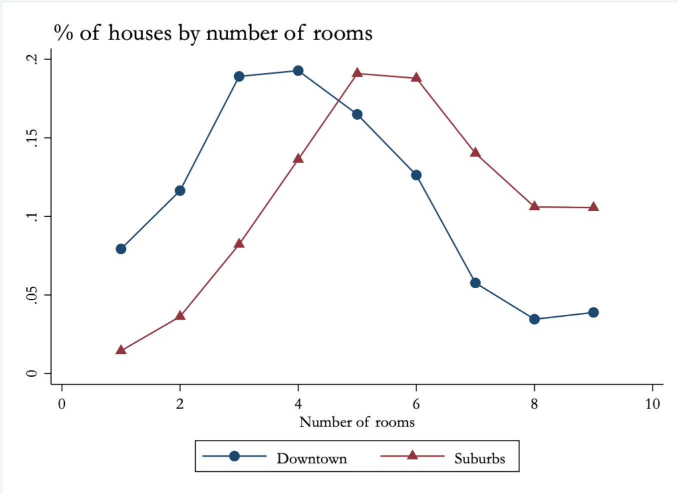
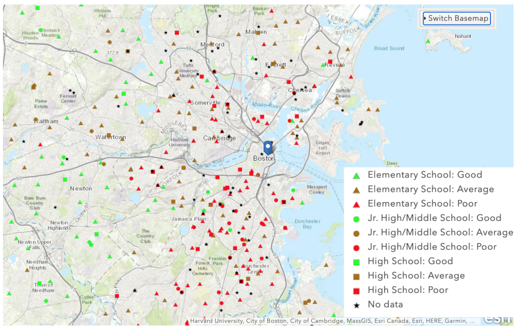


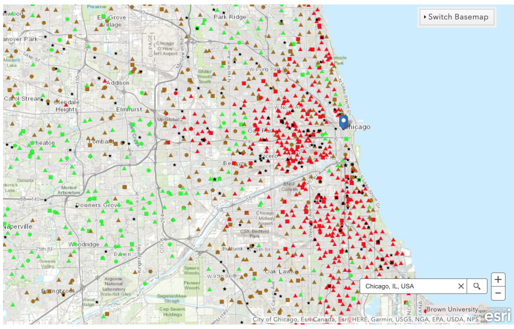
Figure 5: Housing Size

Notes: This figure displays the percentage of houses downtown/in the suburbs by the number of rooms in 2000.

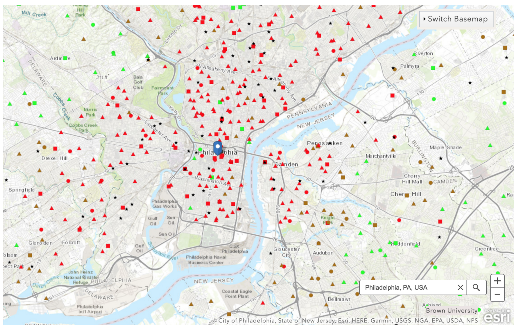
Another reason why families may prefer to relocate to the suburbs is their proximity to children specific amenities that families without children do not value. An important one is the quality of surrounding schools. To illustrate this point, we provide some examples on schools location by quality in Figure 6. We have accessed maps by "Map US Schools", which is part of the American Communities Project at Brown University, led by John Logan. These maps have been constructed using data from the National Center for Education Statistics and contain detailed information on the location and quality of most US schools. All cities accessed show a similar pattern, low school quality for all school education levels in the centre of the city and much high school quality for all education levels in the suburbs. We have selected some well-known cities that are located in treated and non treated states to simply illustrate the pattern.



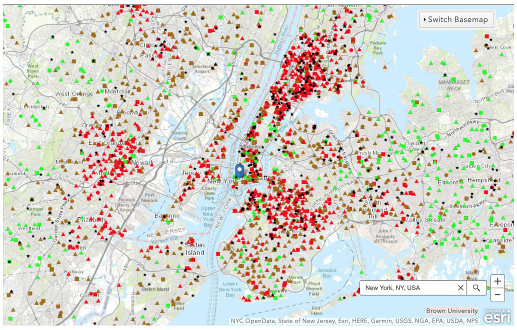
(a) Boston



(b) Chicago



(c) Philadelphia



(d) New York

Figure 6: Schools Quality within the City

Notes: This figures displays the location and quality of schools in different cities in 2000. Source: National Center for Education Statistics, accessed via https://s4.ad.brown.edu/Projects/usschools/index.html