CESifo WORKING PAPERS

9992
2022

October 2022

# Second-Hand Gentrification: Theory and Evidence from High-Speed Rail Extensions

*Gabriel Loumeau, Antonio Russo*

CESifo

**Impressum:**

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9992*

# Second-Hand Gentrification: Theory and Evidence from High-Speed Rail Extensions

## Abstract

This paper investigates whether and how gentrification spreads along intercity transport connections. We consider a model with heterogeneous individuals populating a primary and a secondary city, with commuting within and between cities. By reducing the cost of intercity commuting, the HSR connection induces migration by skilled individuals towards the secondary city. Therefore, house prices rise in the secondary city, and unskilled individuals are pushed to its periphery. We call this effect *second-hand gentrification*. We confirm these predictions using the 2017 expansion of the French HSR network from Paris to Bordeaux and Rennes. We find that the HSR connection made skilled Parisians more likely to move to Bordeaux and Rennes, that these individuals locate over-proportionally in central locations of such cities, and that housing prices there consequently increased by €400 per $m^2$ (i.e., 7%). Remarkably, we also find a negative effect on prices in Paris.

JEL-Codes: R230, R110, R410.

Keywords: gentrification, high speed rail, housing market, Intercity travel.

*Gabriel Loumeau**
*Department of Management, Technology, and Economics, ETH Zürich*
*Leonhardstrasse 21*
*Switzerland – 8092 Zurich*
*loumeau@ethz.ch*

*Antonio Russo*
*Department of Economics*
*University of Sheffield / United Kingdom*
*A.Russo@sheffield.ac.uk*

*corresponding author

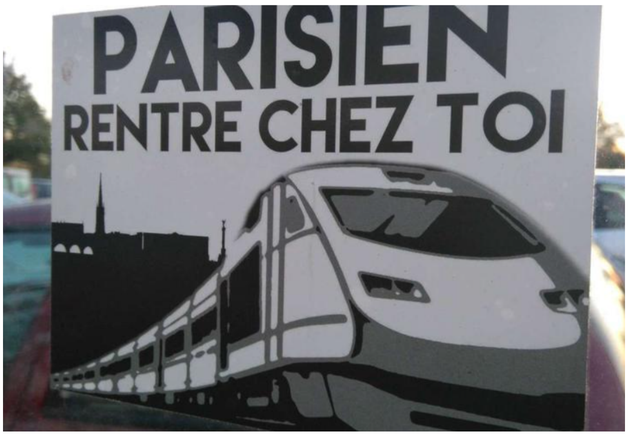
September 29, 2022
We are grateful to participants to presentations at IEB Barcelona, the UEA meeting in London, DTU Copenhagen, VU Amsterdam, Loughborough University, the ITEA conference in Toulouse and the EEA meeting in Milan. In particular, we thank (in alphabetical order) Léa Bou Sleiman, Jonathan Hall, Guillaume Monchambert and Jos van Ommeren for helpful comments and suggestions.

# 1 Introduction

In the past decade, concerns about the spread of gentrification from large urban centers to nearby small- and mid-sized cities have made headlines. Far from being specific to a given country, this phenomenon, labelled *second-hand gentrification*, has received widespread international coverage from the press, politicians and NGOs, as illustrated in Appendix A. Examples include cities such as Hamilton, next to Toronto; Birmingham and Brighton, close to London; Leipzig, close to Berlin; and Bordeaux and Rennes, close to Paris.

Like all forms of gentrification, second-hand gentrification generates significant social unrest, particularly among the original inhabitants of smaller cities. Evidence of this unrest is provided in Figure 1, depicting signs recently posted by locals in Bordeaux. However, given its nature (e.g., the fact that it involves migration across different urban areas), second-hand gentrification has specific features and implications that distinguish it from other instances of gentrification studied so far. Indeed, Figure 1 suggests a link between transport connections to primary cities, such as high-speed railways, and the spread of gentrification. Yet, we know little about how gentrification propagates from large to smaller cities, and about its implications for housing markets and welfare. The aim of this paper is to shed some light on these issues, focusing on the role of transport connections between primary and secondary cities.

Figure 1: Second-hand gentrification: evidence of social unrest in Bordeaux




*Note: the text, in French, reads "Parisian go home."*

As a prime example of intercity transport connections, High Speed Railway (HSR) lines provide the main application and a key source of data for our study. Worldwide, the HSR network is growing rapidly. Since 2000, its size has more than doubled, following important infrastructure investments in China, France, the UK, the United-States, Italy and Spain, among others.[1] By connecting separate urban areas, these transport connections potentially foster economic development, e.g., by integrating distant labor markets. However, the benefits and costs are presumably not spread evenly across all individuals. By reducing travel costs, these connections make the primary city more accessible to people living in secondary ones, but high-skilled workers typically benefit more than low-skilled ones, e.g., because they can take advantage of more substantial wage premia in primary cities (Combes et al., 2008; Baum-Snow and Pavan, 2011; Behrens et al., 2014). Thus, intercity connections can also determine new migration patterns and, in turn, increase the demand for housing in secondary cities by high-skilled individuals. By the same token, however, these connections may also soften the high-skilled demand for housing in primary cities. The resulting changes in housing prices can redistribute welfare across groups (skilled/rich vs. unskilled/poor), within and between cities.

Our study combines theoretical and empirical analysis. We propose a model with a primary and a secondary monocentric cities, with intra- and inter-city commuting. Individuals differ in their skill (wage) level and preference for living in the primary city. In our model, skilled individuals in the secondary city commute to the primary one (possibly only from time to time), either because their job requires them to (e.g., they must attend important meetings at company headquarters) or to take advantage of higher earning opportunities in the primary city (e.g., meet with important clients). We assume intercity commuting only takes place by train and train stations are located within the CBD of each city. In equilibrium, due to their higher value of time, skilled individuals live closer to the CBD of each city than unskilled ones.

Given the above setting, we show that a reduction in the (time) cost of long-distance travel, such as the opening of a HSR line, induces migration of skilled individuals towards the secondary city. Hence, land prices increase there, forcing unskilled individuals to either move to the periphery of the secondary city or migrate to the primary one. Therefore, the model shows that reducing the cost of intercity travel between primary and secondary cities triggers effects that are consistent with second-hand gentrification. The model also suggests a possible negative effect on housing prices in the primary city, particularly in the areas where high-skilled workers that move to the secondary city used to live. Finally, the model predicts that the welfare of skilled

[1] Studies documenting the worldwide growth of HSR since 2000 include, among others, Lawrence et al. (2019) and Egger et al. (2020).

individuals in both cities increases when long-distance travel costs decline. However, unskilled ones benefit only if they live in the primary city, due to the effects on the housing market described above.

We confirm these predictions empirically using a Difference-in-Difference (DiD) identification strategy. The analysis exploits the July 2017 opening of HSR lines connecting Paris to Bordeaux and Paris to Rennes. These connections implied, respectively, a 35% reduction in travel time (from 3h 12min to 2h 04min) and a 32% reduction in travel time (from 2h 04min to 1h 25min). Using fine-grained data on the universe of housing transactions, intra-city density and inter-city migration, we show that the HSR opening had important effects that differed significantly across groups. First, housing prices increase by about €400 in Rennes and Bordeaux (approx. 7% increase). In Paris, housing prices experience a slower increase – by about €245 (approx. 3% of pre-HSR prices) – in the *arrondissements* close to Montparnasse HSR station relative to the rest of Paris. Second, the flow of skilled workers between Paris and Bordeaux/Rennes increase significantly in 2017 and 2018. Finally, the share of skilled workers among in-movers to central Bordeaux/Rennes is increasing by about 10 percentage points.

**Contribution to the literature.** The paper makes several contributions to the literature. From a theoretical standpoint, we build on the literature studying intracity and intercity commuting in systems of cities. Borck and Wrede (2009) proposed a model with two monocentric cities, evaluating the effects of subsidies to intracity and intercity commuting (which they model as a reduction in the time cost of long-distance travel), in presence of agglomeration economies. We adopt a similar setting, but focus on the effects of intercity connections when individuals differ in their wage level. In turn, we analyze the link between transport connections and the internal structure of cities, focusing on the spatial sorting of skilled vs. unskilled individuals. The analysis also contributes to the literature on the distributional effects of transport policies and their impact on the housing market (Borck and Wrede, 2005; Brueckner and Selod, 2006; Borck and Wrede, 2008; De Borger and Russo, 2018). This literature has shown that the "poor" may support subsidies to commuting modes primarily used by the "rich" (e.g., cars), because these subsidies reduce housing prices, particularly in proximity to the CBD (Borck and Wrede, 2005, 2008). Our contribution is to study policies that change *intercity* commuting costs. A reduction in such costs can result in lower housing prices in one of the connected cities, but higher prices in the other. As a result, unskilled individuals can benefit even though they do not use the intercity connection, but only if they reside in the primary city.

This paper also contributes to the literature on neighborhood gentrification (Brueckner and Rosenthal, 2009; Zheng and Kahn, 2013; Guerrieri et al., 2013; Ding et al., 2016), by studying how gentrification spreads along intercity transport networks. The combination of the travel time shock induced by the HSR extension in 2017 and bilateral mobility data allows us to present novel evidence on the causal mechanism at play behind "second-hand gentrification". Namely, we observe an increase in migration by skilled workers towards the secondary city, which in turn increases housing prices and pushes former secondary city residents out of central locations.

We contribute to the literature studying the intracity and intercity effects of transport infrastructure (Baum-Snow, 2007; Michaels, 2008; Duranton and Turner, 2012; Zheng and Kahn, 2013; Ahlfeldt and Feddersen, 2017a; Gibbons et al., 2019; Banerjee et al., 2020; Hayakawa et al., 2021), with a particular focus on railways (Baum-Snow et al., 2017; Charnoz et al., 2018; Donaldson, 2018; Egger et al., 2020; Koster et al., 2021). We exploit individual-level data at a very fine spatial resolution on the universe of intra- and inter-city commuting and migration flows between 2013 and 2019, as well as the unit-level data on the universe of geo-localized daily housing transactions in France between 2016 and 2019.[2] Equipped with such data, we can study both the between and within city effects of HSR extensions at a higher level of precision than previously done in the literature. The fine spatial resolution of the data allows us to observe conunterintuitive effects such as the reduction in housing prices in specific areas of the primary city (i.e., Paris), despite its very large size. Furthermore, the adopted identification strategy permits estimating both the overall causal effects of the HSR extensions, as well as the respective transition paths on commuting, migration and housing prices.

The reminder of the paper is organized as follows. Section 2 describes our theoretical framework and the testable predictions it delivers. Section 3 introduces the empirical context and the data used. Section 4 provides the empirical results. Finally, Section 5 concludes.

[2]Despite the availability of data beyond 2019, we do not extend our analysis beyond this point in time to avoid major confounding factors such as the COVID-19 pandemic. As a reminder, the pandemic arrived in France in the first trimester of 2020, with a first lockdown starting in March 2020.

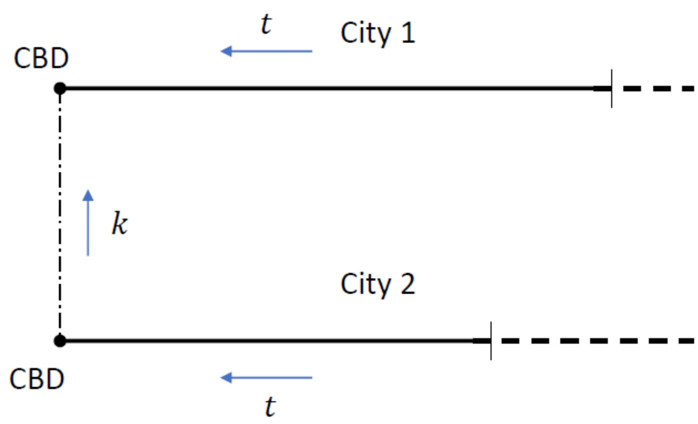
# 2 Theoretical framework

## 2.1 The model

We consider two monocentric and linear cities, indexed by $i \in \{1, 2\}$. We refer to city 1 as the "primary" city (e.g., Paris) and to city 2 as the "secondary" one (e.g., Bordeaux). These cities are inhabited by two groups of individuals, "skilled" and "unskilled", indexed by $j \in \{S, U\}$. We denote by $N_S$ and $N_U$ the given size of each group. The total population is $N = N_S + N_U$. Individuals choose whether to live in city 1 or 2, and also where to live within a given city. Furthermore, individuals can work in city 1 and 2.

All jobs in each city are located in the central business district (CBD), assumed to be point-sized. We model daily commuting costs following Borck and Wrede (2009). Let $\tilde{i} \in \{1, 2\}$ denote the city where an individual goes to work. Consider an individual who goes to work in city $\tilde{i}$ and lives in city $i$ at distance $x$ from its CBD. If $\tilde{i} = i$, the individual sustains only a short-distance commuting (time) cost $tx$, where $t > 0$. If the individual lives in a different city ($\tilde{i} \neq i$), her commuting cost is $tx + k$, where $k$ is the time cost of long-distance travel. We assume that long-distance travel takes place by train, and train stations in each city are located in the CBD. Hence, the individual must first travel $x$ miles to get to the train station. Figure 2 illustrates the spatial structure of the model.

Figure 2: Model structure




To capture the presence of a skill premium in large cities, we assume the skilled earn a weakly lower daily wage in the secondary city ($w_{2S}$) than in the primary one ($w_{1S}$). The wage of unskilled individuals, $w_U$, is instead identical in the two cities and smaller than the skilled wage, so that $w_{1S} \geq w_{2S} \geq w_U$.

All individuals have an exogenous time endowment (number of workdays) that we normalize to one. Although the model allows, in principle, for a flexible structure of commuting patterns, to streamline the exposition we assume that skilled individuals living in city 1 and all unskilled individuals work only in the city where they live ($\tilde{i} = i$). To rationalize this assumption, recall that there is no wage difference across cities for the unskilled. Similarly, the skilled who live in the primary city have little interest in traveling to work to the secondary city, since wages are lower there. Skilled individuals living in city 2, however, work some (possibly all) days in the primary city. Specifically, we assume they work for a given share $\alpha \in (0, 1]$ of days in city 1 and a share $1 - \alpha$ in city 2. For example, an individual's office could be in the secondary city, but she must travel some days to the primary one to participate in meetings at company headquarters or work with clients. In the extreme, if $\alpha = 1$, skilled individuals in city 2 commute to the primary city every day.

We normalize the length of a working day to one and we ignore leisure. Thus, an unskilled individual living in city $i$ at distance $x$ from the CBD earns the following daily income net of commuting costs:

$$I_{iU}(x) = w_U(1 - tx), i = 1, 2. \tag{1}$$

Furthermore, a skilled individual who lives in city 1 earns

$$I_{1S}(x) = w_{1S}(1 - tx). \tag{2}$$

Finally, a skilled individual living in city 2 earns

$$I_{2S}(x) = \alpha w_{1S}(1 - tx - k) + (1 - \alpha) w_{2S}(1 - tx) \tag{3}$$

Each individual spends her income, $I_{ij}(x)$, on a composite consumption good (the numeraire) and housing. Let $c_{ij}(x)$ be the level of consumption by an individual of type $j$ in city $i$ at distance $x$. We assume each individual occupies a lot of unit size and that land is the only input in housing production. Letting $p_i(x)$ be the rental price of a unit of land in city $i$ at distance $x$ from the CBD, we have

$$c_{ij}(x) = I_{ij}(x) - p_i(x). \tag{4}$$

We assume absentee landownership and that agricultural land rent at the boundary of each city equals zero.

Individuals derive utility from the consumption good and have an idiosyncratic preference for living in the primary city. Specifically, we assume that each individual in group $j$ gets a marginal utility $z_j$ from living in city 1, with $z_j \sim U[0, w_j]$. The parameter captures differences among the two cities besides jobs and wages, such as amenities (shops, restaurants, theatres, parks, etc.) and/or public services (schools, libraries, etc.), that individuals may attach different values to. For simplicity, $z_j$ does not depend on the individual's location within the city. The utility of an individual of type $j$, in city $i$ and at distance $x$ from the CBD is therefore

$$V_{ij}(x,z) = c_{ij}(x) + \mathbb{1}_{ij}(1,S)\, z_S + \mathbb{1}_{ij}(1,U)\, z_U, \tag{5}$$

where $\mathbb{1}_{i,j}(1,S) = 1$ (respectively, $\mathbb{1}_{i,j}(1,U) = 1$) if and only if a skilled (resp., unskilled) individual lives in city 1, and zero otherwise. Combining this expression with (4), we can write individual utility given $i$, $j$, $x$, $z_S$, $z_U$ as

$$V_{ij}(x,z) = I_{ij}(x) - p_i(x) + \mathbb{1}_{ij}(1,S)\, z_S + \mathbb{1}_{ij}(1,U)\, z_U. \tag{6}$$

## 2.2 Discussion of the setup

The model assumes that commuting patterns are exogenous to streamline the exposition, but this assumption is not crucial. For instance, we could let the share of days worked in the primary city by the skilled living in the secondary, $\alpha$, be endogenous. This would complicate the exposition, because the choice of $\alpha$ would depend on factors including as the distance from the CBD, the wage, the cost of long- and short-distance travel, and would be intertwined with the choice of location within a city. However, this complication would not change the results we present below regarding the effects of changes in the cost of long-distance travel. What really matters for those results is that the skilled in the secondary city commute long-distance at least sometimes ($\alpha > 0$).

We have also assumed that the wage that skilled individuals earn when working in the primary city does not depend on where they live. Alternatively, we could let these individuals earn a different wage, to capture, e.g., lower productivity from not fully developing connections at the company headquarters, compared to individuals who live (and work regularly) there. We could also let skilled individuals in city 2 earn the same daily wage, $w_{2,S}$, regardless of whether they work in the primary city. Again, allowing for these possibilities would not change the analysis in a fundamental way, as long as skilled individuals earn a higher wage than the unskilled.

Another simplifying assumption we have made is to ignore agglomeration economies, which are likely to be relevant for skilled jobs in the largest city. To capture such economies, we could let the wage of skilled individuals depend on how many such individuals work in each city. This assumption would not change our main findings regarding the effect of changes in long-distance travel costs, as we argue in Section 2.5.

We concentrate on a closed system of cities (exogenous total population) to focus on the redistribution of population and welfare within the two cities connected by the HSR line. The opening of an HSR connection may, however, also result in changes in the overall population of the two cities. We present an extension allowing for this possibility in Section 2.5.

The idiosyncratic utilities from living in the primary city, $z_j$, play an important role in our model for two reasons. First, they imply that we do not need equal utility of individuals (conditional on skill) among the two cities as a condition for equilibrium, which would impose a rigid structure on the allocation of population in the model.[3] Second, the idiosyncratic utilities ensure the effect of changes in the cost of commuting long-distance on utility depends on where individuals live.

The assumption that $z_j$ does not depend on the individual location within the primary city is not strictly necessary, but simplifies the exposition. Alternatively, we could have assumed all amenities and services are in located the CBD. Since in our model individuals commute to the CBD anyway, this assumption would just require an additional cost of commuting to $t$, with little change to the analysis.[4]

## 2.3 Solving the model

We now characterize the equilibrium of our model. To ease exposition, we are going to present the analysis focusing on the case where skilled individuals earn the same wage in both cities, denoted by $w_S$. That is, we set $w_{1,S} = w_{2,S} = w_S$. The generalized analysis, where $w_{1,S} \geq w_{2,S}$, yields qualitatively similar results and is available in Appendix B.

[3]The main issue is that the model would be overidentified, given that population sizes, $n_1$ and $n_2$, would have to satisfy three independent equations at the same time. Namely, the conditions requiring that skilled and unskilled get equal utilities in the two cities and that $n_1 + n_2 = N$.

[4]We assume the parameters $z_j$ are distributed uniformly for ease of exposition. It is not obvious that any other distribution would describe the preferences for living in the primary city in a more realistic way.

#### 2.3.1 Allocation of population between cities

We begin by characterizing the allocation of population between the two cities. Let $n_i$ denote the number of individuals that live in city $i$, and $n_{ij}$ be the number of individuals of type $j$ that live there, with $n_i = \sum_{j=S,U} n_{ij}$. As we show in Appendix B, this allocation can be obtained by using the equilibrium conditions requiring that (i) individuals of each group be indifferent as to their location within a city, and (ii) that individuals live in the primary city if and only if their idiosyncratic preference for such city is above a threshold, $\bar{z}_j$. Figure 3 provides an illustration of how the equilibrium populations are characterized (as customary, in the figure we focus on the utility of individuals living at the boundary of each city). We obtain that[5]

$$n_1 = \frac{tN^2 + N + \alpha k N_S}{1 + 2tN}, \quad n_2 = \frac{tN^2 - \alpha k N_S}{1 + 2tN}, \tag{7}$$

$$n_{1S} = N_S \frac{1 + tN + \alpha k(1 + 2tN_U)}{1 + 2tN}, \quad n_{2S} = N_S \frac{tN - \alpha k(1 + 2tN_U)}{1 + 2tN}, \tag{8}$$

$$n_{1U} = N_U \frac{1 + t(N - \alpha k 2N_S)}{1 + 2tN}, \quad n_{2U} = N_U \frac{t(N + \alpha k 2N_S)}{1 + 2tN}. \tag{9}$$

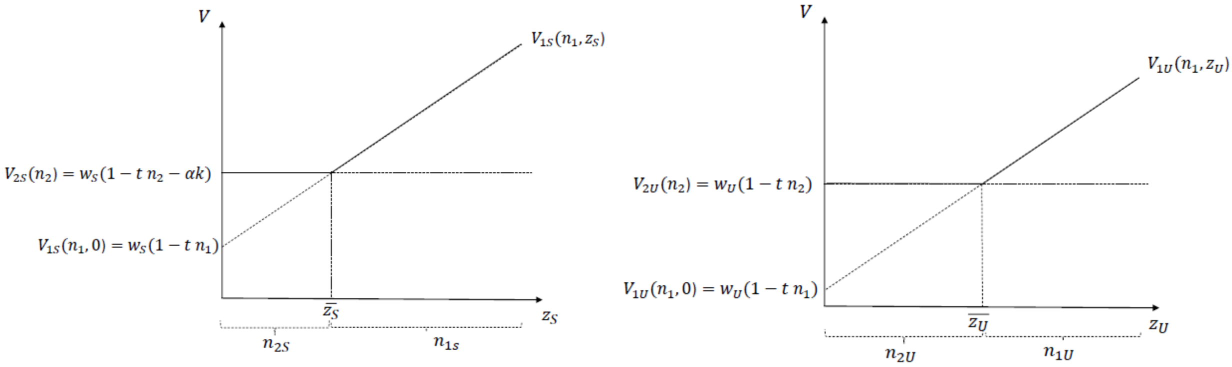
A first observation from the above expressions is that $n_1 > n_2$ in equilibrium. This outcome is intuitive, since the primary city offers higher wages and lower commuting cost to the skilled individuals.[6] A second observation from expressions (7)-(9) is that the cost of long-distance travel, $k$, affects the choice of city in a way that differs among skilled and unskilled individuals. Consider a reduction in $k$ (as in, e.g., the opening of a HSR line). This change has a direct effect on commuting expenses of the skilled individuals, but not on the unskilled ones (who only travel short-distance). This effect encourages the skilled individuals to live in the secondary city. However, by making city 2 relatively more attractive, lower long-distance travel costs also make housing relatively less affordable there, particularly to the unskilled. Therefore, not only the size, but also the composition of the population in the two cities change. In terms of

[5]The condition $N > max(\alpha k(\frac{1}{t} + 2N_U), \alpha k 2N_S - \frac{1}{t})$ is necessary and sufficient for all population quantities to be positive. We assume throughout the analysis that this condition holds. We also assume that $tN < 1$, which is sufficient for all individuals to achieve a positive level of utility in equilibrium. Both conditions require that commuting costs be not exceedingly large.

[6]As Figure 3 shows, for a given skill level, the disposable income (net of commuting and housing costs) when living in the secondary city is higher than in the primary city. Individuals live in the primary city only if their idiosyncratic utility, , $z_j$, compensates this gap. This outcome is a consequence of the assumption (made for notational convenience) that such utilities are positive, but does not drive our results. We show in Appendix E that allowing for negative values of $z_j$ would not change the analysis in a significant way.

Figure 3: Allocation of skilled and unskilled individuals among cities

(a) Skilled (b) Unskilled




overall population size, though, the direct effect dominates, so city 2 (resp. 1) gets larger (smaller) when $k$ goes down. We summarize these findings in the following Proposition:

**Proposition 1** *The size of the primary city and the share of skilled population therein increase with the cost of long-distance travel, $k$. By contrast, the size of the secondary city and the share of skilled individuals therein decrease with $k$.*

### 2.3.2 Allocation of population within cities, city structure and land rents

To determine the equilibrium allocation of individuals within each city, we first characterize the bid-rent functions, starting from city 1. Given the expressions for individual utility in (6), and that land rent at the boundary is zero, the bid rent functions for each group of individuals, $p_{1j}(x)$, are

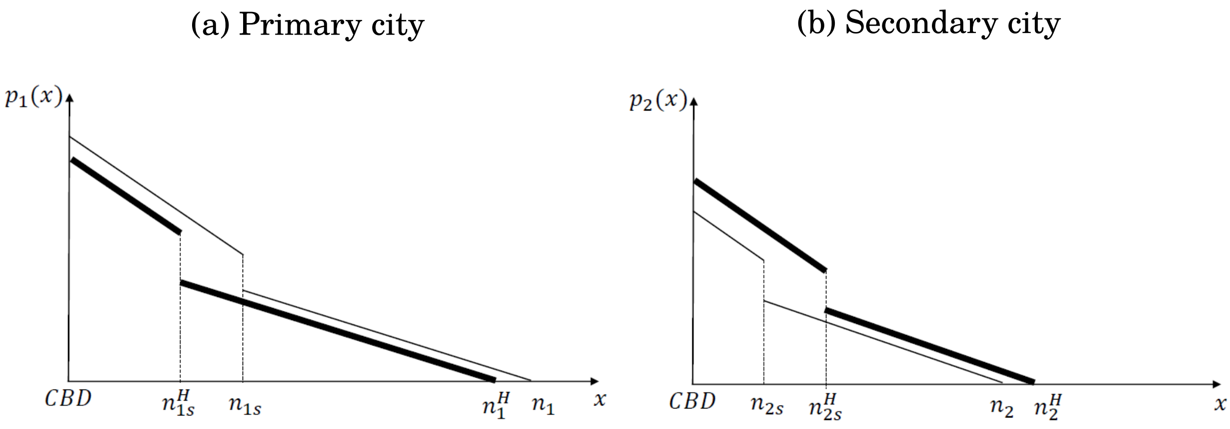
$$p_{1S}(x) = w_S(1 - tx) + z_S - w_S(1 - tn_1) - z_S = w_S t(n_1 - x), \quad p_{1U}(x) = w_U t(n_1 - x). \tag{10}$$

Similarly, we obtain the bid-rent functions in city 2:

$$p_{2S}(x) = w_S t(n_2 - x), \quad p_{2U}(x) = w_U t(n_2 - x). \tag{11}$$

Note that, since neither long-distance travel costs nor the idiosyncratic preference for city 1 depend on location within cities, these parameters do not enter the bid-rent functions. Moreover, given $w_S > w_U$, the bid rent functions of skilled individuals are everywhere steeper than those of the unskilled. Hence, there is full sorting of individuals by skill: skilled individuals outbid the unskilled ones for residential plots closest to the CBD. Furthermore, there is a discrete change in the price of land at the

Figure 4: Effect of reducing the long-distance travel cost on the rental price of land




*Notes:* Bold lines represent the rental price of land given lower long-distance travel costs. Variables conditional on lower long-distance travel costs denoted with superscript "-".

border between the areas occupied by skilled and unskilled individuals. Formally, the equilibrium rental price of land in each city is as follows

$$p_i(x) = \begin{cases} w_S t(n_i - x) & \text{if } 0 < x \leq n_{iS}, \\ w_U t(n_i - x) & \text{if } n_{iS} < x \leq n_i, \\ 0 & \text{if } x > n_i, \end{cases} \quad i = 1, 2. \tag{12}$$

Consider now the effect of reducing the cost of long-distance travel, $k$. As shown in Figure 4, the rental price in city 2 increases everywhere, because the total population of that city expands (Proposition 1). Furthermore, the number of skilled individuals increases as well, and so does the area occupied by this group. In addition, land rents in the secondary increase, particularly in the area newly occupied by skilled individuals and formerly inhabited by unskilled ones. The opposite effects apply in city 1. As we discuss further below, these findings suggest a relation between the opening of HSR connections (which reduce $k$) and gentrification in the secondary city. Quite interestingly, the findings also suggest that, by the same token, the pressure on the housing markets in the primary city - particularly in proximity to the HSR station - should be somewhat relieved.

**Proposition 2** *The rental price of land and the area occupied by skilled individuals in the primary (resp. secondary) city increase (resp. decrease) with the cost of long-distance travel.*

## 2.4 Individual utility and long-distance travel costs

Starting from the individual utility in (6), focusing again on individuals at the city boundary and given the equilibrium population sizes in (7), we can write the equilibrium utilities of skilled individuals as follows

$$V_S(z_S) = \begin{cases} V_{1,S}(z_S) = w_S(1 - t\frac{N+tN^2+\alpha kN_S}{1+2tN}) + z_S & \text{if } z_S \geq \bar{z_S}, \\ V_{2,S}(z_S) = w_S(1 - t\frac{tN^2-\alpha kN_S}{1+2tN} - \alpha k) & \text{if } z_S < \bar{z_S}, \end{cases} \tag{13}$$

where

$$\bar{z_S} = w_S\left(t\frac{N + 2\alpha kN_S}{1 + 2tN} - \alpha k\right). \tag{14}$$

From these expressions, we can determine how the utility of skilled individuals changes with the cost of commuting long-distance:

$$\frac{\partial V_S}{\partial k} = \begin{cases} -w_S\alpha\frac{tN_S}{1+2tN} & \text{if } z_S \geq \bar{z_S}, \\ -w_S\alpha\frac{1+t(N+N_U)}{1+2tN} & \text{if } z_S < \bar{z_S}, \end{cases} \tag{15}$$

and where $\frac{\partial \bar{z_S}}{\partial k} = -w_S\frac{1+2tN_U}{1+2tN}$. The utility of skilled individuals decreases with the cost of long-distance travel. Although land rents decrease in city 2, there is a direct loss to the skilled in the form of higher travel costs, and the net effect is negative. The utility of skilled individuals living in the primary city decreases as well, because a higher $k$ induces net migration towards that primary city, which raises land rents there.

Consider now the utility of unskilled individuals. We obtain that

$$V_U(z_U) = \begin{cases} V_{1,U}(z_U) = w_U(1 - t\frac{N+tN^2+\alpha kN_S}{1+2tN}) + z_U & \text{if } z_U \geq \bar{z_U}, \\ V_{2,S}(z_U) = w_U(1 - t\frac{tN^2-\alpha kN_S}{1+2tN}) & \text{if } z_U < \bar{z_U}, \end{cases} \tag{16}$$

where

$$\bar{z_U} = w_U\left(t\frac{N + 2\alpha kN_S}{1 + 2tN}\right). \tag{17}$$

Starting from these expressions, we can determine how the utility of unskilled individuals varies with the cost of commuting long-distance:

$$\frac{\partial V_U}{\partial k} = \begin{cases} -w_U\alpha\frac{tN_S}{1+2tN} & \text{if } z_U \geq \bar{z_U}, \\ w_U\alpha\frac{tN_S}{1+2tN} & \text{if } z_U < \bar{z_U}, \end{cases} \tag{18}$$

and $\frac{\partial \bar{z_U}}{\partial k} = -w_U\frac{2tN_S}{1+2tN}$. Unskilled individuals who live in the primary city are worse-off when $k$ increases, since land in that city becomes more expensive. By the same token, however, unskilled individuals in the secondary city benefit from the reduction in equilibrium land prices there. We summarize these results in the following Proposition:

**Proposition 3** *The equilibrium utility of skilled individuals decreases with the cost of long-distance travel, regardless of where they live. The utility of unskilled individuals decreases with the cost of long-distance travel if and only if they live in the primary city, and increases otherwise.*

## 2.5 Extensions

Before summarizing the implications of the above analysis, we briefly present two extensions to the model.

**Agglomeration economies.** In Appendix C, we propose a version of the model that incorporates agglomeration economies. Specifically, we let the wage earned by skilled individuals in the primary city be an increasing function of the number of such workers that live there and, in our model, work there regularly. To capture the key implications of agglomeration economies in the simplest possible way, we assume that the unskilled all have the same preference for living in the primary city. Furthermore, we retain the assumption that the skilled wage in the secondary city is exogenous. We show that a reduction in $k$ results again in an increase in the overall population in the secondary city, migration of skilled individuals from the primary one and increased (resp. decreased) land rents in the secondary (resp. primary) city. The intuition is that changes in $k$ still make the primary city more accessible to individuals living in the secondary. Fundamentally, agglomeration does not change the fact that skilled individuals stand to benefit from such accessibility more than unskilled ones.[7]

**Open city system.** In Appendix D, we extend the model allowing the size of skilled and unskilled groups, $N_S$ and $N_U$, to be endogenous. We assume that, in a preliminary stage, individuals choose whether to settle in the system formed by cities 1 and 2, or elsewhere. This decision is based on the utility they expect to receive in the system (see (13) and (16)). The analysis then unfolds as in the baseline model. We find that changes in $k$ have generally an ambiguous effect on $N_S$ and $N_U$. This finding is fairly intuitive when considering unskilled individuals, since, as shown in (18), the direct effect of $k$ on their utility can be negative or positive, depending on which city they live

[7]As we argue in Appendix C, it is possible to make the model more complex by letting the skilled wage in the primary increase in the total number of workers there (including those who commute long distance irregularly) and to allow for agglomeration economies in skilled wages also in the secondary city. Neither of these modifications would change the results substantially.

in. The skilled instead benefit when $k$ goes down (see (15)), which, by itself, should attract more such individuals in the system. However, if $N_U$ increases, there is a countervailing effect on the utility of the skilled, because of the ensuing increase in land rents. As a result, the predictions from Propositions 1-3 become less clear-cut in an open system of cities.

### 2.6 Implications of the analysis and testable hypotheses

Propositions 1-3 predict several interesting effects of connecting primary and secondary cities with HSR lines. The time gains from these connections amount to a reduction in the cost of long-distance travel, $k$. The model predicts that the HSR connection should induce a reduction in the size of the primary city, and an increase in the size of the secondary one. Furthermore, this connection should induce migration of skilled workers from the primary to the secondary city, and migration of unskilled workers in the opposite direction, so that the share of skilled individuals in the secondary city increases (Proposition 1). The flow of long-distance commuters to the primary should also increase, with most of the increase being among skilled workers.

In our model, skilled individuals tend to live in the central areas of both cities, unlike unskilled ones. The HSR connection should induce an expansion of the residential area occupied by skilled individuals in city 2, with unskilled ones moving further out. In addition, the connection should result in an increase in land prices in city 2, with the strongest increase in the areas where skilled individuals replace unskilled ones (Proposition 2). The model suggests that the opposite effects should arise in city 1.

These predictions are consistent with the hypothesis that HSR connections foster gentrification in secondary cities, by inducing migration by skilled individuals from the primary city. The key force driving this effect is that the HsR connection makes the primary city more accessible when living in the secondary city. In our model, this accessibility is more valuable to the skilled than to the unskilled, given that the skilled benefit from greater opportunities to work and earn higher wages in the primary city.

A less intuitive prediction of the model is that, by inducing net migration of skilled individuals out of the primary, the HSR connection also tends to alleviate the pressure on the housing market therein - particularly in the areas previously inhabited by the skilled that end up moving to the secondary city. From an empirical perspective, it is perhaps unlikely to find such an effect in the housing market of the entire primary city, given the typical size of such cities and the presence of many possible confounding factors. However, as we argue when presenting the empirical analysis (Section 4), it is reasonable to expect to find effects consistent with the model predictions at least in the area in proximity to the HSR station.

Finally, the model predicts a redistribution of welfare among the different groups as a result of the HSR connection. Specifically, this connection should be beneficial to skilled individuals in both cities. The skilled in city 2 use the HSR for commuting to the primary city. Although they pay more for housing, the net effect on their utility is positive. Skilled and unskilled individuals in the primary city, despite not commuting long-distance, should also benefit via the reduced pressure on the housing market. On the other hand, unskilled individuals in the secondary city suffer, because they do not use the rail connection, but see their housing expenditures increase due to the rise in land rents (Proposition 3).

In sum, the theoretical analysis yields the following testable implications:

- H1: the HSR line should have a positive effect on floor prices in the secondary city.
- H2: the HSR line should have a negative effect on floor prices in the primary city.
- H3: the HSR line should induce migration of skilled workers to the secondary city from the primary city.
    - Corr. H3: long-distance travel should increase particularly among skilled workers.
- H4: the HSR line should increase the share of skilled individuals moving into the central neighborhoods of the secondary city and discourage "native" individuals from moving there.

# 3 Context and data

To test the above hypotheses, we exploit the extension of the French HSR in July 2017 and study its effects on Paris (primary city) and Bordeaux and Rennes (secondary cities). This section describes the development of the high-speed rail network in France since its opening in 1981, briefly presents how different French metropolitan areas are impacted by the HSR extension, before finally documenting the data used in the empirical analysis. In describing the context, we pay particular attention to the institutional details used in the identification strategy presented in Section 4.

## 3.1 Context

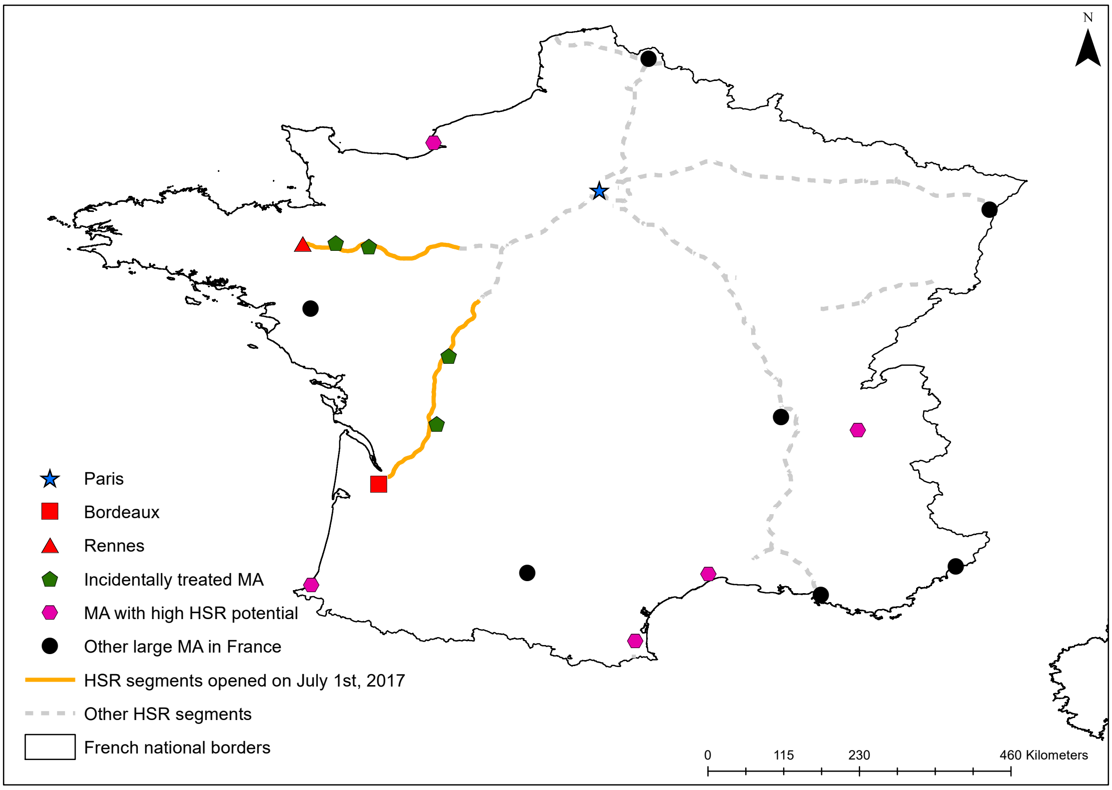
**France's high-speed rail network.** The French high-speed rail network is operated by the French National Railway Company (SNCF). The first high-speed connection opened in September 1981 between Paris and Lyon. Since then, many of France's largest cities have been connected via high-speed rail to Paris. In Figure 5, dash grey lines trace the HSR network operational before July 2017. These include Lyon, Marseille, Lille, and Strasbourg. Running with top operational speeds between 300km/h and 350km/h, passengers can cover large distances within a short amount of time. For instance, the HSR takes 1h47min to cover slightly more than 400km between Paris and Lyon. Whereas HSR in France is not cheap, it remains accessible to most families, as well as for people travelling for business.[8] Relative to air travel, the HSR network has the significant advantage of departing and arriving in the city center. Hence, in many cases, travelling by HSR will be faster than air travel when measuring travel time door-to-door. Finally, HSR is more expensive than the long-distance bus network, but much faster.

**The 2017 high-speed rail extension.** Since July 2017, the high-speed network connects Bordeaux and Rennes to Paris, respectively. Bold yellow lines in Figure 5 locate the HSR extensions. The HSR connection between Paris and Bordeaux implied a *35% reduction in travel time from 3h 12min to 2h 04min*. Similarly, the connection between Paris and Rennes implied a *32% reduction in travel time from 2h 04min to 1h 25min*. In Figure 5, Bordeaux and Rennes are marked with a red square and a red triangle, respectively. As illustrated in Figure A7, Bordeaux and Rennes host a single HSR station. Rennes train station is located in the very center, whereas Bordeaux' HSR station is slightly south of the center. In both case, the HSR stations can easily be reached from anywhere in the city.

**Specificities of HSR in Paris.** Paris hosts four HSR stations (cf. Figure 6). Each connects to different regions of the country. Station "*Montparnasse*" in the south-west of Paris is the only one connecting to destinations in the West and South-West of the country, *including Bordeaux and Rennes*. Note that provincials residing in Paris are over-proportionally located near the station connecting to their place of origin. For in-

[8] Large families benefits from important reductions when using the French National Railway network ranging from 30% to 75%. Benefits are function of the number of under-18 children in the household. For instance, with three children, ticket prices are discounted by 30%.

Figure 5: HSR extensions on July 1st, 2017



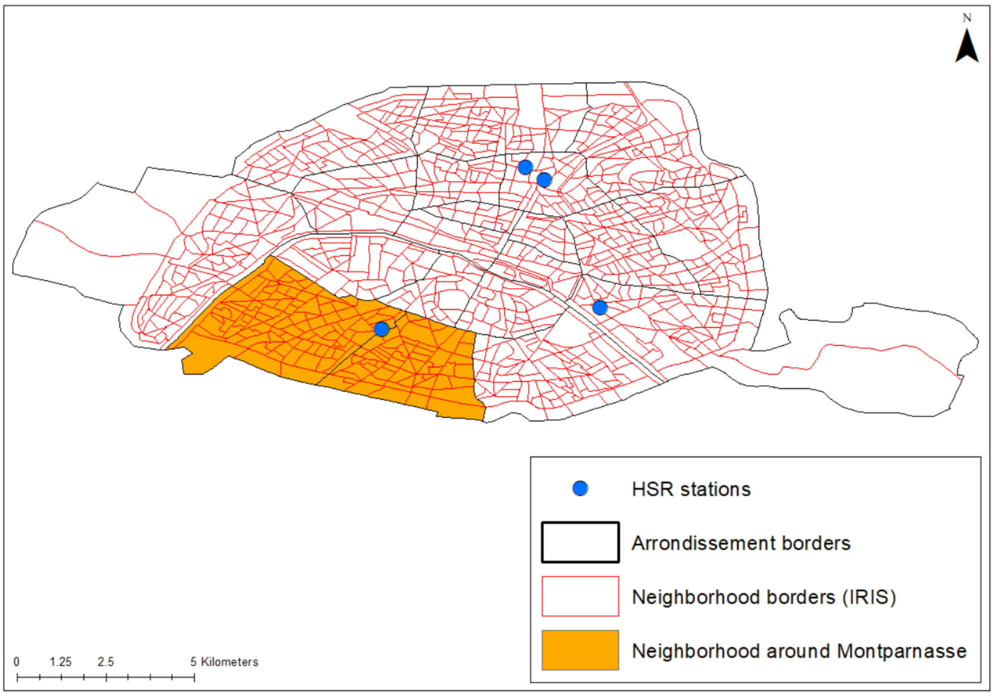

*Notes:* Authors' own illustration based on shapefiles from `https://www.data.gouv.fr/en/datasets`. Incidentally treated metropolitan areas are: Poitiers, Angoulème, Laval and Vitré. Metropolitan areas considered with high potential for HSR by the French National Railway Company (SNCF) are: Le Havre, Biarritz, Chambery, Montpellier and Perpignan (see, Charnoz et al., 2018). "MA" stands for Metropolitan Area.

stance, individuals originating from the West of the country have over-proportionally located close to Montparnasse. This pattern is not new. Already in 1914, Gallouédec noted that "*the provincials are located in Paris by district. They generally settle near the stations where the lines leading to their province of origin end: the Bretons near the Montparnasse station, ...*". Figure A6 highlights the region of origin of provincials residing in Paris at the arrondissement level in 1914. As we argue below, this is one of the observations motivating our focus on the arrondissements close to Montparnasse HSR station when studying the effects of the HSR on outcomes in Paris.

## 3.2 Other metropolitan areas used for identification

Beyond a primary city (i.e., Paris) and secondary cities (i.e., Bordeaux and Rennes), we define three additional sets of metropolitan areas which will be used for identification in Section 4.

Figure 6: *Arrondissements*, Neighborhoods, and HSR stations in Paris




*Notes:* Authors' own illustration based on shapefiles from https://www.data.gouv.fr/en/datasets.

**Incidentally treated metropolitan areas.** Rennes and Bordeaux are not the only cities directly impacted by the new HSR lines. On the way to Bordeaux, Poitiers and Angoulème (green pentagon in Figure 5) got connected to the HSR network. Similarly, on the way to Rennes, Laval and Vitré (green pentagon in Figure 5) also became HSR stations. These four cities were treated, but not primarily targeted. They received a HSR connection only thanks to their geographical placement on the way to either Bordeaux or Rennes. Such setting – often labelled *incidental* or *inconsequential* treatment – has commonly been used in the literature to identify transport investments effects free of endogeneity problems due to possible non-random infrastructure placement (see, among others, Chandra and Thompson, 2000; Holl, 2004; Melo et al., 2010; Ghani et al., 2016; Ahlfeldt and Feddersen, 2017b; Gibbons et al., 2019).

**Metropolitan areas with high HSR potential.** The French National Railway Company (SNCF) has listed cities with high potential for high-speed rail: Le Havre, Biarritz, Chambery, Montpellier and Perpignan. Updating the rail network to these cities – highlighted by purple hexagon in Figure 5 – is not currently in progress, but it may be in the near future. The choice of these cities is mostly due to either a proximity to Paris without near HSR alternative (i.e., Le Havre) or to international travel possibilities towards Spain and Italy.[9]

**Large French metropolitan areas.** Finally, Figure 5 illustrates the location of the top 10 largest metropolitan areas in France. When using this set of metropolitan areas in Section 4 to build a control group, we exclude Nantes and Toulouse. The former is excluded because of its historical rivalry with Rennes; the second for its historical rivalry with Bordeaux. Hence, it appears likely that possible HSR effects on Rennes/Bordeaux would also affect Nantes/Toulouse. Strasbourg is also excluded from the control group because it connected via HSR in 2016. We do not include it in the treatment group as the travel time gain in the case of Strasbourg was much lower (<15%).

## 3.3 Data

We primarily draw on four local data-sets. In what follows, we provide a short summary of the different data sources and the data construction processes.[10]

**Housing transaction data.** We use the official housing price data on the universe of housing property sold in France between Jan 1st, 2014 and Dec 31st, 2020.[11] This data-set is a repeated cross-section produced and made publicly available by `https://app.dvf.etalab.gouv.fr/`. Table A1 presents simple descriptive statistics of the data – focusing on the number of observations, as well as the mean and standard deviation of the studied variables. For each transaction, we know the price in €/m$^2$, the size (i.e., floor space in m$^2$), the number of rooms, and the type (house or apartment). Transaction date refers to the date of the change in ownership. For the period Jan 1st, 2016 to Dec 31st, 2020, the data also includes the exact coordinates of each property. To illustrate the within-city coverage and scale of the database used, Figure A8 illustrates the location of each transaction observed in Paris, Rennes and Bordeaux separately.

To study the evolution of housing prices, we use primarily two metrics. First, we look at the *transaction price* in € per m$^2$, which is the price at which the property was effectively sold. Second, we study the *hedonic price* which we compute by regressing

[9] Charnoz et al. (2018) use this setting – on an earlier period – to identify the effect of communication costs on the organization multi-plant firms.

[10] All data used are publicly and freely available.

[11] Time fixed effects are used to control for the arrival of COVID-19 in France in S1 and S2 2020. Before 2020, it appears safe to assume that economic agents didn't anticipate the arrival of the pandemic.

separately for each period the transaction price per $m^2$ on the characteristics of the property: type of housing (i.e., house or apartment), size of the housing unit, number of rooms, size of garden.[12] We then extract the residual – to which we add the regression constant – as outcome. The hedonic price can be seen as the price of a reference dwelling.

**Municipality-to-municipality migration data.** We use the French National Institute of Statistics and Economic Studies (INSEE) record on individuals' location of residence at year $t$ and the location of residence on January 1st at year $t-1$. The universe of individuals is covered each year between 2013 and 2018. *Locations are defined at the municipal level* (and *arrondissements* for Paris, see Figure A7a). Individual information further records the level of education for each individual. We use this information to define as skilled all individuals with a tertiary education. Finally, note that, as we look at dense urban municipalities, mobility flows are sufficiently large that issues of censoring (from below) due to confidentiality considerations are not a concern in the present setting.

**Municipality-to-municipality commuting data.** We also measure the municipality of residence and municipality of work for the universe of workers in France on a yearly basis between 2013 and 2018. As in the migration data, individual information further records the level of education for each individual. We use this information to define as skilled all individuals with a tertiary education.

**Residential density data.** We make use of the INSEE residential database which records individual information at the place of residence at the level of the city neighborhood (formally labeled *IRIS*). Neighborhoods are much more fine grained units than municipalities, as illustrated in Figure A7. The INSEE delinaetes IRIS such that: "*population generally falls between 1,800 and 5,000. The unit is homogeneous in terms of living environment and the boundaries of the unit are based on the major dividing lines provided by the urban fabric (main roads, railways, bodies of water etc).*"[13] For each individual, we then know their IRIS of residence, level of education, their inter- and intra-national migratory background, etc. Relative to the municipality-to-municipality migration data, the residential database is more precise on the location

[12] On the hedonic price index approach using French housing data, see Gourieroux and Laferrere (2009), Musiedlak and Vignolles (2016), Combes et al. (2019), and Tricaud (2021).

[13] *Source:* `https://www.insee.fr/en/information/2568929`.

of residence but only records the administrative region (NUTS2) of residence at $t-1$.[14]

**Miscellaneous.** Finally, we complement these data-set with municipal level information, including population count, area and age composition.

# 4 Empirical Analysis

In this section, we exploit the extension to the French HSR network in July 2017 to test the hypotheses listed in Section 2.6, using the data described above. For each hypothesis, we start by describing our identification strategy, before presenting the results.

## 4.1 Are housing prices increasing in Bordeaux/Rennes? (H1)

Hypothesis 1 states that *opening a HSR line should increase housing prices in the secondary cities*.

**Identification strategy.** We adopt an Event-Study Difference-in-Difference (ES-DiD) identification strategy following Schmidheiny and Siegloch (2019). Over the standard Difference-in-Difference approach, this strategy is able to capture and illustrate precisely the timing of the effect. Formally, the treatment effect is allowed to vary over time. We are then interested in studying its dynamics over a window ranging from $\underline{j} < 0$ periods preceding the event to $\overline{j} > 0$ after the treatment. $\beta_j, \forall j \in (\underline{j}; \overline{j})$ are then the coefficients of interest. We estimate the following empirical model for housing prices:

$$p_{oit} = \sum_{j=\underline{j}}^{\overline{j}} \beta_j T_{it}^j + \boldsymbol{X}'_{oit}\gamma + \mu_i + \theta_t + \varepsilon_{oit}, \tag{19}$$

where $p_{oit}$ is the (transaction or hedonic) price (in € per m$^2$) of unit $o$ in city $i$ at semester $t$, $T_{it}^j$ are $j-$specific interactions between a time indicator and the treatment status indicator, $\boldsymbol{X}_{oit}$ refers to a set of housing unit specific characteristics, $\mu_i$ and $\theta_t$ are city- and time- specific fixed effects, respectively. $\varepsilon_{oit}$ is a error term. When presenting the results below, we also report the value of a $\beta$ coefficient referring to

[14]Mainland France is divided into 12 administrative regions. Paris is in the main city in the Ile-de-France region, Bordeaux is the main city in the Nouvelle Acquitaine region and Rennes is the main city in the Brittany region.

the DiD coefficient, i.e., when estimating a simpler version of (19) with a treatment-post HSR extension dummy, instead of treatment-semester-specific dummies.[15]

Our treatment group is made by all housing transactions in Bordeaux and Rennes. We consider two control groups. First, we use all other cities among the 10 largest cities in France (i.e., Marseille, Lyon, Lille, and Nice), excluding Paris, Nantes, Toulouse and Strasbourg.[16] As alternative control group, we consider all cities labeled as "high-HSR-potential" by the French National Railway Company (SNCF), i.e., Le Havre, Biarritz, Chambéry, Perpignan and Montpellier. The use of this control group is motivated by the fact that there may be similar underlying socio-economic forces that lead a city to receive an HSR connection. As these cities are *next in line*, they constitute a suitable control group to account for such effects.

**Results.** Figure 9 reports the evolution of the effect of the HSR opening on housing prices in Bordeaux and Rennes. The black estimates refer to the first control group and the grey estimates refer to the second one. Housing prices are defined in € per $m^2$ on a semester basis between January 2014 and December 2020. Panel (a) and (b) use the transaction and the hedonic price as outcome, respectively.

Overall, we observe – independently of the control group – that the HSR opening had a strong and sharp effect on housing prices. Whereas trends are flat in S1 and S2 2016, housing prices increase by €200 per $m^2$ already in S1 2017. The effect then reaches an average increase of €400 per $m^2$.[17]

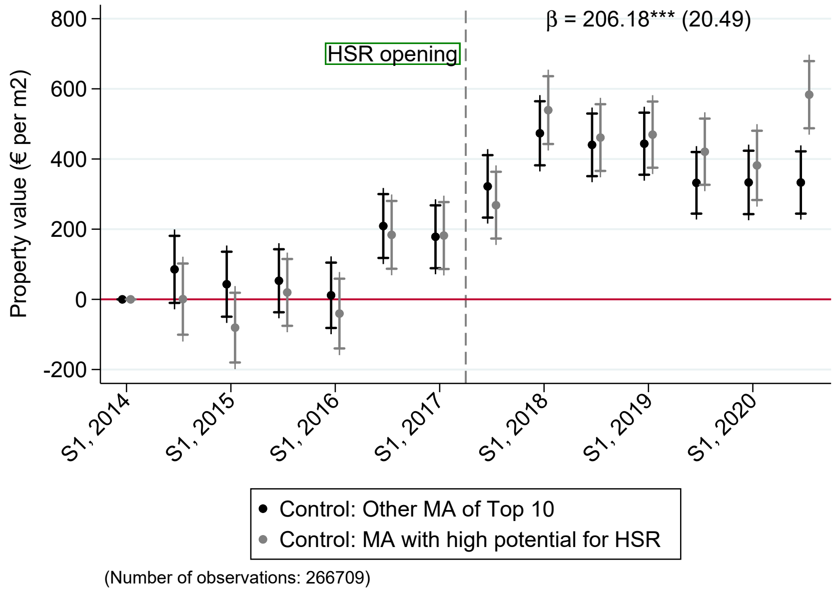
[15] As is standard when studying housing prices locally, (19) implies that the time-specific treatment effects on housing prices (€/$m^2$) are estimated based on different transactions. To reduce biases arising due to differences in housing units on the market over time, we already account for housing unit specific characteristics via unit specific controls or the hedonic price approach. An alternative to these approaches is to study the evolution of average housing price aggregated at the level of grid cells. We use grid cells of $100m^2$ and $1km^2$. Results are discussed below and presented in Figures A14 (for Bordeaux and Rennes) and A15 (for Paris). Results using grid cell aggregation provide qualitatively and quantitatively similar results.

[16] As described in Section 3.2, whereas Paris is excluded for clear reasons, we also exclude Nantes and Toulouse as they are historical rivals of Rennes and Bordeaux, respectively. The rivalry between each pair may induce important externalities which could bias the results. Strasbourg is also excluded from the control group because it connected via HSR in 2016. We do not include it in the treatment group as the travel time gain in the case of Strasbourg was much lower (<15%).

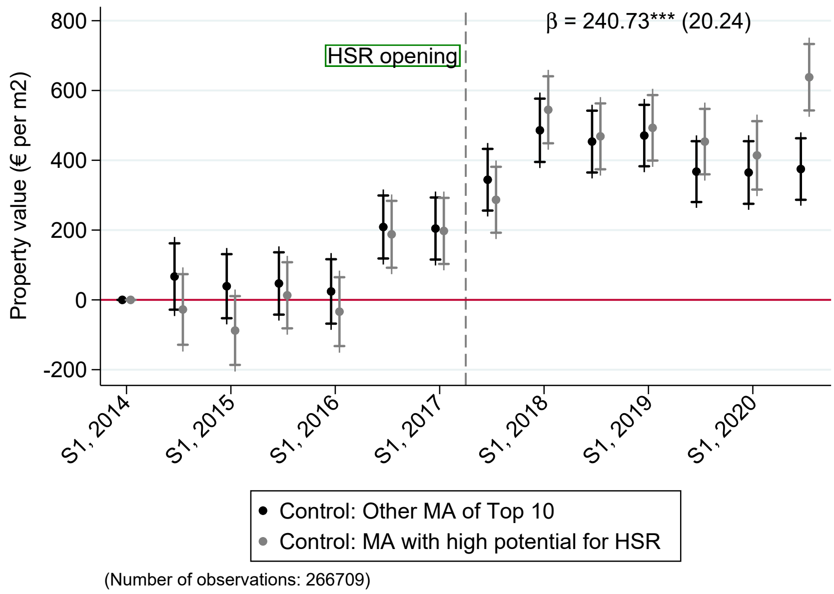
[17] Figure A14 report qualitatively and quantitatively similar results when using average housing prices within $100m^2$ (Figure A14a) and $1km^2$ (Figure A14b) grid cells.

Figure 7: Average HSR effect on housing prices in Bordeaux and Rennes (€ per $m^2$)

(a) Transaction price




(b) Hedonic price




*Notes:* ES-DiD model (19) using 95% CIs. Property values (on a quarterly basis) in Bordeaux and Rennes are compared to property values in Marseille, Lyon, Lille, and Nice (i.e., all other cities of the top 10 French largest cities excluding Nantes, Paris and Toulouse for incidental treatment reasons). Strasbourg is also excluded from the control group because it connected via HSR in 2016. The *transaction price* is the price at which the property was effectively sold. The *hedonic price* is computed by regressing separately for each period the transaction price per $m^2$ on the characteristics of the property: type of housing (i.e., house or apartment), size of the housing unit, number of rooms, size of garden. We then extract the residual – to which we add the regression constant – as outcome.

**Housing market effects by type of housing units.** The granularity of the housing transaction data allows us to study whether the treatment effect differs by type of housing unit. Figure A10 studies the effect of the HSR extension on the value of apartments (A10a) and houses (A10b). We use again transaction prices as outcomes. As in Figure 7, we use two different control groups: the largest non-treated French cities, and cities labeled high-HSR potential by the French national railway company. Overall, the effect appears significantly larger for houses, which is consistent with the hypothesis that households moving away from Paris put a premium on larger dwelling space.

**Housing market effects on incidentally treated secondary cities.** The flat pre-trends observed in 2014, 2015 and S1 of 2016 in Figure 7 are reassuring regarding endogeneity concerns due to non-random HSR placement. Yet, a further test of the effect of HSR on secondary cities can be performed by focusing on incidentally treated metropolitan areas. As discussed in Section 3, incidentally treated secondary cities are cities that received a HSR connection *solely because they are located on the way between Paris and Bordeaux/Rennes*.[18]

Figure A9 presents the results of estimating (19) with incidental cities as treatments units. Following the arrival of HSR, housing prices have increased in incidentally treated metropolitan areas. However, with an average treatment effect of €69 per $m^2$, the magnitude of the treatment effect is smaller than in Bordeaux and Rennes. Two reasons may explain this smaller magnitude. First, the frequency of high-speed trains to Paris is smaller; hence, a smaller treatment effect. Second, though Rennes and Bordeaux are much smaller than Paris, both still offer a relevant bundle of urban amenities (i.e., exhibitions, theaters, concerts, ...). This is not the case of Laval, Vitré, Poitiers or Angoulème. Shorter travel times via HSR and lower housing prices may not be sufficient to attract as many parisians to small urban centers; hence, the smaller magnitude of the effect.

## 4.2 Are housing prices affected in Paris? (H2)

Hypothesis 2 states that *opening a HSR line should have a negative effect on housing prices in the primary city*.

[18]These are: Poitiers and Angoulème (on the way to Bordeaux), and Laval and Vitré (on the way to Rennes).

**Identification strategy.** To study the effect of the HSR on housing prices in Paris, we estimate the same model as in (19), using again transaction and hedonic prices alternatively as outcomes. The definition of the treatment and control groups, however, differ. In this part of the analysis, we compare housing prices in the arrondissements around Montparnasse station (i.e., $14^{\text{th}}$ and $15^{\text{th}}$) to either the housing prices in all other arrondissements (control group 1) or to housing prices in the other arrondissements with an HSR station (control group 2). Recall that the other HSR stations in Paris are not connected to the new HSR lines. Figure 6 illustrates the urban structure of Paris, and locates the *arrondissements*, neighborhoods and HSR stations within Paris *intra muros*.

We focus exclusively on housing transactions within Paris to form our control groups. This is due to the particularly dynamic and competitive nature of the city's housing market, relative to other cities in France. Our choice of the arrondissements around Montparnasse station as treatment group is motivated by several reasons. To begin, our theoretical model suggests that, after the HSR line opens, the skilled individuals that move to the secondary city would otherwise have lived relatively close to the primary city's HSR station. Although the model is highly stylized, this prediction should hold more generally. The HSR connection makes the primary city more accessible from the secondary one, but this applies particularly to areas close to the "treated" train station in the primary (Montparnasse). Thus, the new connection makes the secondary city more attractive mainly to the individuals that value access to the area around that station (e.g., because they work nearby, even from time to time), and would otherwise have settled either in relative proximity to it, or close to good transport connections to such area, e.g., metro stations. Accordingly, we shall also estimate specifications of the model where our treatment group are housing units close to metro stations within different radii from Montparnasse.

An additional reason for our choice of treatment group is the demographic trend – observed since the late 19th century (Gallouédec, 1914) – whereby non-originally Parisian French workers moving to Paris tend to reside in the neighborhood close to the train station connecting to their location of origin. For instance, as discussed in Section 3, the neighborhood around Montparnasse has long hosted communities from Bretagne (the region in western France where Rennes is located). Individuals living in the Montparnasse area should thus be relatively more sensitive to the opportunity to relocate to such cities, compared to residents of other areas in Paris.

**Results.** Figure 8 studies the average HSR effect on housing prices in Paris around Montparnasse. Independent of the control group, we observe a significant negative effect on housing prices – by about €250 per m$^2$ – already in S2 2017. Pre-trends are flat between 2014 and S1 2017. Over time, the effect appears stable. This effect is directly in line with the theoretical predictions, and thus confirms hypothesis H2.

Importantly, the negative effect estimated is relative to the control groups. In absolute terms, housing prices in the 14th and 15th arrondissements have increased even after July 2017, *but less than in the other arrondissements*. As intuitive illustration, the simple average transaction price per m2 in the 14th and 15th arrondissements was about €8,710 before July 2017 and €10,060 after. In the rest of Paris, it was €8,690 before and €10,330 after.[19] This simple average approach is qualitatively and quantitatively in line with the results in Figure 8.

The analysis at the 100m$^2$ and 1km$^2$ cell level – reported in Figure A15 – confirms this result. As Paris' housing market is mostly constituted of apartments (see Table A1), the effect is driven by apartment transactions.

**How spatially spread is the effect in Paris?** In Figure 8, we used *arrondissement* boundaries to define the treatment status. However, the effect could reach (or go beyond) arrondissement borders. To study how far in space does the effect in Paris spread, we estimate a set of treatment effects defined as the interaction of a post dummy (equals to unity from S2 2017 onwards, and 0 otherwise) and 1km rings of Paris Montparnasse HSR station. Locations at more that 6km of Paris Montparnasse constitute the omitted category. Transaction prices are used as outcomes.

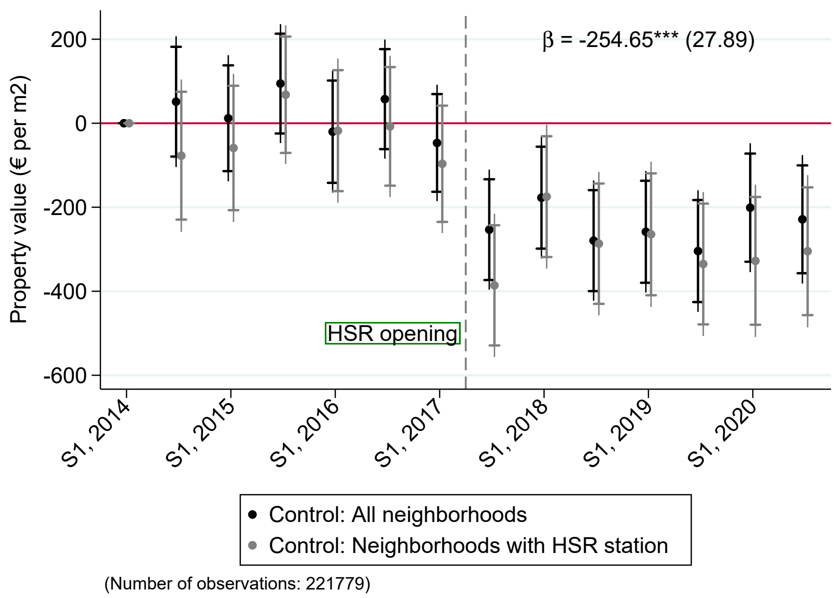
Figure A11 reveals that the negative effect on housing prices is Paris is not restricted to the absolute vicinity of Montparnasse. Instead, we observe the effect on housing prices within 3km of the station as opposed to locations further away. Yet, not all housing units within 3km of Montparnasse have experienced a similar effect. Figure A12 reveals that the effect on housing prices is mostly driven by location within 100m of a metro station on a direct line to Montparnasse.[20]

[19] Before refers to the period January 2014 to June 2017, and after to the period July 2017 to December 2020.

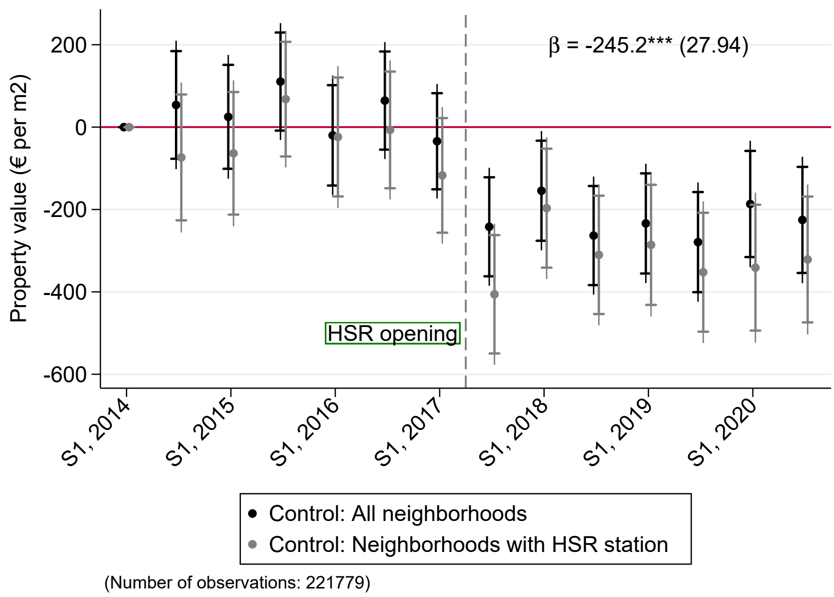
[20] Direct metro lines to Montparnasse are lines 4, 6, 12 and 13.

Figure 8: Average HSR effect on housing prices around Gare Montparnasse (Paris' train station to Rennes and Bordeaux, € per m²)

(a) Transaction price




(b) Hedonic price




*Notes:* ES-DiD model (19) using 90% and 95% CIs. Figure studies the average HSR effect on housing prices in Paris around Montparnasse. We alternatively define the control group as all other Parisian arrondissements, and all other arrondissements hosting a HSR station. The *transaction price* is the price at which the property was effectively sold. The *hedonic price* is computed by regressing separately for each period the transaction price per $m^2$ on the characteristics of the property: type of housing (i.e., house or apartment), size of the housing unit, number of rooms, size of garden. We then extract the residual – to which we add the regression constant – as outcome.

## 4.3 Are Parisians more likely to move to Bordeaux/Rennes? (H3)

Hypothesis 3 states that *opening a HSR line should induce migration of skilled workers to the secondary city from the primary one*.

**Identification strategy.** We adopt a Triple Difference (TD) identification strategy in which we study the residential flow of skilled workers from Paris to Bordeaux/Rennes relative to the same flows to other cities in France and relative to unskilled individuals. This strategy aims at capturing the increased flow of *skilled workers to Bordeaux and Rennes* in 2017 and afterwards. The strategy allows to distinguish the effect of the HSR on the propensity of skilled workers to move to Bordeaux and Rennes, from the propensity of skilled workers to move in general (relative to unskilled workers).

Denote the skill level by $\omega$, the location of residence at year $t$ by $i$, and the location of residence at $t-1$ by $i'$. Then $y_{\omega i' it}$ is the flow of workers of skill $\omega$ who moved from $i'$ to $i$ between years $t-1$ and $t$. $Skill_\omega$ is a dummy equal to one if individuals hold a tertiary education degree, $HSR_i$ a dummy equal to one if the destination of the move is a HSR-treated secondary city (i.e., Bordeaux or Rennes), $D_t$ is a dummy equal to unity for years after (and including) 2017 and zero otherwise. We estimate the following empirical model for outcome $y_{\omega i' it}$:

$$\begin{aligned} y_{\omega i' it} = {} & \alpha_1(D_t \times Skill_\omega) + \alpha_2(D_t \times HSR_i) + \alpha_3\big(Skill_\omega \times HSR_i\big) \\ & + \alpha_4\big(D_t \times Skill_\omega \times HSR_i\big) + \mu_i + \nu_{i'} + \theta_t + \rho_\omega + \varepsilon_{\omega i' it}, \end{aligned} \tag{20}$$

where $\alpha_4$ is the parameter of interest. $\mu_i$, $\nu_{i'}$, $\theta_t$ and $\rho_\omega$ are destination-, origin-, year- and skill-specific fixed effects, respectively. $\varepsilon_{\omega i' it}$ is the error term.

Consistently with our approach in 4.2, we concentrate on residential origin locations in Paris within the arrondissements around Montparnasse HSR station (14$^{\text{th}}$ and 15$^{\text{th}}$). Furthermore, in line with our approach in Section 4.1, we define the alternative option to Bordeaux and Rennes using two groups of cities. First, we consider other cities of the top 10 largest French MSA. Second, we consider cities labeled *high-HSR potential* by the French national railway company.

Moreover, to precisely understand the effect of the HSR on relocation, we complement the analysis on residential mobility flows with an analysis of the commuting flows of skilled workers between Bordeaux/Rennes and Paris. This aims at estimating the share of workers that relocated "residentially" to Bordeaux/Rennes, but kept their official workplace in Paris. To do so, we estimate (20) with $i'$ denoting the residence location and $i$ the workplace location.

Table 1: Migratory and commuting flows of skilled workers

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Outcome: | **Migration flows** | | **Commuting flows** | |
| Control: | Top 10 MA | HSR potential | Top 10 MA | HSR potential |
| DiDiD ($\alpha_4$) | 15.206* | 24.011** | 2.256* | 3.158* |
| | (8.024) | (12.060) | (1.309) | (1.773) |
| Avg. flow (pre-HSR) | 15.198 | 24.765 | 5.591 | 6.907 |
| Obs. | 494 | 131 | 2,493 | 977 |
| $R^2$ | 0.78 | 0.83 | 0.37 | 0.49 |

*Notes:* Triple-difference model (20). Avg. flow (pre-HSR) is computed across all years between 2013 and 2016. Parisian arrondissements around Montparnasse ($14^{th}$ and $15^{th}$) constitute the origin locations. Furthermore, in line with our approach in Section 4.1, we define the alternative option to Bordeaux and Rennes, using two groups of cities. First, we consider other cities of the top 10 largest French MSA (i.e., Marseille, Lyon, Lille, and Nice). Second, we consider cities labeled *high-HSR potential* by the French national railway company (i.e., Le Havre, Biarritz, Chambery, Montpellier and Perpignan). Robust standard errors in parentheses.

**Results.** Table 1 presents the results on residential mobility (columns 1 and 2) and commuting (columns 3 and 4) flows. For interpretation purposes, we also report the average flow size for each sample defined on the pre-HSR period (i.e., 2013-2016). Periods are defined on a yearly basis between 2013 and 2018. The table reveals that the yearly residential relocation flow of skilled workers from the *arrondissements* around Montparnasse to Bordeaux/Rennes increased significantly in 2017. We estimate a triple interaction effect on the average flow between 15 and 24 workers, which implies an almost 100% increase in the relocation flows of skilled workers. This finding confirms hypothesis H3. At the same time, the commuting flows from Bordeaux/Rennes to Paris increased by 2 to 3 workers for the average bilateral flow.[21] This corresponds

[21] As workplace location in Paris, we consider all locations within Paris, but also to Paris central business district: *La Défense*.

to a 50% increase in the flow of long distance commuters. Furthermore, it also implies that the share of movers that kept their main workplace in Paris is (at most) 15%.

## 4.4 Do in-moving skilled workers locate in central locations? (H4)

Hypothesis 4 states that *the HSR line should increase the share of skilled individuals moving into the central neighborhoods of the secondary city and discourage "native" individuals from moving there.*

**Identification strategy.** To empirically test hypothesis 4, we adopt an ES-DiD identification strategy as in Section 4.1. Using a similar notation, we estimate the following empirical model for outcome $y_{it}$:

$$y_{it} = \sum_{j=\underline{j}}^{\bar{j}} \beta_j T_{it}^j + \mu_i + \theta_t + \varepsilon_{it}. \tag{21}$$

The outcome variable $y_{it}$ is the share of skilled workers among in-movers into a neighborhood of Bordeaux and Rennes.[22] To define neighborhoods, we follow the INSEE definition of IRIS (cf. Section 3). Central locations in Bordeaux and Rennes are alternatively defined as the three and five most central neighborhoods in those cities.[23]

The variables $T_{it}^j$ are $j-$specific interactions between a yearly indicator and the treatment status indicator. $\mu_i$ and $\theta_t$ are neighborhood- and time- specific fixed effects, respectively. $\varepsilon_{it}$ is an error term. Equation (21) has a similar structure to (19), though the units of observation are individuals and not housing transactions. Another important difference is that we can estimate (21) following the same individuals over time.

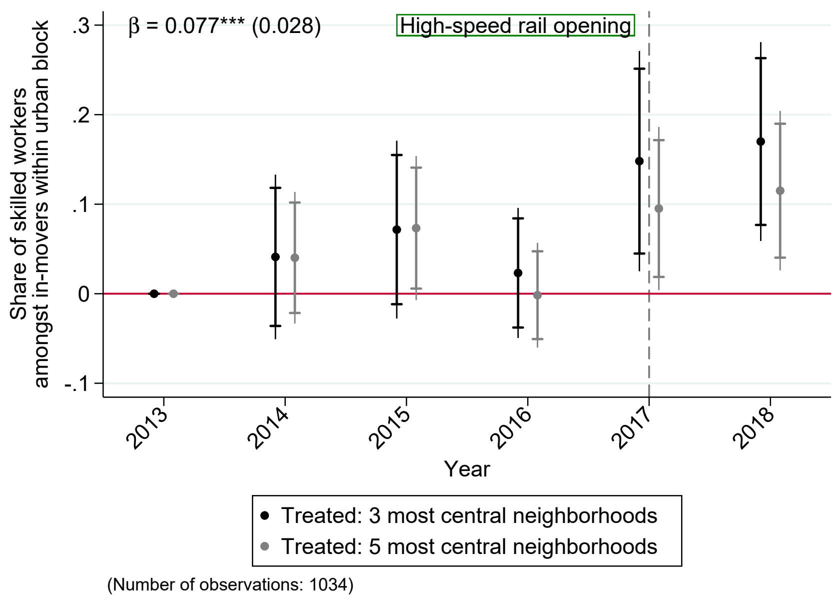
**Results.** Figure 9 compares the composition of in-movers to Bordeaux and Rennes between central and non-central locations. Neighborhoods in the very center of Rennes and Bordeaux experience a 10 percentage point increase in the share of skilled

[22]Note that, given that we study neighborhoods within Bordeaux/Rennes, the available data does not allow to study movers strictly from Paris. Instead, we study the flow of in-movers from one of the non-Bordeaux/Rennes French regions (Paris included).

[23]We use the IRIS last two digits to rank central neighborhoods.

among in-movers in 2017. This coincides with the opening of the new HSR line. Interestingly, the effect dissipates when increasing the number of neighborhoods composing the "city center"; thus, highlighting the very concentrated nature of location decisions by skilled workers.

Figure 9: Share of skilled among in-movers to central locations in Bordeaux and Rennes




*Notes:* ES-DiD model (21) using 95% CIs. The Figure compares the composition of in-movers to Bordeaux and Rennes between central and non-central locations. Outcome is the share of skilled in-movers in the total of in-movers to the neighborhood.

Figure 9 focuses on the in-movers from outside the region of Bordeaux and Rennes. As a complement to this analysis, in Figure A13 we study the probability that local individuals move from the greater periphery to the municipalities of Bordeaux and Rennes. Specifically, we estimate a similar equation to (21).[24] Consistently with our theory, we find that, from 2017 onwards, locals from the periphery are significantly less likely to move towards the municipalities of Bordeaux and Rennes. We estimate the decrease to reach -9 percentage points already in 2018. This effect is key in understanding the social unrest and local protest that followed the opening of the HSR line, and documented in the press (e.g., Figures 1 and A1).

[24]Greater periphery of Bordeaux and Rennes is defined as their respective departments. This definition is more generous than the metropolitan area definition and insures that the results are not affected by the arbitrary commuting threshold used to delineate metropolitan areas.

# 5 Conclusion

This paper investigates whether and how transport infrastructure induces gentrification to spread across cities. We consider a model with a primary and a secondary city, with intra- and inter-city commuting. Individuals differ in their skill (wage) level and preference for living in the primary city, with skilled individuals in the secondary city commuting – at least infrequently – to the primary one. The HSR line reduces the cost of long-distance travel, inducing migration of skilled individuals towards the secondary city. Hence, floor prices increase therein, forcing unskilled individuals to either move to the periphery of the secondary city or to migrate to the primary one. We call this effect *second-hand gentrification*. Interestingly, the model predicts also a negative effect on housing prices in the primary city.

We confirm these predictions empirically by exploiting the July 2017 opening of HSR lines between the cities of Paris, Bordeaux and Rennes. Using data on the universe of housing transaction in France, we show that housing prices have increased in Bordeaux and Rennes by €400 per m2 (secondary cities), but were negatively affected in relative proximity to the HSR station of Paris by €245 per m2 (primary city). We further show that, following the HSR opening, skilled parisians have been more likely to relocate to Bordeaux and Rennes; and that, when relocating, they reside in the very center of these cities.

The estimation period in this paper precedes the COVID-19 health crisis. Yet, the development of *remote working* that took place during this crisis provides further motivation for the analysis presented. Indeed, this paper would suggest that, if the need to commute to work decreases, second-hand gentrification is likely to strengthen and affect all cities within reasonable distance of large employment centers.

## References


Ahlfeldt, G. M., and Feddersen, A. (2017a). “From periphery to core: measuring agglomeration effects using high-speed rail.” *Journal of Economic Geography*, *18*(2), 355–390.

Ahlfeldt, G. M., and Feddersen, A. (2017b). “From periphery to core: measuring agglomeration effects using high-speed rail.” *Journal of Economic Geography*, *18*(2), 355–390.

Banerjee, A., Duflo, E., and Qian, N. (2020). “On the road: Access to transportation infrastructure and economic growth in china.” *Journal of Development Economics*, *145*, 102442.

Baum-Snow, N. (2007). “Did Highways Cause Suburbanization?*.” *The Quarterly Journal of Economics*, *122*(2), 775–805.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., and Zhang, Q. (2017). “Roads, Railroads, and Decentralization of Chinese Cities.” *The Review of Economics and Statistics*, *99*(3), 435–448.

Baum-Snow, N., and Pavan, R. (2011). “Understanding the City Size Wage Gap.” *The Review of Economic Studies*, *79*(1), 88–127.

Behrens, K., Duranton, G., and Robert-Nicoud, F. (2014). “Productive cities: Sorting, selection, and agglomeration.” *Journal of Political Economy*, *122*(3), 507–553.

Borck, R., and Wrede, M. (2005). “Political economy of commuting subsidies.” *Journal of Urban Economics*, *57*(3), 478–499.

Borck, R., and Wrede, M. (2008). “Commuting subsidies with two transport modes.” *Journal of Urban Economics*, *63*(3), 841–848.

Borck, R., and Wrede, M. (2009). “Subsidies for intracity and intercity commuting.” *Journal of Urban Economics*, *66*(1), 25–32.

Brueckner, J. K., and Rosenthal, S. S. (2009). “Gentrification and Neighborhood Housing Cycles: Will America’s Future Downtowns Be Rich?” *The Review of Economics and Statistics*, *91*(4), 725–743.

Brueckner, J. K., and Selod, H. (2006). “The political economy of urban transport-system choice.” *Journal of Public Economics*, *90*(6-7), 983–1005.

Chandra, A., and Thompson, E. (2000). "Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system." *Regional Science and Urban Economics*, *30*(4), 457–490.

Charnoz, P., Lelarge, C., and Trevien, C. (2018). "Communication costs and the internal organisation of multi-plant businesses: Evidence from the impact of the french high-speed rail." *The Economic Journal*, *128*(610), 949–994.

Combes, P.-P., Duranton, G., and Gobillon, L. (2008). "Spatial wage disparities: Sorting matters!" *Journal of Urban Economics*, *63*(2), 723–742.

Combes, P.-P., Duranton, G., and Gobillon, L. (2019). "The Costs of Agglomeration: House and Land Prices in French Cities." *Review of Economic Studies*, *86*(4), 1556–1589.

De Borger, B., and Russo, A. (2018). "The political economy of cordon tolls." *Journal of Urban Economics*, *105*, 133–148.

Ding, L., Hwang, J., and Divringi, E. (2016). "Gentrification and residential mobility in philadelphia." *Regional Science and Urban Economics*, *61*, 38–51.

Donaldson, D. (2018). "Railroads of the raj: Estimating the impact of transportation infrastructure." *American Economic Review*, *108*(4-5), 899–934.

Duranton, G., and Turner, M. A. (2012). "Urban Growth and Transportation." *The Review of Economic Studies*, *79*(4), 1407–1440.

Egger, P., Loumeau, G., and Loumeau, N. (2020). "China's dazzling transport-infrastructure growth: Measurement and effects."

Gallouédec, L. (1914). *Nouveau cours de géographie*. Hachette.

Ghani, E., Goswami, A. G., and Kerr, W. R. (2016). "Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing." *Economic Journal*, *126*(591), 317–357.

Gibbons, S., LyytikÃ¤inen, T., Overman, H. G., and Sanchis-Guarner, R. (2019). "New road infrastructure: The effects on firms." *Journal of Urban Economics*, *110*, 35–50.

Gourieroux, C., and Laferrere, A. (2009). "Managing hedonic housing price indexes: The french experience." *Journal of Housing Economics*, *18*(3), 206–213.

Guerrieri, V., Hartley, D., and Hurst, E. (2013). “Endogenous gentrification and housing price dynamics.” *Journal of Public Economics*, *100*, 45–60.

Hayakawa, K., Koster, H. R., Tabuchi, T., and Thisse, J.-F. (2021). “High-speed Rail and the Spatial Distribution of Economic Activity: Evidence from Japan’s Shinkansen.” Discussion papers 21003, Research Institute of Economy, Trade and Industry (RIETI).

Holl, A. (2004). “Manufacturing location and impacts of road transport infrastructure: empirical evidence from spain.” *Regional Science and Urban Economics*, *34*(3), 341–363.

Koster, H. R., Tabuchi, T., and Thisse, J.-F. (2021). “To be connected or not to be connected? The role of long-haul economies.” *Journal of Economic Geography*.

Lawrence, M., Bullock, R., and Liu, Z. (2019). *China’s High-Speed Rail Development*. World Bank.

Melo, P., Graham, D., and Noland, R. (2010). “Impact of transport infrastructure on firm formation: Evidence from the portuguese municipalities.” *Transportation Research Record Journal of the Transportation Research Board*, *2163*, 133–143.

Michaels, G. (2008). “The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System.” *The Review of Economics and Statistics*, *90*(4), 683–701.

Musiedlak, Y., and Vignolles, B. (2016). “Les mouvements des prix immobiliers dans l’ancien au cours des annÃ©es 2000: des marchés locaux différenciés.” *Document de travail du SOeS*, *22631*.

Schmidheiny, K., and Siegloch, S. (2019). “On Event Study Designs and Distributed-Lag Models: Equivalence, Generalization and Practical Implications.” Tech. rep.

Tricaud, C. (2021). “Better Alone? Evidence on the Costs of Intermunicipal Cooperation.” Tech. rep.

Zheng, S., and Kahn, M. E. (2013). “Does government investment in local public goods spur gentrification? evidence from beijing.” *Real Estate Economics*, *41*(1), 1–28.

# Appendix for
# "*Second-Hand Gentrification: Theory and Evidence from High-Speed Rail Extensions*"
# (for online publication only)

G. Loumeau and A. Russo

---

# A Coverage of second-hand gentrification (Press, politicians, NGOs)

Figure A1: Second-hand gentrification in the press (The Guardian, March 2018)

## Double trouble? How big cities are gentrifying their neighbours

▲ 'I see Bordeaux changing and not for the better'. ore than 70% of Bordeaux's new arrivals are from the Paris region. Photogra

From Bordeaux to Leipzig to Hamilton, cities are facing a new problem: secondhand gentrification thanks to more expensive cities nearby

As the spring sunshine beams down on the honey-coloured buildings of Bordeaux, few of the tourists on the terraces pay much attention to the sticker on the lamppost. "*Parisien rentre chez toi*," it declares - "Parisian go home" - accompanied by a graphic of the new high-speed train that now connects Bordeaux with the French capital in just over two hours.

Last summer's opening of the TGV route sped up more than just the travel time between the cities. Gentrification had long been underway in Bordeaux, a by-product of mayor Alain Juppé's much-vaunted regeneration of the city - including the cleaning up of the famed architecture, a new tram system and the mega-museum La Cité du Vin. Since the arrival of the high-speed rail link, however, well-heeled Parisians - lured by the sunny skies, slower pace of life and lower property prices - have been moving en masse.

Figure A2: Second-hand gentrification in the press (The News Tribune, August 2018)

LOCAL

**Would bullet trains spur affordable housing or gentrification along a Northwest line?**

BY JAMES DREW

UPDATED AUGUST 20, 2019 2:55 PM

Washington, B.C. want to make connection via high-speed rail

1:06

In this March 2018 file video, governments in Washington and British Columbia have committed money to study the feasibility of a high-speed rail line from Vancouver, B.C. to Portland. It would include stops in Olympia and Tacoma. BY JIM DONALDSON

Figure A3: Second-hand gentrification: The Response of Bordeaux's mayor (Europe 1, March 2018)

ACCUEIL / SOCIÉTÉ

**"La gentrification de Bordeaux est un fantasme", assure Alain Juppé**

14h33, le 16 mars 2018 , modifié à 14h58, le 16 mars 2018

AA

Figure A4: HSR and Gentrification (NRDC, June 2018)

NRDC

OUR WORK OUR EXPERTS OUR STORIES GET INVOLVED ABOUT US DONATE

onEarth › CULTURE & POLITICS

## When Public Transportation Leads to Gentrification

Why transit-oriented development projects need to include affordable housing amid all those luxury condos and cafés.

June 01, 2018 | Jeff Turrentine

Passengers at San Diego's MTS Green Line, Old Town Transit Center

# B Derivation of equilibrium population sizes, utility levels and effects of $k$

In this appendix, we derive the equilibrium population quantities (7), (8) and (9). We also characterize the equilibrium utility levels and the effects of changes in $k$ on such variables (expressions (15)-(18)). Note that we present these derivations in the generalized model where $w_{1,S} > w_{2,S}$. The expressions shown in the main text are obtained by replacing $w_{1,S} = w_{2,S} = w_S$.

In equilibrium, an individual living in a city must attain the same level of utility regardless of her distance from the CBD. To characterize this utility level, it is useful to consider individuals of each group living at the boundary of each city, where land rent equals zero. Combining the expressions for utility, starting from (6), with $n_i = \sum_{j=S,U} n_{ij}$, we can obtain the equilibrium population sizes in the two cities. Given fixed lots of unit size, if $n_i$ is the population in city $i$, it also equals the distance of the boundary from the CBD. Hence, we have $p_i(n_i) = 0$. Combining (6) with (1)-(3), we can thus write

$$V_{1,S}(n_1, z_S) = w_{1,S}(1 - tn_1) + z_S, \quad V_{1,U}(n_1, z_U) = w_U(1 - tn_1) + z_U, \tag{22}$$

$$V_{2,S}(n_2, z_S) = w_\alpha(1 - tn_2) - \alpha k w_{1,S}, \quad V_{2,U}(n_2, z_U) = w_U(1 - tn_2), \tag{23}$$

where $w_\alpha = \alpha w_{1,S} + (1-\alpha)w_{2,S}$. Recall that the idiosyncratic utility component $z_j$ is independent of distance from the CBD, $\forall j$. To determine the equilibrium allocation of individuals of group $j$ among cities 1 and 2, we first characterize the value of $z_j$ such that these individuals are indifferent between the two. Let $\bar{z_j}$ denote this value. Given the above expressions for utility, we have

$$\bar{z_S} = w_\alpha - w_{1,S} + t(w_{1,S}n_1 - w_\alpha n_2) - \alpha k w_{1,S}, \quad \bar{z_U} = w_U t(n_1 - n_2). \tag{24}$$

Since all skilled individuals such that $z_S \geq \bar{z_S}$ live in city 1, we get

$$n_{1S} = N_S \cdot Pr[z_S \geq \bar{z_S}] = N_S\left(1 - \frac{w_\alpha - w_{1,S} + t(w_{1,S}n_1 - w_\alpha n_2) - \alpha k w_{1,S}}{w_{1,S}}\right), \tag{25}$$

$$n_{2S} = N_S - n_{1S}. \tag{26}$$

Similarly, we get the following expressions regarding the unskilled group

$$n_{1U} = N_U \cdot Pr[z_U \geq \bar{z_U}] = N_U\left(1 - t(n_1 - n_2)\right), \quad n_{2U} = N_U t(n_1 - n_2). \tag{27}$$

Combining the above expressions with $n_i = \sum_{j=S,U} n_{ij}$, we obtain the following

$$n_1 = \frac{(N_S + N(1+tN_U))\, w_{1S} - N_S(1-tN)w_\alpha + kN_S\alpha w_{1S}}{(1+t(N_U+N))w_{1S} + tN_S w_\alpha}, \tag{28}$$

$$n_2 = \frac{(-N_S + tN^2)\, w_{1S} + N_S(w_\alpha - k\alpha w_{1S})}{(1+t(N_U+N))w_{1S} + tN_S w_\alpha}. \tag{29}$$

$$n_{1S} = N_S \frac{(2+tN_U(3-tN))w_{1S} - (1-tN-t(3-tN)N_U)w_\alpha + k(1+2tN_U)\alpha w_{1S}}{(1+t(N_U+N))w_{1S} + tN_S w_\alpha}, \tag{30}$$

$$n_{1U} = N_U \frac{((1-tN)^2 + t(3-tN)N_U)w_{1S} + tN_S((3-tN)w_\alpha - 2k\alpha w_{1S})}{(1+t(N_U+N))w_{1S} + tN_S w_\alpha}. \tag{31}$$

$$n_{2S} = N_S \frac{(-1+tN-t(2-tN)N_U)w_{1S} + (1+t(2-tN)N_U)w_\alpha - k(1+2tN_U)\alpha w_{1S}}{(1+t(N_U+N))w_{1S} + tN_S w_\alpha}, \tag{32}$$

$$n_{2U} = tN_U \frac{(-2N_U + N(3-tN+tN_U))w_{1S} + N_S((-2+tN)w_\alpha + 2k\alpha w_{1S})}{(1+t(N_U+N))w_{1S} + tN_S w_\alpha}. \tag{33}$$

Setting $w_{1,S} = w_{2,S} = w_S$ (which implies that $w_\alpha = w_S$), these expressions boil down to (7), (8) and (9).

Replacing $n_1$ and $n_2$ above in (22), (23) and (24), and rearranging we get:

$$\bar{z_S} = w_{1S} \frac{(-1+tN-t(2-tN)N_U)w_{1S} + (1+tN_U(2-tN))w_\alpha - (1+2tN_U)k\alpha w_{1S}}{(1+t(N_U+N))w_{1S} + tN_S w_\alpha}. \tag{34}$$

$$V_{1S}(z_S) = w_{1S} \frac{(1-tN+t(2-tN)N_U)w_{1S} + tN_S((2-tN)w_\alpha - k\alpha w_{1S})}{(1+t(N_U+N))w_{1S} + tN_S w_\alpha} + z_S. \tag{35}$$

$$V_{2S} = w_{1S} \frac{(1+tN(2-tN))w_\alpha - (1+t(N+N_U))k\alpha w_{1S}}{(1+t(N_U+N))w_{1S} + tN_S w_\alpha}. \tag{36}$$

$$\bar{z_U} = tw_U \frac{(-2N_U + N(3-tN_S))w_{1S} - N_S((2-tN)w_\alpha + 2k\alpha w_{1S})}{(1+t(N_U+N))w_{1S} + tN_S w_\alpha}. \tag{37}$$

$$V_{1U}(z_U) = w_U \frac{(1-tN+t(2-tN))w_{1S} + tN_S((2-tN)w_\alpha - k\alpha w_{1S})}{(1+t(N_U+N))w_{1S} + tN_S w_\alpha} + z_U. \tag{38}$$

$$V_{2U} = w_U \frac{(1+tN(2-tN))w_{1S} + ktN_S\alpha w_{1S}}{(1+t(N_U+N))w_{1S} + tN_S w_\alpha}. \tag{39}$$

Setting $w_{1,S} = w_{2,S} = w_S$ (which implies that $w_\alpha = w_S$) these expressions boil down to (13), (14), (16), and (17).

Let us now evaluate the effect of changes in $k$ on the population quantities obtained above. We have

$$\frac{\partial n_1}{\partial k} = \frac{\alpha w_{1,S} N_S}{(1 + t(N_U + N))w_{1S} + tN_S w_\alpha} \tag{40}$$

$$\frac{\partial n_2}{\partial k} = -\frac{\alpha w_{1,S} N_S}{(1 + t(N_U + N))w_{1S} + tN_S w_\alpha} \tag{41}$$

$$\frac{\partial n_{1U}}{\partial k} = -\frac{\alpha w_{1,S} 2tN_U N_S}{(1 + t(N_U + N))w_{1S} + tN_S w_\alpha} \tag{42}$$

$$\frac{\partial n_{2U}}{\partial k} = \frac{\alpha w_{1,S} 2tN_U N_S}{(1 + t(N_U + N))w_{1S} + tN_S w_\alpha} \tag{43}$$

$$\frac{\partial n_{1S}}{\partial k} = \frac{\alpha w_{1,S}(1 + 2tN_U)N_S}{(1 + t(N_U + N))w_{1S} + tN_S w_\alpha} \tag{44}$$

$$\frac{\partial n_{2S}}{\partial k} = -\frac{\alpha w_{1,S}(1 + 2tN_U)N_S}{(1 + t(N_U + N))w_{1S} + tN_S w_\alpha} \tag{45}$$

$$\frac{\partial n_{1U}}{\partial k} = -\frac{\alpha w_{1,S} 2tN_U N_S}{(1 + t(N_U + N))w_{1S} + tN_S w_\alpha} \tag{46}$$

$$\frac{\partial n_{2U}}{\partial k} = \frac{\alpha w_{1,S} 2tN_U N_S}{(1 + t(N_U + N))w_{1S} + tN_S w_\alpha} \tag{47}$$

Setting $w_{1,S} = w_{2,S} = w_S$ (which implies that $w_\alpha = w_S$), these derivatives describe the effect of $k$ on (7), (8) and (9) in the main text.

Finally, consider the effect of $k$ on equilibrium utilities. We have

$$\frac{\partial V_{1,S}}{\partial k} = -\frac{\alpha t w_{1,S}^2 N_S}{(1 + t(N_U + N))w_{1S} + tN_S w_\alpha} \tag{48}$$

$$\frac{\partial V_{2,S}}{\partial k} = -\frac{\alpha t w_{1,S}^2(1 + t(N + N_U))}{(1 + t(N_U + N))w_{1S} + tN_S w_\alpha} \tag{49}$$

$$\frac{\partial V_{1,U}}{\partial k} = -\frac{\alpha t w_{1,S} w_U N_S}{(1 + t(N_U + N))w_{1S} + tN_S w_\alpha} \tag{50}$$

$$\frac{\partial V_{2,U}}{\partial k} = \frac{\alpha t w_{1,S} w_U N_S}{(1 + t(N_U + N))w_{1S} + tN_S w_\alpha} \tag{51}$$

Setting $w_{1,S} = w_{2,S} = w_S$ (which implies that $w_\alpha = w_S$), these expressions boil down to (15) and (18) in the main text.

# C Agglomeration economies

We now propose a version of the model that accounts for economies of agglomeration in the primary city. Specifically, we modify the baseline setting by assuming that the skilled wage in city 1, $w_{1S}$ is an increasing function of the number of skilled individuals that work regularly there. That is, we assume that $w_{1S} = w + \beta n_{1S}$, where $w$ and $\beta$ are positive constants. Skilled individuals living in the secondary city earn a wage $w_{2S} = w$. Assuming the skilled wage a function of the number of individuals in a city makes computations significantly heavier. Hence, we simplify the baseline setting slightly by assuming that all unskilled individuals have the same preference for the primary city. That is, $z_U = z > 0$ for these individuals. For consistency with the above assumptions, we also assume that $z_S$ is distributed uniformly on the $[0, w+\beta N_S]$ interval.

Note that we ignore agglomeration economies from skilled workers that do not work in the primary city regularly (i.e., those that live in the secondary city). We also ignore the presence of agglomeration economies in the secondary city. These assumptions are made for simplicity, but do not drive the main results. Alternative versions of the model that relax these assumptions are available upon request.

Given the above assumptions, we can write the utilities of individuals of different types and conditional on where they live as follows

$$V_{1S}(n_1, z_S) = w_{1S}(1 - tn_1) + z_S, \quad V_{1U}(n_1, z) = w_U(1 - tn_1) + z,$$

$$V_{2S}(n_2) = w_{2S}(1 - tn_2) - \alpha k w_{2S}, \quad V_{2U}(n_2) = w_U(1 - tn_2).$$

In equilibrium, unskilled individuals must attain the same utility level regardless of their city. Given this condition and the identity $n_1 + n_2 = N$, we get

$$n_1 = N/2 + z, \quad n_2 = N/2 - z. \tag{52}$$

Hence, we can characterize the skilled individual indifferent between living in city 1 and 2 as such that

$$\bar{z_S} = w_{2S}[1 - t(N/2 - z)] - \alpha k w_{2S} - w_{1s}[1 - t(N/2 + z)]. \tag{53}$$

Combining the above expression with $w_{1S} = w + \beta n_{1S}$, $w_{2S} = w$, and the fact that $n_{1S} = N_S \cdot Pr[z_S \geq \bar{z_S}]$, we obtain the equilibrium expressions for skilled population living in city 1 and 2:[25]

$$n_{1S} = N_S \frac{2w(1-2t+\alpha k)+\beta}{2w+N_S t(N+2z)} \quad n_{2S} = N_S(1 - \frac{2w(1-2t+\alpha k)+\beta}{2w+N_S t(N+2z)}). \tag{54}$$

These expressions confirm the main predictions of the baseline model (Proposition 1) that the skilled population living in city 1 (resp. 2) increases (decreases) with $k$. Furthermore, since $n_{iU} = n_i - n_{iS}$, for $i = 1, 2$, the opposite effects apply to the unskilled population.

Given the simplified expressions for total population in the two cities, (52), this version of the model does not predict a change in land rents everywhere in the two cities. However, the skilled occupy the location closest to the CBD in both cities (full skill sorting). Hence, as in the baseline model, a reduction in $k$ implies that the area occupied by the skilled in the secondary city expands, and so land rents increase for the plots of land newly occupied by the skilled. That is, the essential prediction of gentrification in the secondary city still holds. We conclude that allowing for agglomeration economies does not change the fundamental predictions of the baseline model.

# D Analysis of the Open City System case

We relax the assumption that the size of the two groups, $N_U$ and $N_S$, are exogenous. For simplicity, we concentrate on the simplified scenario where $w_{1S} = w_{2S} = w_S$ that we consider in the main text.

Assume there is a preliminary stage (stage 0) where two groups of individuals, $j \in S, U$, of exogenous size $M_j$, decide whether to settle in the system formed by cities 1 and 2 or settle elsewhere. Let $N_j$ be the number of individuals from each group that decides to live in the system. The action then unfolds as in the baseline model, where the $N_S$ and $N_U$ individuals decide where to live and work among city 1 and 2.

In stage 0, we assume that individuals in group $j$ get an exogenous utility $V_j^o$ from settling outside the system. Furthermore, we assume the idiosyncratic preference for city 1 versus city 2, $z_j$, is not yet realized at this stage. Hence, individuals expect to receive a utility $E[V_j]+m$, where $m$ is an idiosyncratic preference parameter uniformly distributed on the $[0, w_j]$ interval, and $E[V_j]$ is the expected utility at stage 1, defined

[25] To ensure positive quantities of population in equilibrium in both cities and for both groups, we impose the following conditions on the parameters: $t(N/2+z) < 1$, $\alpha k < 2tz$ and $w > \beta N_S \frac{1-tN/2-tz}{2tz-\alpha k}$. In words, these conditions require that commuting costs be not too large and that agglomeration effects be not too strong, for otherwise no skilled individual would live in city 2.

as

$$E[V_j] = \int_0^{\bar{z}_j} \frac{V_{2j}(n_2)}{w_j} dz_j + \int_{\bar{z}_j}^{w_j} \frac{V_{1j}(n_1, z_j)}{w_j}) dz_j, \tag{55}$$

where

$$V_{1,S}(n_1, z_S) = w_S(1 - tn_1) + z_S, \quad V_{1,U}(n_1, z_U) = w_U(1 - tn_1) + z_U,$$

$$V_{2,S}(n_2) = w_S(1 - tn_2) - \alpha k w_S, \quad V_{2,U}(n_2) = w_U(1 - tn_2),$$

Recall that in equilibrium these utilities do not vary with the individual's distance from the CBD, $x$, so there is no loss in focusing on utilities calculated at the city boundaries.

Proceeding by backward induction, given $N_S$ and $N_U$, Stage 1 yields the same equilibrium as in the baseline model. Therefore, the populations in city 1 and 2 are as characterized in (7), (8) and (9). Furthermore, equilibrium utilities are as characterized in (13) and (16). We can also characterize the indifferent individuals as in (14) and (17).

Consider now Stage 0. Anticipating the equilibrium utility levels at the next stage, an individual in group $j$ chooses to settle in the system of city 1 and 2 if and only if $E[V_j] + m \geq V_j^o$. Given our assumptions, the equilibrium pair $(N_S, N_U)$ satisfies the following system of equations[26]

$$N_j - M_j \frac{w_j - V_j^o + E[V_j]}{w_j} = 0, \quad j = S, U. \tag{56}$$

The left hand side of each equation is a function of $(N_S, N_U)$, since the equilibrium utilities $V_{i,j}$ at stage 1 depend on such quantities (see (13) and (16)). Given $N_j \in [0, M_j]$ and that these functions are continuous and map into the same interval, the above system has at least one solution (fixed point).

Let us now characterize how changes in $k$ affect the equilibrium group sizes, $N_S$ and $N_U$. To do so, we start from the system in (56) and apply the implicit function theorem to get

$$\frac{\partial N_S}{\partial k} = -\frac{det \begin{bmatrix} -\frac{N_S}{w_S}\frac{\partial E[V_S]}{\partial k} & -\frac{N_S}{w_S}\frac{\partial E[V_S]}{\partial N_U} \\ -\frac{N_U}{w_U}\frac{\partial E[V_U]}{\partial k} & 1 - \frac{N_U}{w_U}\frac{\partial E[V_U]}{\partial N_U} \end{bmatrix}}{det \begin{bmatrix} 1 - \frac{N_S}{w_S}\frac{\partial E[V_S]}{\partial N_S} & -\frac{N_S}{w_S}\frac{\partial E[V_S]}{\partial N_U} \\ -\frac{N_U}{w_U}\frac{\partial E[V_U]}{\partial N_S} & 1 - \frac{N_U}{w_U}\frac{\partial E[V_U]}{\partial N_U} \end{bmatrix}}, \quad \frac{\partial N_U}{\partial k} = -\frac{det \begin{bmatrix} 1 - \frac{N_S}{w_S}\frac{\partial E[V_S]}{\partial N_S} & -\frac{N_S}{w_S}\frac{\partial E[V_S]}{\partial k} \\ -\frac{N_U}{w_U}\frac{\partial E[V_U]}{\partial N_S} & -\frac{N_U}{w_U}\frac{\partial E[V_U]}{\partial k} \end{bmatrix}}{det \begin{bmatrix} 1 - \frac{N_S}{w_S}\frac{\partial E[V_S]}{\partial N_S} & -\frac{N_S}{w_S}\frac{\partial E[V_S]}{\partial N_U} \\ -\frac{N_U}{w_U}\frac{\partial E[V_U]}{\partial N_S} & 1 - \frac{N_U}{w_U}\frac{\partial E[V_U]}{\partial N_U} \end{bmatrix}}. \tag{57}$$

[26]For consistency, we assume that $0 < V_j^o < w_j + E[V_j]$.

To determine the sign of $\frac{\partial N_S}{\partial k}$ and $\frac{\partial N_U}{\partial k}$, we need to study the signs of the derivatives $\frac{\partial E[V_j]}{\partial k}$, $\frac{\partial E[V_j]}{\partial N_S}$ and $\frac{\partial E[V_j]}{\partial N_U}$. Differentiating (55) and using the fact that, by definition, $\bar{z}_j = V_{2j}(n_2) - w_j(1 - tn_1)$, we have

$$\frac{\partial E[V_j]}{\partial k} = \frac{\partial V_{2j}}{\partial k}\frac{\bar{z}_j}{w_j} + (1 - \frac{\bar{z}_j}{w_j})(-w_j t\frac{\partial n_1}{\partial k}), \quad j = S, U. \tag{58}$$

The derivatives in (15) and (18) indicate that $\frac{\partial V_{2S}}{\partial k} < 0$ and $\frac{\partial V_{2U}}{\partial k} > 0$. Furthermore, $\frac{\partial V_{1S}}{\partial k} = -w_S t\frac{\partial n_1}{\partial k} < 0$ and $\frac{\partial V_{1U}}{\partial k} = -w_U t\frac{\partial n_1}{\partial k} < 0$. Hence, we have

$$\frac{\partial E[V_S]}{\partial k} < 0, \quad \frac{\partial E[V_U]}{\partial k} \gtreqless 0. \tag{59}$$

An increase in $k$ is costly to the skilled individuals who live in the primary city because $\frac{\partial n_1}{\partial k} > 0$, which implies that their housing costs increase. The skilled who live in city 2 suffer too because, although their housing expenses decrease ($\frac{\partial n_2}{\partial k} < 0$), their cost of commuting increases, and the net effect is negative. The effect of $k$ on the expected utility of the unskilled is instead ambiguous, because these individuals benefit from an increase in $k$ if they end up living in city 2 (since $n_2$ decreases, and so does the land rent there), but, by the same token, they suffer if they live in city 1.

Consider now the effect of group sizes, $(N_S, N_U)$, on individual utility. We have

$$\frac{\partial E[V_j]}{\partial N_S} = \frac{\partial V_{2j}}{\partial N_S}\frac{\bar{z}_j}{w_j} + (1 - \frac{\bar{z}_j}{w_j})(-w_j t\frac{\partial n_1}{\partial N_S}), \quad j = S, U. \tag{60}$$

$$\frac{\partial E[V_j]}{\partial N_U} = \frac{\partial V_{2j}}{\partial N_U}\frac{\bar{z}_j}{w_j} + (1 - \frac{\bar{z}_j}{w_j})(-w_j t\frac{\partial n_1}{\partial N_U}), \quad j = S, U. \tag{61}$$

Note that, in the above expressions, $\frac{\partial V_{2j}}{\partial N_U} = -w_j t\frac{\partial n_2}{\partial N_U}$ and $\frac{\partial V_{2j}}{\partial N_S} = -w_j t\frac{\partial n_2}{\partial N_S}$. It can be shown that $\frac{\partial n_i}{\partial N_U} > 0$ and $\frac{\partial n_i}{\partial N_S} > 0$ hold, $\forall i$ (derivations available upon request). In words, an expansion in the total size of either population group expands the size of both cities. Hence, we have

$$\frac{\partial E[V_j]}{\partial N_U} < 0, \quad \frac{\partial E[V_j]}{\partial N_S} < 0, \quad j = S, U. \tag{62}$$

Given (59) and (62), the determinants of the matrices in expression (57) is generally ambiguous. To get some intuition, consider the effect of $k$ on $N_S$. An increase in the cost of long-distance travel, all else given, tends to reduce the utility of skilled individuals, as discussed above. This effect should induce fewer of the skilled to settle in the two-city system. However, the overall effect of $k$ on $N_S$ also depends on how $k$ affects the utility of unskilled individuals, and we have seen that this effect is ambiguous. Furthermore, then net change in $N_S$ also depends on how $N_U$ changes, and this change can also be either positive or negative. Therefore, changes in the cost of long-distance travel can either increase or decrease the total population living in the two-city system.

In light of the above findings, let us consider the overall effect of $k$ on the size of each city and on equilibrium utilities in an open system. We have

$$\frac{dn_i}{dk} = \frac{\partial n_i}{\partial k} + \frac{\partial n_i}{\partial N_S}\frac{\partial N_S}{\partial k} + \frac{\partial n_i}{\partial N_U}\frac{\partial N_U}{\partial k} \quad i = 1, 2. \tag{63}$$

The first term on the right hand side is positive if and only if $i = 2$ (this is immediately seen from (7) reported in the text). However, the other two terms are ambiguous as discussed above, and so is the overall effect. A similar conclusion applies when considering the effect of $k$ on the size of skilled and unskilled groups, $n_{ij}$, in each city. It also follows that the effect of $k$ on land rents in the two cities (which increase with $n_i$) is ambiguous.

Finally, we study the effect of $k$ on individual utility, as expressed in (13) and (13). We have

$$\frac{dV_{ij}}{dk} = \frac{\partial V_{ij}}{\partial k} + \frac{\partial V_{ij}}{\partial N_S}\frac{\partial N_S}{\partial k} + \frac{\partial V_{ij}}{\partial N_U}\frac{\partial N_U}{\partial k} \quad i = 1, 2, \quad j = S, U. \tag{64}$$

The first derivative on the right hand side is as presented in expressions (15) and (18) in the main text. The other terms are ambiguous in sign, however, because the effect of $k$ on $N_S$ and $N_U$ is ambiguous. Thus, again, we cannot make a definitive statement about the effect of $k$ on individual utility when taking into account the adjustment in the overall size of the population in the system.

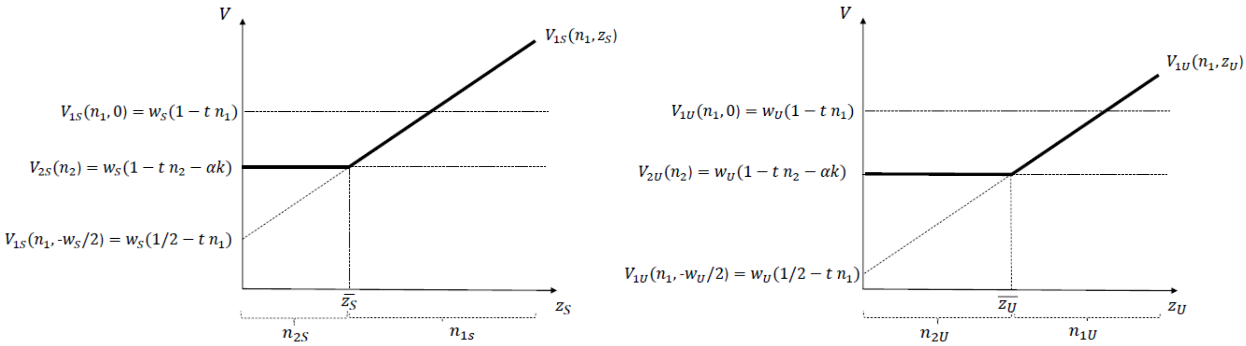
# E Allowing for negative values of $z_U$ and $z_S$

In this Appendix we provide a modified version of the baseline model where we allow for negative preferences for living in the primary city. Specifically, we assume that each individual in group $j$ gets a marginal utility $z_j$ from living in city 1, with $z_j \sim U[-\frac{w_j}{2}, \frac{w_j}{2}], j = S, U$. We make no other changes with respect to the baseline model. For concreteness, and focus on the setting where $w_{1S} = w_{2S} = w_S$ for concreteness.

As in the baseline model (following the same steps as in Appendix B), the allocation of population in the cities can be obtained by using the equilibrium conditions requiring that (i) individuals of each group be indifferent as to their location within a city, and (ii) that individuals live in the primary city if and only if their idiosyncratic preference for such city is above a threshold, $\bar{z}_j$. The only difference lies in the

Figure A5: Allocation of skilled and unskilled individuals among cities

(a) Skilled (b) Unskilled




distribution of $z_j$. We obtain that[27]

$$n_1 = N/2 + \frac{\alpha k N_S}{1+2tN}, \quad n_2 = N/2 - \frac{\alpha k N_S}{1+2tN}, \tag{65}$$

$$n_{1S} = N_S \frac{1/2 + tN + \alpha k(1+2tN_U)}{1+2tN}, \quad n_{2S} = N_S \frac{1/2 + tN - \alpha k(1+2tN_U)}{1+2tN}, \tag{66}$$

$$n_{1U} = N_U \frac{1/2 + tN - \alpha k N_S}{1+2tN}, \quad n_{2U} = N_U \frac{1/2 + tN + \alpha k N_S}{1+2tN}. \tag{67}$$

It can be verified that, given the above expressions, individuals of a given type earn a higher income (net of commuting/housing costs) in the primary city than in the secondary. See Figure A5.

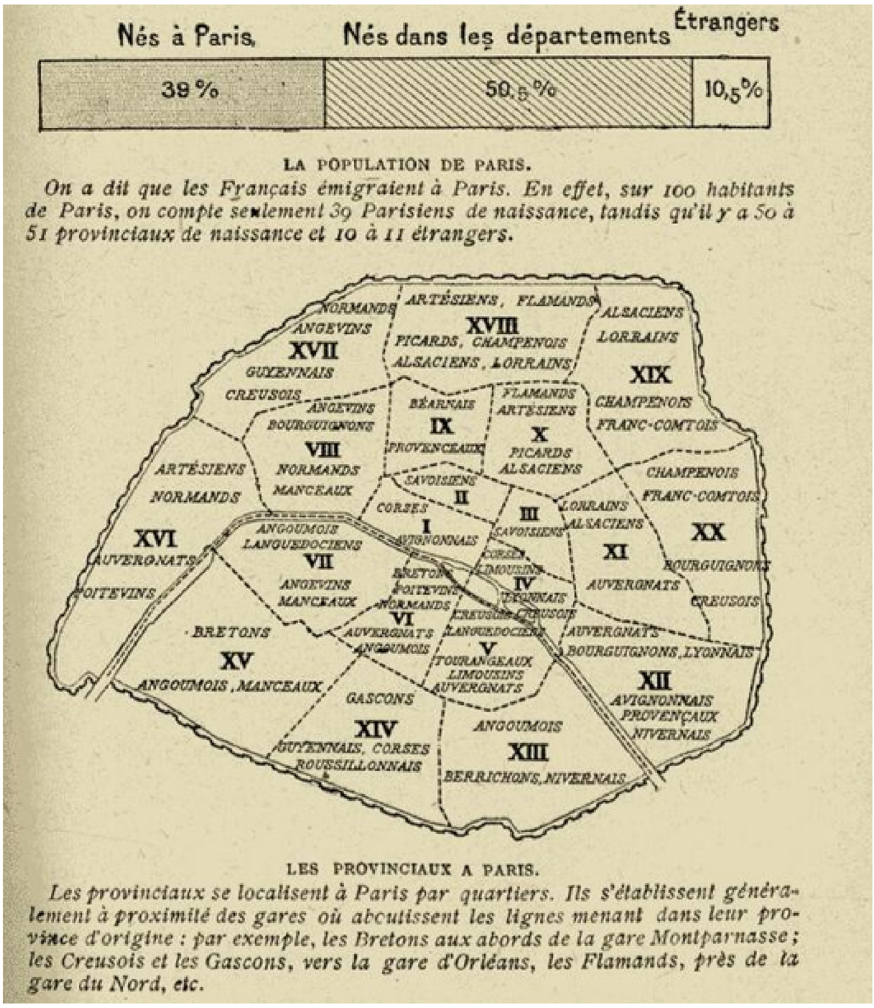
These expressions show that, as in the baseline model, a reduction in $k$ brings to an increase in population in city 2 overall, and also an increase in the skilled population in such city, while city 1 loses population overall and loses skilled individuals, with the opposite effects applying to the unskilled individuals. The remainder of the analysis unfolds exactly as in the baseline model and is therefore not repeated here.

[27]The condition $\alpha k < min(\frac{N(1+2tN)}{2N_S}, \frac{1+2tN}{1+2tN_U})$ is necessary and sufficient for all population quantities to be positive. We assume throughout the analysis that this condition holds. We also assume that $tN < 1$, which is sufficient for all individuals to achieve a positive level of utility in equilibrium. Both conditions require that commuting costs be not exceedingly large.

# F Supporting material

## F.1 Figures

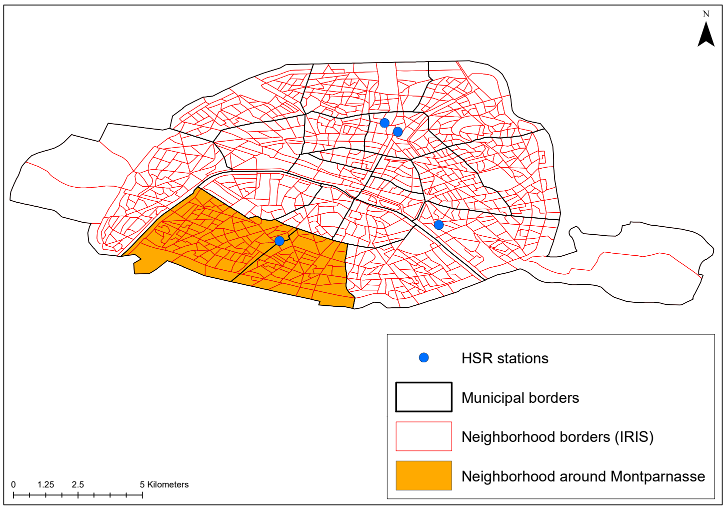
Figure A6: Location of provincials in Paris (1914)



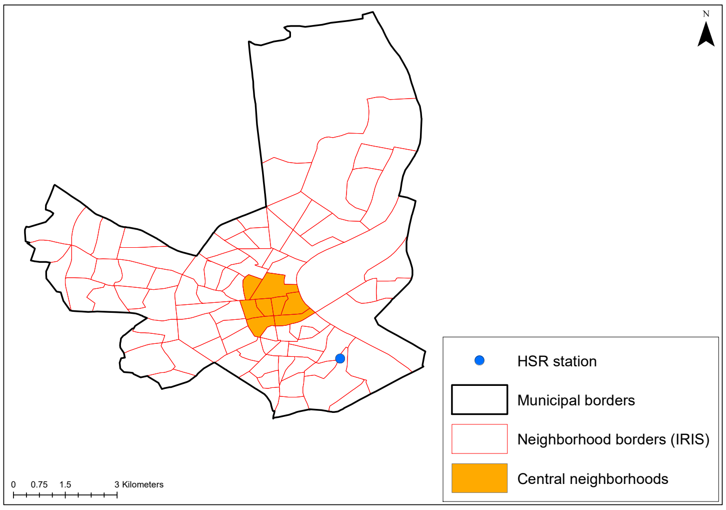

*Notes:* Bottom text can be translated to: *"The provincials are located in Paris by district. They generally settle near the stations where the lines leading to their province of origin end: the Bretons near the Montparnasse station, ...". Source:* Gallouédec (1914).

Figure A7: Urban and administrative structure of Paris, Bordeaux and Rennes, including location of HSR stations

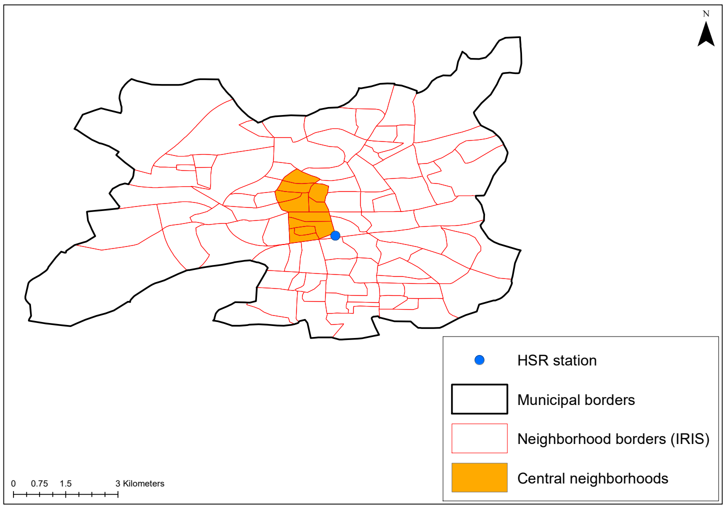
(a) Paris




(b) Bordeaux




(c) Rennes

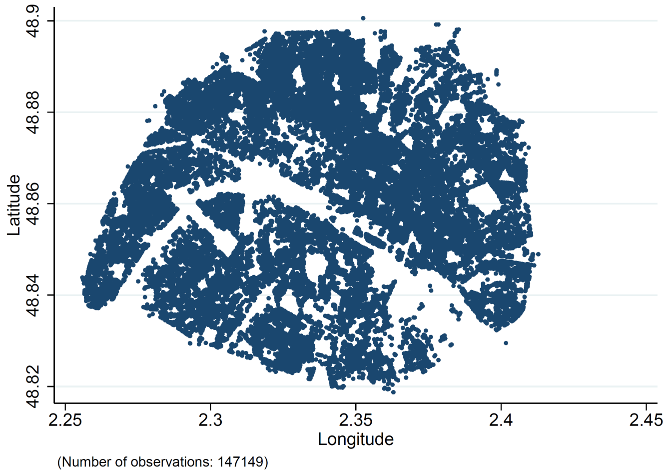




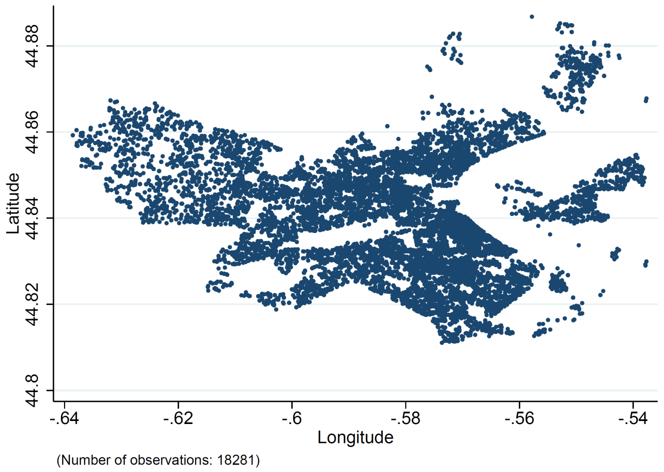

*Notes:* Authors' own illustration based on shapefiles from https://www.data.gouv.fr/en/datasets.

Figure A8: Location of housing unit transactions (2016- 2020)

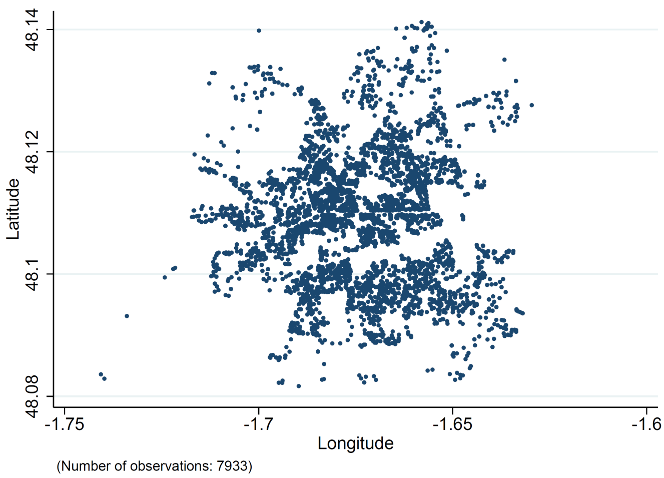
(a) Paris




(b) Bordeaux

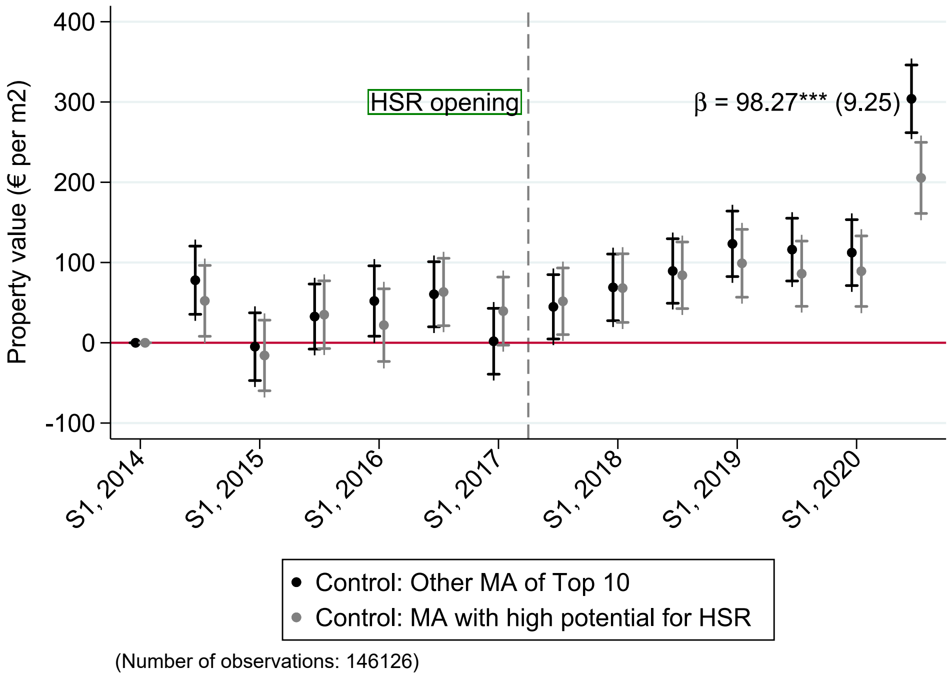



(b) Rennes




*Notes:* Official housing price data on the universe of housing property sold in France between Jan 1st, 2016 and Dec 31st, 2020. This data-set includes the exact coordinates of each property. It is produced and made publicly available by https://app.dvf.etalab.gouv.fr/.

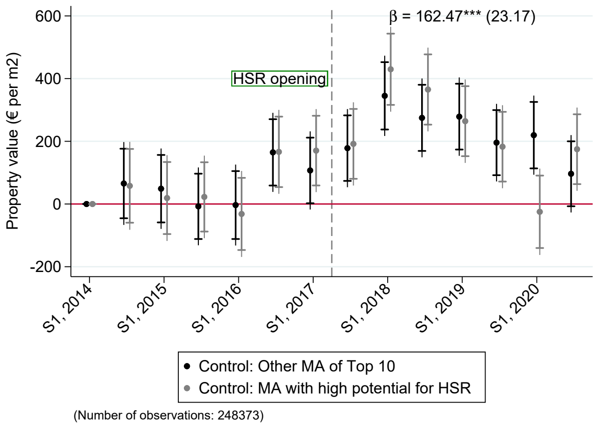
Figure A9: Average HSR effect on housing prices in incidentally treated secondary cities (€ per m²)



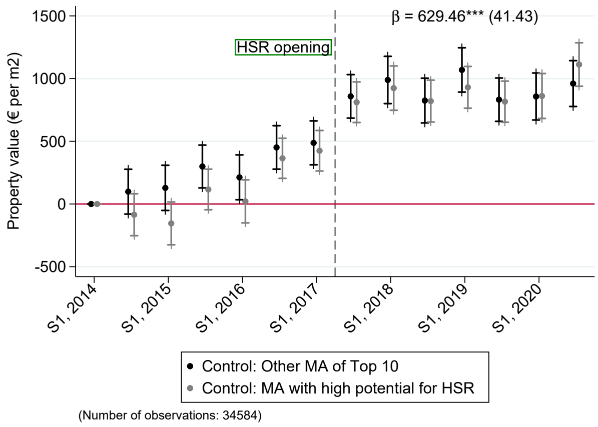

*Notes:* ES-DiD model (19) using 95% CIs.

Figure A10: High-speed rail and property value of houses and appartments

(a) Property value of appartments (€ per m$^2$)

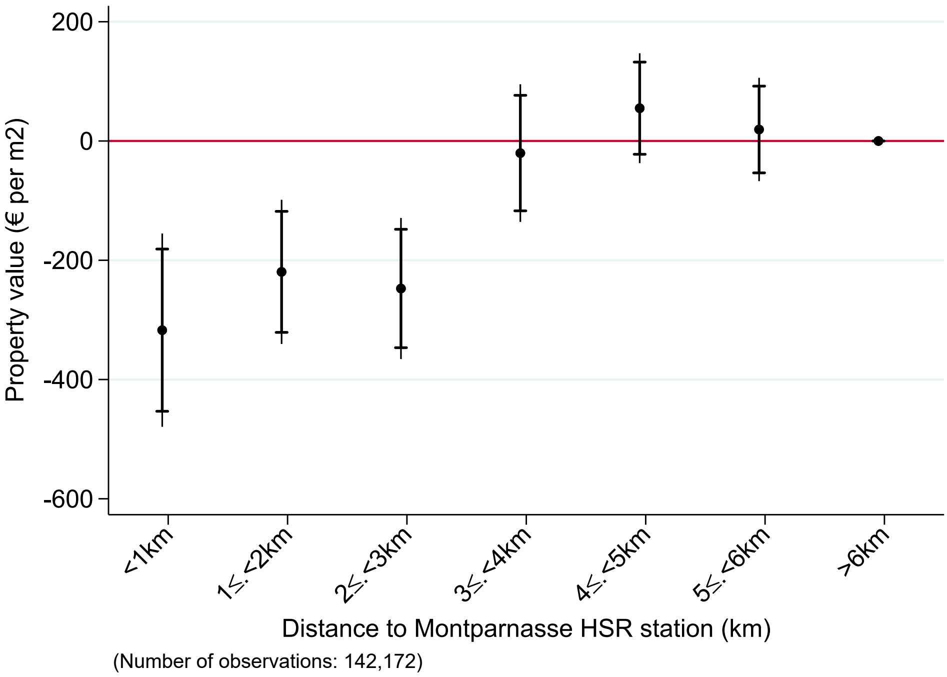



(b) Property value of houses (€ per m$^2$)



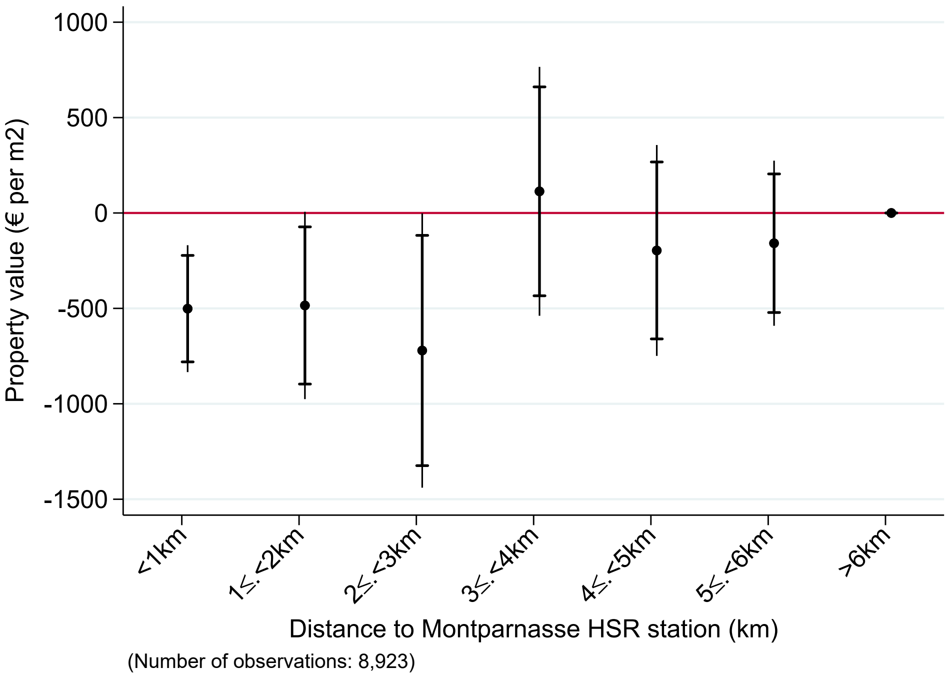

*Notes:* ES-DiD model (19) using 95% CIs. Property values (on a semester basis) in Bordeaux and Rennes are compared to property values in Marseille, Lyon, Lille, Nice and Strasbourg (i.e., all other cities of the top 10 French largest cities excluding Nantes, Paris and Toulouse for incidental treatment reasons).

Figure A11: Average HSR effect on housing prices as a function of distance to Gare Montparnasse (Paris' train station to Rennes and Bordeaux, € per m²)



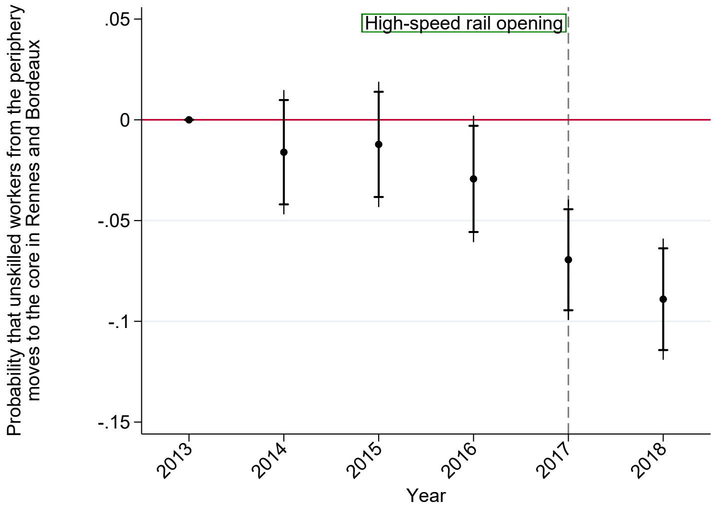

*Notes:* ES-DiD model (19) using 90% and 95% CIs.

Figure A12: AVERAGE HSR EFFECT ON HOUSING PRICES CLOSE TO METRO STATIONS AS A FUNCTION OF DISTANCE TO GARE MONTPARNASSE
(PARIS' TRAIN STATION TO RENNES AND BORDEAUX, € PER M²)




*Notes:* ES-DiD model (19) using 90% and 95% CIs.

Figure A13: Periphery residents moving to core in Bordeaux and Rennes




*Notes:* ES-DiD model (21) using 95% CIs. Figure studies the probability that locals from Bordeaux and Rennes' greater peripheries move to the municipalities of Bordeaux and Rennes. Greater periphery of Bordeaux and Rennes is defined as their respective departments.

## F.2 Tables

Table A1: Descriptive statistics on housing unit transactions (2016-2020)

| | Variable | N | Mean | SD |
|---|---|---|---|---|
| **Paris** | Price (€/$m^2$) | 147,149 | 10,021.98 | 2775.90 |
| | Size ($m^2$) | 147,149 | 53.49 | 39.99 |
| | Number of rooms | 147,149 | 2.41 | 1.29 |
| | Share of apartments (%) | 147,149 | 100 | - |
| | Month of transaction | 147,149 | 6.91 | 3.40 |
| **Bordeaux** | Price (€/$m^2$) | 18,281 | 4,573.67 | 993.82 |
| | Size ($m^2$) | 18,281 | 71.24 | 46.58 |
| | Number of rooms | 18,281 | 2.89 | 1.54 |
| | Share of apartments (%) | 18,281 | 74 | - |
| | Month of transaction | 18,281 | 7.06 | 3.38 |
| **Rennes** | Price (€/$m^2$) | 7,933 | 3,994.12 | 835.51 |
| | Size ($m^2$) | 7,933 | 63.71 | 39.83 |
| | Number of rooms | 7,933 | 2.92 | 1.66 |
| | Share of apartments (%) | 7,933 | 77 | - |
| | Month of transaction | 7,933 | 7.19 | 3.31 |

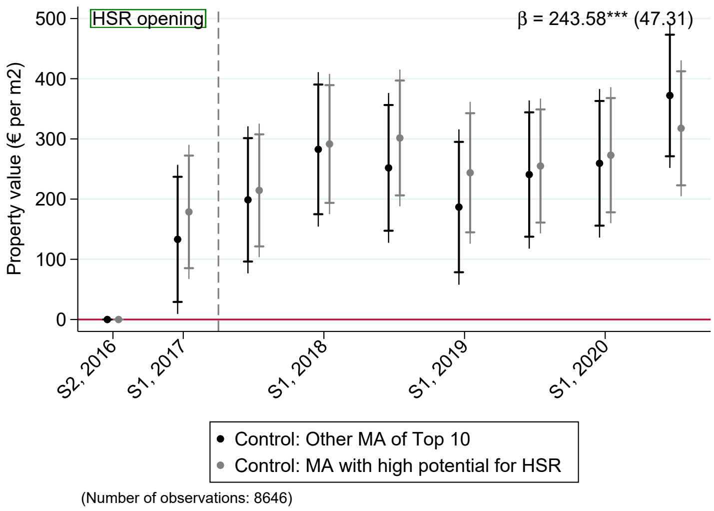
*Notes:* Official housing price data on the universe of housing property sold in France between Jan 1st, 2016 and Dec 31st, 2020. This data-set is produced and made publicly available by https://app.dvf.etalab.gouv.fr/.

# G Robustness checks

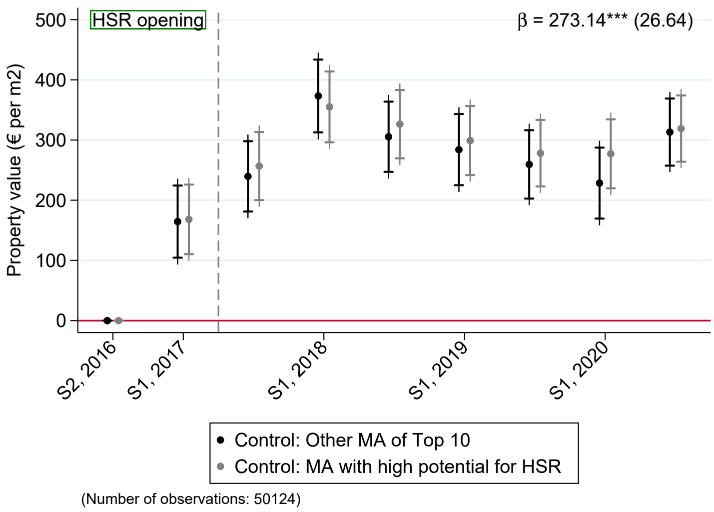
## G.1 Grid cells as units of observation

Figure A14: High-speed rail and property values at the *grid cell level* in Bordeaux and Rennes

(a) Property value with 100m$^2$ cells (€ per m$^2$)




(b) Property value with 10m$^2$ cells (€ per m$^2$)

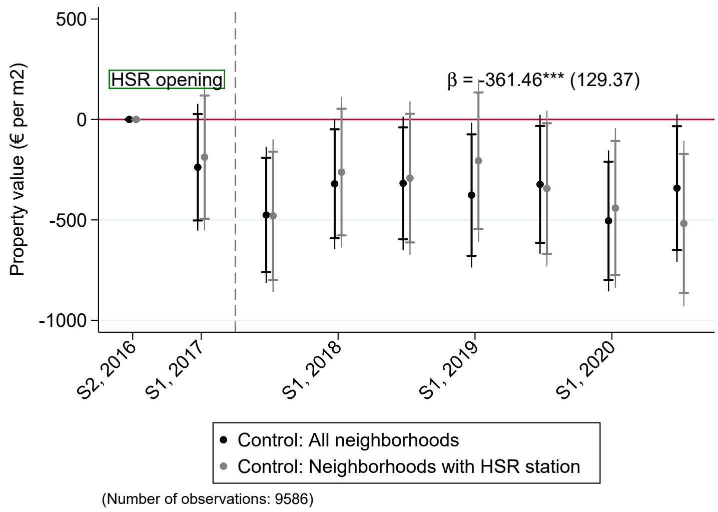



*Notes:* DiD model using 95% CIs. Property values (on a quarterly basis) in Bordeaux and Rennes are compared to property values in Marseille, Lyon, Lille, Nice and Strasbourg (i.e., all other cities of the top 10 French largest cities excluding Nantes, Paris and Toulouse for incidental treatment reasons).

Figure A15: HIGH-SPEED RAIL AND PROPERTY VALUES AT THE *GRID CELL LEVEL* IN PARIS

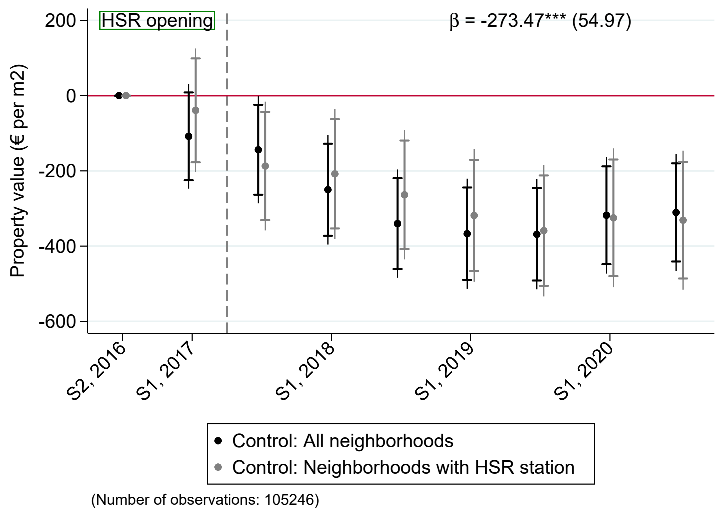
(a) Property value with $100m^2$ cells

(€ per $m^2$)




(b) Property value with $10m^2$ cells

(€ per $m^2$)




*Notes:* DiD model using 95% CIs. Property values (on a quarterly basis) in Bordeaux and Rennes are compared to property values in Marseille, Lyon, Lille, Nice and Strasbourg (i.e., all other cities of the top 10 French largest cities excluding Nantes, Paris and Toulouse for incidental treatment reasons).