

This is the **accepted version** of the article:

Garcia-López, Miquel-Àngel; Hémet, Camille; Viladecans-Marsal, Elisabet.
«How does transportation shape intrametropolitan growth? An answer from
the Regional Express Rail». The Journal of Regional Science, Vol. 57 Núm. 5
(novembre 2017), p. 758-780. DOI 10.1111/jors.12338

This version is available at <https://ddd.uab.cat/record/220864>

under the terms of the  **IN**
COPYRIGHT license

HOW DOES TRANSPORTATION SHAPE INTRAMETROPOLITAN GROWTH? AN ANSWER FROM THE REGIONAL EXPRESS RAIL

Miquel-Àngel Garcia-López^{*†}

Universitat Autònoma de Barcelona and Institut d'Economia de Barcelona

Camille Hémet^{*‡}

Paris School of Economics (ENS), Institut d'Economia de Barcelona and CEPR

Elisabet Viladecans-Marsal^{*§}

Universitat de Barcelona, Institut d'Economia de Barcelona and CEPR

January 2017

ABSTRACT: This paper analyzes the influence of transportation infrastructure, and in particular of the Regional Express Rail (RER), on employment and population growth in the Paris metropolitan area between 1968 and 2010. In order to make proper causal inference, we rely on historical instruments and control for all other transportation modes that could be complement or substitute to the RER. Our results show that proximity to a RER station increases employment and population density and, in particular, employment and population growth. The latter effects are higher in municipalities located near RER stations and close to employment (sub)centers. They are also found to be particularly strong for jobs in the service sector, for factory workers and for highly educated population. We find no impact of the RER expansion on employment growth during the first part of the period, while the effect on population growth appears earlier but declines over time.

Key words: urban growth, urban spatial structure, transportation

JEL classification: R11, R12, R42, L91

^{*}We are very grateful to Gilles Duranton, Laurent Gobillon, Miren Lafourcade, Daniel McMillen and Jean-Claude Prager, as well as seminar and conference participants and two anonymous referees for their comments and suggestions. We also want to thank Corentin Trévien and Thierry Mayer for their very helpful GIS maps and data on the transportation networks. Financial support from the Société du Grand Paris, the Ministerio de Ciencia e Innovación (research projects ECO2010-20718 and ECO2014-52999-R (M.A. Garcia-López), and ECO2010-16934 and ECO2013-41310-R (E. Viladecans-Marsal)), Generalitat de Catalunya (research projects 2014SGR1326 (M.A. Garcia-López) and 2014SGR420 (E. Viladecans-Marsal)), and the "Xarxa de Referència d'R+D+I en Economia Aplicada" is gratefully acknowledged.

[†]Department of Applied Economics, Universitat Autònoma de Barcelona, Edifici B, Facultat d'Economia i Empresa, 08193 Cerdanyola del Vallès, Spain (miquelangel.garcia@uab.cat).

[‡]Paris School of Economics, 48 Boulevard Jourdan, 75014 Paris, France (camille.hemet@psemail.eu).

[§]Department Economics, Universitat de Barcelona, John M Keynes 1-11, 08034 Barcelona, Spain (eviladecans@ub.edu).

HOW DOES TRANSPORTATION SHAPE INTRAMETROPOLITAN GROWTH?

AN ANSWER FROM THE REGIONAL EXPRESS RAIL

1 Introduction

During recent decades the city of Paris and its metropolitan area have undergone major demographic and socioeconomic changes. Specifically, between 1968 and 2010, the Paris metropolitan area strengthened its position both as the most populated area (around 12 million inhabitants) and as the largest economic region in the country (with one fifth of total employment). However, this growth has not been homogeneously distributed throughout the metropolitan area. In fact, even though the levels of employment and population of the whole area respectively grew by 32 percent and 27 percent over this period, the central business district (CBD) has seen employment fall by 7 percent and its population by 13 percent. As a result of these changes, the share of total metropolitan area employment located in the CBD fell from 45 percent in 1968 to 32 percent in 2010. The same pattern is observed for the population of the CBD, which today represents around 19 percent of the whole area, compared to 28 percent at the end of the 1960s. All in all, these trends indicate that the Paris metropolitan area has undergone a marked process of suburbanization, accompanied by the emergence of employment subcenters (areas of high employment density outside the CBD).

There is a long tradition in the literature of studies seeking to explain the determinants of city structure and city growth. An important amount of research has been devoted to understanding why some cities are more successful and grow more rapidly than others, and how the urban structure changes, leading to a variety of responses: while some authors emphasize the importance of human capital and skills ([Moretti, 2004](#); [Rosenthal and Strange, 2008](#), are good examples), others focus on the role of the weather ([Glaeser et al., 2001](#)) or the availability of consumer amenities ([Carlino and Saiz, 2008](#)) as attractors for population. Ultimately, however, the literature agrees that a city's density largely ac-

counts for its capacity to be productive and to attract better firms and workers (Combes et al., 2012). For urban economists, agglomeration economies are therefore considered an important source of city growth. What is also clear is that the advantages provided by agglomeration economies increase with a reduction in the transportation costs for goods and people (Glaeser and Gottlieb, 2009). Yet, even though transportation seems to be a key element for employment and population distribution and growth, its impact on urban growth has long been ignored in the empirical literature (see the pioneering work by Duranton and Turner, 2012).

The main objective of this paper is thus to contribute to this relatively recent branch of the literature by analyzing the spatial influence on employment and population growth of a major improvement in transportation infrastructure. More precisely, our study relies on the case of the Paris metropolitan area, where the transportation network has been substantially improved with the introduction of a new suburban train, the *Réseau Express Régional* (RER henceforth). Since its inauguration in 1975, the total length of the RER network has increased by around 550 km, and now operates 257 stations connecting more than 170 municipalities. Yet, this is not the only change in the transportation infrastructure that the region has undergone over the past forty years. Likewise, the metro and tramway networks have also been expanded throughout the area. Most notably, the area's main road system (highways) has been extended by 600 km, while the number of ramps and accessibility to many other municipalities have been increased. Thus, although we focus our attention on the expansion of the RER (1975–2010) and its effects on the location of new jobs and inhabitants, we need to take into account all the other transportation modes that might complement or substitute the RER system. We therefore make considerable efforts to control for all the changes in the area's other transportation modes, in order to estimate unbiased estimates of the RER effect. We first analyze the changes in Paris metropolitan area's urban spatial structure between 1968 and 2010 and the location role of the CBD and employment subcenters. We then turn to analyze the spatial influence of transportation

infrastructure on the 2010 intrametropolitan distribution of employment and population. We finally estimate whether transportation fostered employment and population growth during this period.

Doing so, this paper makes three important contributions to the literature. First, we analyze the impact of RER improvements on employment and population growth for each of the 1,300 municipalities in the Paris metropolitan area, which has witnessed an important improvement in its transportation system in recent decades. Second, to the best of our knowledge, this is the first work to undertake an analysis of the causal effects of improvements to an infrastructure system on city growth controlling for all the possible modes of transportation (railroads, metro, tramways and highways). Third, our empirical strategy allows us to solve the causality problem that is common in this type of approach through the use of historical instruments.

Our work builds on seminal theoretical contributions which help frame the way transportation infrastructures should affect population and employment growth: the classical monocentric city model developed by [Alonso \(1964\)](#), [Mills \(1967\)](#) and [Muth \(1969\)](#) shows that transportation (accessibility) is the main factor that determines urban land use ([Duranton and Puga, 2015](#)). [Anas and Moses \(1979\)](#) and [Baum-Snow \(2007a\)](#) extend this model by considering two competing transportation infrastructures: a classical transportation infrastructure based on a dense network of radial streets and a high speed transportation infrastructure based on sparse radial corridors ([Anas and Moses, 1979](#); [Baum-Snow, 2007a](#)). Depending on the cost of alternative transportation modes, the authors find that population and employment spread out along the sparse corridors, increasing surrounding land rents and densities.

This paper can thus be viewed as an empirical test of these theoretical predictions regarding the role of transportation infrastructures. While we focus on the effect on the spatial distribution and growth of population and employment across municipalities of a large metropolitan area, other empirical studies analyze the effects of railroads on a wide

range of alternative outcomes. A great deal of papers study for instance the impact of new railroads and stations on the housing market and the property values of the affected area. [Baum-Snow and Kahn \(2000\)](#), [Baum-Snow and Kahn \(2005\)](#), [Bowes and Ihlanfeldt \(2001\)](#), [Gibbons and Machin \(2005\)](#) and [Ryan \(2005\)](#) are good examples, but are mostly applied to U.S. cities. More recent studies, such as [Diao et al. \(2016\)](#) and [Li et al. \(2016\)](#), focus instead on Chinese cities and analyze the impact of the high speed train on the housing markets. On the same topic, other recent approaches explore the incidence of railroads on land use ([Hurst and West, 2014](#); [Schuetz, 2015](#)). Closer to our paper, [Chatman and Noland \(2014\)](#) study the impact on employment density and productivity in U.S. central cities, and [Kotavaara et al. \(2011\)](#) focus on population growth in Finish municipalities, but do not take stand on the spatial structure of these cities. Some recent research departs from these traditional topics in urban economics, and has started looking into the effect of railroad on largely unexplored aspects, such as air pollution ([Chen and Whalley, 2012](#)) or cities' gentrification ([Grube-Cavers and Patterson, 2015](#)). The evidence of these analyses is quite diverse, but a majority of papers point to a positive or neutral effect of these investments on the different analyzed outcomes. Our paper distinguishes itself from this large literature by analyzing the incidence of railroads investments on the internal spatial structure of cities, which remains largely unexplored, to the best of our knowledge.

Our study is also related to recent empirical studies that have examined other aspects of transportation infrastructure. Sharing our intrametropolitan approach, some papers focus on the effect of rail or road transportation on the suburbanization process, both in the U.S. ([Baum-Snow, 2007b](#)) and in Europe ([Garcia-López et al., 2015a,b](#)), while other studies, conducted at the county level, consider alternative outcomes, such as workers' earnings ([Michaels, 2008](#)) or employment growth ([Jiwattanakulpaisarn et al., 2009](#)). [Durananton and Turner \(2011\)](#) and [Hsu and Zhang \(2014\)](#) provide intermetropolitan evidence for the effect of highway improvements on congestion in the U.S. and Japan, respectively. In the development economic literature as well, some recent papers analyze the effect of

infrastructures on various economic outcomes, in China ([Banerjee et al., 2012](#); [Faber, 2014](#)) or India ([Donaldson, 2015](#)).

Analyzing the impact of infrastructure improvements on city growth typically involves two problems of inference. First, all types of infrastructure take time to be built and their effects on city growth are not immediate. This problem can be solved by using long differences for both employment or population and infrastructure changes. Second, the location of new infrastructures is not random. Rather, it is most likely endogenous to employment and population growth: planners may decide to improve the connection of deprived areas in order to boost their economic activity or attract population, or they may on the contrary connect areas that are expected to grow, anticipating on future needs for transportation. It means that any naive estimate of the effect of the distance to a transportation infrastructure on population or employment growth will be biased, which is one of the main issues in the literature. Only recently have a few papers proposed various inference strategies to address this problem. A first category of solution consists in using exogenous deviations from initial transportation plans (see for instance [Baum-Snow, 2007b](#); [Michaels, 2008](#)). Special mention needs to be made to [Mayer and Trévién \(2015\)](#) who also focus on the Paris region. They exploit the deviation from the initial investment plan for the RER network resulting from budgetary and technical constraints to evaluate the impact of the opening and of the progressive extension of the RER between 1975 and 1990 on employment and population at the municipal level. An alternative identification strategy has been proposed by [Jiwattanakulpaisarn et al. \(2009\)](#), who use the lagged levels of highway lane-mile density as instruments for highway infrastructure investments. Finally, the use of historical instruments is another popular strategy to tackle this causality issue. For instance, [Duranton and Turner \(2012\)](#) instrument road infrastructure using the U.S. railway network at the end of the 19th century and the routes taken by major expeditions of the United States between 1518 and 1850; [Hsu and Zhang \(2014\)](#) rely on the historical railway network plan of 1890 and the planned national express way exten-

sion as exogenous sources of variation of highway location in Japan; and [Garcia-López et al. \(2015a\)](#) use the Roman roads and the 1760 Postal routes as instruments for Spanish highways. Our empirical strategy follows this approach, as we rely on two historical instruments, the 1870 railways and the Roman roads, as a source of exogenous variation of current infrastructure' location.

Our results show that the RER network influences the location of employment and population, even after controlling for other modes of transportation. Getting 1 kilometer closer to a RER station is found to increase employment and population density by around 5 percent and 4 percent, respectively. Furthermore, a dynamic analysis reveals that improving the RER network significantly increases municipal employment and population growth: for each kilometer closer to a RER station, employment increases by 2 percent and population by 1 percent. Although this impact seems limited, it is considerably reinforced once we introduce heterogeneity in the analysis: (a) For municipalities located less than 1 kilometer from a RER station, each kilometer closer to a RER station increases employment and population growth by 8 percent and 6 percent respectively; (b) For non-central municipalities located less than 5 kilometers from an employment (sub)center, each kilometer closer to a RER station increases employment and population growth by 12 percent and 9 percent respectively; (c) The effects on employment growth are higher for services (3 percent) and factory jobs (3 percent), the effects on population growth are higher for inhabitants with high school (3 percent) and university (2 percent) degree; (d) We find no impact of the RER expansion on employment growth during the first part of the covered period, while the impact on population growth was sizeable much earlier but declined over time.

The remainder of the paper is organized as follows. Section 2 first describes the changes in the urban spatial structure in the Paris metropolitan area and highlights the suburbanization process of the area. It then turns to presenting the main changes in the different transportation infrastructures in the Paris metropolitan area and their influence on the

intrametropolitan location of employment and population. Section 3 presents our main results, and Section 4 concludes.

2 Urban spatial structure and transportation in Paris metropolitan area

2.1 Urban spatial structure and suburbanization process

The *Paris metropolitan area* is one of the 22 administrative regions in continental France, known as *Ile de France*, to which the city of Paris belongs. It is divided into eight *départements* (administrative sub-regions) and 1,300 municipalities. Note that the city of Paris has been a *département* of its own since 1968, and is divided into 20 municipalities called *arrondissements*. The municipality is the unit of analysis of this paper. It is actually the smallest administrative division that we can use, since smaller divisions were not introduced in the French census before 1999. This is however a reasonable unit of analysis given our research agenda as French municipalities are particularly small:¹ in *Ile de France*, the average municipal surface is 9.3 km², and the median is 7.6 km². The metropolitan area of Paris is the densest and most populated region in France, with 986.7 inhabitants per square kilometer in 2011 for a total of 11,852,851 inhabitants. It is also the main employment center in the country: with a total of 5,660,253 jobs in 2011, it accounted for more than a fifth of total employment in continental France. Among them, 0.2 percent work in the agricultural sector, 5.2 percent in the construction sector, 8.4 percent in industry and the remaining 86.2 percent in the tertiary sector (trade, services, public administration, education, among others).

With its 20 *arrondissements* expanding over 105.4 km² for a density of 21,347 inhabitants per square kilometer, the city of Paris constitutes the CBD of the Paris metropolitan area. In 2011, about 2,250,000 inhabitants lived in Paris, corresponding to 19 percent of the metropolitan area's total population. The CBD also accounted for 32 percent of the

¹Mainland France comprises more than 36,500 municipalities.

metropolitan area's employment, with about 1,800,000 jobs, concentrated in the tertiary sector: 67.9 percent in trade, transportation and services and another 24.4 percent in public administration, education, health and social services. We can notice from these figures that tertiary sector jobs are over-represented in the CBD: 92.3 percent of all jobs in Paris compared to 86.2 percent of all jobs in the whole metropolitan area.

As most large agglomerations, the Paris metropolitan area also includes several employment subcenters. We identify them using the method first developed by [McDonald and Prather \(1994\)](#) and improved by [McMillen \(2001\)](#). The general idea is to estimate densities following a monocentric spatial pattern. The predicted densities obtained are then subtracted from the corresponding real densities. From these residuals, those that are positive and statistically significant are selected and defined as subcenters.² This method, which is described more thoroughly in Appendix A, enables us to identify 21 employment subcenters made of 88 municipalities in 1968 and 34 subcenters including 89 municipalities in 2010.

After identifying the main characteristics of the urban spatial structure in Paris metropolitan area, we now study its temporal and spatial trends: we characterize the importance of the CBD, the subcenters, and the other municipalities, before looking at the spatial influence of the CBD and the subcenters on the location of firms and households.

All information about population and employment used in this paper comes from various censuses provided by the French statistical institute, the INSEE.³ Each census wave

²The empirical literature has proposed alternative procedures to identify subcenters. Among them, the most used are those based on density and employment thresholds ([Giuliano and Small, 1991](#); [Giuliano et al., 2007](#); [Muñiz et al., 2008](#)) and on employment and/or population density peaks ([McDonald, 1987](#); [McDonald and Prather, 1994](#); [Sivitanidou, 1996](#); [Craig and Ng, 2001](#); [Muniz et al., 2003](#); [Redfearn, 2007](#); [Garcia-López, 2010](#); [Craig et al., 2016](#)). Although they sometimes identify different sets of subcenters in the same city (for instance, [Redfearn \(2007\)](#) finds different subcenters in Los Angeles when using his own methodology than when relying on [Giuliano and Small \(1991\)](#) and [McMillen \(2001\)](#)'s procedures), there are examples in which different methods lead to similar sets of subcenters (see for instance [Craig et al. \(2016\)](#) in Houston when using [McMillen \(2001\)](#) and [Redfearn \(2007\)](#)'s strategies). In our case, we rely on [McMillen \(2001\)](#)'s method not only because it is consistent with theory but also because it is directly linked and based on (deviations from) the monocentric model, which is consistent with the historical spatial structure of the Paris metropolitan area.

³Between the early 1960s and the late 1990s, one census surveying all individuals living in France was conducted about every decade. Since 2004, the design and sampling methodology of the census has changed

provides us with the number of individuals living and working in each municipality, and enables us to determine their socio-economic and demographic characteristics (e.g. level of education, socio-economic category of the job, type of occupation, age, gender, nationality, birth country, marital status, household size). Because our study is based on the 1968-2010 period, during which the railroad network underwent significant improvements (further details on this topic are given in section 2.2), we therefore rely on the 1968, 1975, 1982, 1990, 1999 and 2010 waves of the census.

Table 1 reports the number of jobs (Panel A) and inhabitants (Panel B) in the CBD, the identified subcenters and the remaining municipalities in 1968 and 2010. We alternatively refer to subcenters identified in 1968 (first three columns) or in 2010 (last three columns). The total numbers in the bottom line of each panel reveal that the Paris metropolitan area as a whole grew by about one third over the period, both in terms of employment (32.6 percent) and population (27.4 percent). Disaggregating these figures between CBD, subcenters and other locations enables us to detect the suburbanization process experienced by the Paris metropolitan area since 1968. Indeed, we see that the number of jobs in the CBD decreased by 7.1 percent, while population fell by 13.4 percent, to the benefit of subcenters and other municipalities. This evolution reflects an absolute suburbanization process. We can also note that the CBD's share of total employment and population dropped respectively from 45.3 percent to 31.7 percent and from 28 percent to 19 percent, revealing that the CBD's decentralization process was also relative. Taking a closer look at the subcenters and comparing the 1968 situation of the subcenters identified in 1968 with the 2010 situation of those identified in 2010, we can observe that they gained in terms of employment, both in absolute and relative terms (from 33 percent to 35 percent), illustrating a process of absolute and relative employment centralization in the subcenters. On the other hand, subcenters lost in terms of population, both in absolute and relative terms

completely, and is now conducted annually over a fraction of the population, so that the census data labeled "year n " is in fact collected over five years ($n - 2$ to $n + 2$). More details on this new sampling methodology can be found in English on the [INSEE webpage](#).

(from 34 percent to 26 percent), to the benefit of other municipalities: the subcenters have themselves been undergoing a population suburbanization process towards the smaller municipalities.

Table 1: Employment and population in metropolitan Paris, 1968–2010

Panel A: Employment	1968 Subcenters			2010 Subcenters		
	1968	2010	1968–2010	1968	2010	1968–2010
Paris	1,935,716 (45.26%)	1,797,678 (31.71%)	-138,038 (-7.1%)	1,935,716 (45.26%)	1,797,678 (31.71%)	-138,038 (-7.1%)
Subcenters	1,419,072 (33.18%)	1,762,894 (31.10%)	343,822 (24.2%)	1,132,124 (26.47%)	1,978,722 (34.91%)	846,598 (74.8%)
Non-central municipalities	921,992 (21.56%)	2,108,330 (37.19%)	1,186,338 (129%)	1,208,940 (28.27%)	1,892,502 (33.38%)	683,562 (56.5%)
Total	4,276,780	5,668,902	1,392,122 (32.6%)	4,276,780	5,668,902	1,392,122 (32.6%)
Panel B: Population	1968 Subcenters			2010 Subcenters		
	1968	2010	1968–2010	1968	2010	1968–2010
Paris	2,590,771 (28.01%)	2,243,833 (19.04%)	-346,938 (-13.4%)	2,590,771 (28.01%)	2,243,833 (19.04%)	-346,938 (-13.4%)
Subcenters	3,153,224 (34.10%)	3,472,991 (29.46%)	319,767 (10.1%)	2,370,046 (25.63%)	3,103,007 (26.33%)	732,961 (30.9%)
Non-central municipalities	3,504,637 (37.89%)	6,069,410 (51.50%)	2,564,773 (73.2%)	4,287,815 (46.36%)	6,439,394 (54.63%)	2,151,579 (50.2%)
Total	9,248,632	11,786,234	2,537,602 (27.4%)	9,248,632	11,786,234	2,537,602 (27.4%)

Note: Metropolitan shares and growth rates in parentheses.

In order to get an idea of whether the CBD and the subcenters influence the intrametropolitan distribution of employment and population, we regress the 2010 employment and population densities (in log) on the distance to CBD and the distance to the nearest employment subcenter (where we alternatively use subcenters identified in 1968 and 2010), controlling for a vector of geographic characteristics (land area, altitude, ruggedness index and elevation range). Municipal employment and population densities are found to be larger as the municipality gets closer to the CBD or to a subcenter (this trend is particularly marked for population density), revealing that both CBD and subcenters influence the spatial pattern of employment and population location. If we perform the same exercise considering population or employment density *growth* as the dependent variable, we observe that this growth was larger for municipalities closer to the CBD and subcenters, suggesting an increasing influence of both CBD and subcenters on the location of jobs and

residences despite the suburbanization process.⁴

2.2 *Transportation and the location of employment and population in 2010*

The transportation infrastructure of the Paris metropolitan area today is based on both a railroad network and a main road system. There are four types of railway networks. First, a suburban train (train henceforth) that connects Paris to the suburbs, including some of the most remote parts of the region. This network was initiated during the first half of the nineteenth century, and has been continuously expanded since then. An important modernization wave took place in the 1960s, with all steam trains being replaced by electric trains. Panel A of Table B.1 in Appendix B shows the most recent evolution of the train network. This network, composed of five lines with a total length of 788 km and with 231 stations located in 196 municipalities in 2010, has undergone a slight reduction in the past 40 years.

The Paris region is also endowed with a regional express network, the RER (*Réseau Express Régional*) which started operating during the second half of the 1970s. Figure B.1 in Appendix B shows the evolution of the RER network between 1975 and 2010. Like the train, the RER connects Paris to the suburbs, but for a shorter maximum distance: the furthest RER stations are located about 30 km away from Paris. Most of the RER lines follow the train lines and were designed to improve the former network. An important distinction between the train and RER networks is that the latter has connections within Paris. The RER thus enables passengers to commute from one part of the Paris metropolitan area to another, going through Paris, but without having to switch to another train or metro to cross the city. This represents a clear improvement to regional transit overall. Between 1975 and 2010, the RER network increased its number of lines from 1 to 5, its total length from 39 to 587 km, its number of stations from 22 to 243, and the number of municipalities having an RER station grew from 16 to 167 (see Appendix B Table B.1 Panel B). We refer to

⁴The corresponding results are not reported in the paper for the sake of brevity, but are available upon request to the authors.

[Mayer and Trévien \(2015\)](#) for a detailed history of the rail network in the Paris metropolitan area, and for a thorough explanation of the differences between the two regional train networks.

In addition to these regional railroad networks, Paris is endowed with a very dense subway system (*métro* henceforth), which was opened in 1900 and mainly connects areas within Paris. Between 1968 and 2010, the *métro* network was further expanded with the addition of two new lines that increased its length by 44 km, and 34 new stations were added connecting 13 new municipalities (Appendix B Table B.1 Panel C). Today, a few *métro* stations extend beyond Paris, but they remain within a very limited range (contiguous municipalities mostly). Finally, the Paris metropolitan area also enjoys a tramway network, which is much more recent: the first segments started operating in the beginning of the 1990s, and the network is still expanding. This network is mostly located at the fringe of Paris, with some segments running in the first ring of municipalities around Paris. Note that while the main regional trains have a radial structure, linking Paris to the suburbs, the tramway is much more circular, the various lines forming a circle around the CBD and along its borders. In 2010, this network was based on 4 lines with a total length of 40 km, with 70 stations connecting 19 municipalities (Appendix B Table B.1 Panel D).

In the case of the main road system, we focus on the highway network (and include some other main roads). Although France's first highway projects date from the 1920s and the 1930s, the real expansion of the French network took place during the second half of the 20th century. In the Paris metropolitan area (Appendix B Table B.1 Panel E), the number of highways increased from 11 to 41 between 1968 and 2010, expanding the network from 229 km with 46 ramps in 40 municipalities to 821 km with 168 ramps in 133 municipalities.

We now want to investigate the spatial influence of transportation on the intrametropolitan distribution of employment and population in the Paris metropolitan area in 2010. As mentioned in the introduction, this analysis is likely to be flown by the identification issue

that is pervasive in the literature: transportation and its improvements are not decided and located randomly. On the contrary, they are endogenous to employment and/or population growth: planners may decide to improve the connection of deprived areas in order to boost their economic activity or attract population, or they may on the contrary connect areas that are expected to grow, anticipating on future needs for transportation. Following recent literature ([Duranton and Turner, 2012](#); [Garcia-López, 2012](#); [Garcia-López et al., 2015b](#)), we rely on historical railways and roads to instrument the location of modern transportation infrastructures in the Paris metropolitan area. More precisely, we use Roman roads and the 1870 railroads as alternative instrumental variables. We provide evidence of the validity in terms of exogeneity and relevance of these instruments in Appendix C.

Since our focus is on the RER, we regress the log of the 2010 employment and population densities on the distance to the nearest RER station in 2010, controlling for the distance to other types of transportation infrastructures and for the distance to the nearest 2010 employment center, where distances are alternatively included as such or in logarithm⁵. We also control for geographical characteristics (land area, altitude, ruggedness index and elevation range), as well as for historical dummy variables indicating (1) whether municipalities were Roman settlements, (2) whether they used to be major towns between the 10th and the 15th centuries and (3) between the 16th and the 19th centuries, (4) whether they had a monastery built between the 12th and 16th centuries, and (5) whether they hosted important fairs between the 10th and the 16th centuries.⁶

⁵As [Redfearn \(2009\)](#) and [McMillen and Redfearn \(2010\)](#) highlight, it is difficult to choose functional forms in spatial data. As a result, we use the most common functional forms in the literature: direct and logged distances.

⁶These variables come from the [Digital Atlas of Roman and Medieval Civilizations](#), with the exception of the major cities of the 16th to 19th centuries which are identified in [Bairoch \(1988\)](#).

$$\begin{aligned}
2010 \ln(\text{density}) = & \delta_0 + \delta_1 \times 2010 f(\text{distance to RER station}) \\
& + \delta_2 \times 2010 f(\text{distance to non-RER station or ramp}) \\
& + \delta_3 \times f(\text{distance to the nearest 2010 employment center}) \\
& + \sum_i (\delta_{4,i} \times \text{geography}_i) + \sum_i (\delta_{5,i} \times \text{history}_i).
\end{aligned} \tag{1}$$

Because we either use direct measures of distances in kilometers or their logarithm, the f function stands alternatively for the identity function or the logarithm function, respectively. The coefficients δ_1 and δ_2 represent density gradients when we use direct distances, and elasticities when we use logged distances. They capture the extent to which density increases with proximity to the nearest RER station and to the nearest non-RER station or highway ramp, respectively. In order to address the endogeneity issue, we estimate this equation using a two-stage least square (TSLS) procedure, where Roman roads and 1870 railroads are used as instruments for the RER and the non-RER variables (see details in [Appendix C](#)).

Table 2 reports results for Equation (1) in terms of employment density (columns 1 to 5) and population density (columns 6 to 10) when using direct distances (Panel A) and logged distances (Panel B). In all cases, we find that transportation infrastructures do influence the location of employment and population: the estimated coefficients are always negative and significant. In particular, our results show that getting closer to a RER station by one kilometer (or 10 percent if we read our results in terms of elasticities) increases employment density by around 6 percent and population density by 4 to 5 percent. Results also show higher (but less significant) effects for non-RER transportation: each additional kilometer closer to a non-RER station or ramp increases employment and population densities by 7 to 8 percent and 6 to 7 percent, respectively. In terms of elasticities, we can say that getting 10 percent closer to a non-RER infrastructure leads to a 5 to 10 percent rise

in employment density and to a 4 to 9 percent rise in population density. In the rest of the paper, we focus on results controlling only for the group of non-RER infrastructures, including highways ramps (i.e. the specifications used in columns 5 and 10 of Table 2), since these non-RER coefficients are of the same order of magnitude (considering their standard errors), and do not affect the RER coefficient much.

Table 2: Urban spatial structure and proximity to RER and other transportation, TSLS

Dependent variable:	2010 ln(Employment density)					2010 ln(Population density)				
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Panel A: Direct distances										
2010 Distance to the nearest RER station	-0.065 ^a (0.007)	-0.063 ^a (0.007)	-0.063 ^a (0.007)	-0.052 ^a (0.009)	-0.055 ^a (0.008)	-0.050 ^a (0.005)	-0.050 ^a (0.006)	-0.050 ^a (0.005)	-0.041 ^a (0.007)	-0.043 ^a (0.006)
2010 Distance to the nearest commuter train		-0.074 ^c (0.039)					-0.061 ^b (0.030)			
2010 Distance to the nearest non-RER station			-0.079 ^c (0.042)					-0.066 ^b (0.032)		
2010 Distance to the nearest highway ramp				-0.072 ^c (0.038)					-0.060 ^b (0.030)	
2010 Distance to the nearest non-RER stat/ramp					-0.084 ^c (0.044)					-0.069 ^b (0.034)
First-stage statistic	226.48	38.46	34.00	23.68	54.16	226.48	38.46	34.00	23.68	54.16
Panel B: Logged distances										
2010 ln(Dist to the nearest RER station)	-0.679 ^a (0.067)	-0.620 ^a (0.066)	-0.605 ^a (0.067)	-0.542 ^a (0.084)	-0.550 ^a (0.075)	-0.550 ^a (0.052)	-0.498 ^a (0.051)	-0.485 ^a (0.051)	-0.429 ^a (0.069)	-0.436 ^a (0.057)
2010 ln(Dist to the nearest commuter train)		-0.432 ^a (0.150)					-0.380 ^a (0.116)			
2010 ln(Dist to the nearest non-RER station)			-0.472 ^a (0.161)					-0.415 ^a (0.125)		
2010 ln(Dist to the nearest highway ramp)				-1.029 ^b (0.410)					-0.905 ^a (0.338)	
2010 ln(Dist to the nearest non-RER stat/ramp)					-0.536 ^a (0.185)					-0.471 ^a (0.145)
First-stage statistic	172.46	48.71	40.21	10.13	35.23	172.46	48.71	40.21	10.13	35.23
Distance to the nearest 2010 center	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Geography	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
History	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Instrument:										
Distance to the nearest 1870 railroad line	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Distance to the nearest Roman road	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: 1300 observations for each regression. Geography variables are land area, altitude, index of terrain ruggedness, and elevation range. History variables are dummy variables for municipalities (1) that were Roman settlements (based on DARMC maps), (2) that were major towns between the 10th and the 15th centuries (based on DARMC maps), (3) that were major towns between the 16th and the 19th centuries (based on Bairoch, 1988), (4) with a monastery built between the 12th and 16th centuries (based on DARMC maps), and (5) that hosted important fairs between the 10th and the 16th centuries (based on DARMC maps). Robust standard errors are in parentheses. ^a, ^b, and ^c indicates significant at 1, 5, and 10 percent level, respectively.

3 The effect of the RER on local growth in Paris metropolitan area

By investigating the effects of transportation's improvements on local growth in employment and population in the Paris metropolitan area, our paper brings several new insights to the related literature. First, it focuses on the intrametropolitan level, that is, on the municipalities that make up the Paris metropolitan area, while most previous studies are at the city-metropolitan level. Second, we study the effects on both employment and population growth, while previous studies focus on just one or the other. Finally, although our main interest is the effect of the RER, we also control for other modes of transportation, while most of the previous studies consider just one type of infrastructure.

In this section, we start by studying the effects of RER on local growth for all 1300 municipalities that make up Paris metropolitan area. We then explore whether the RER effects are heterogeneous across space by grouping municipalities according to their proximity to a RER station and to their proximity to employment (sub)centers. Afterwards, we study heterogeneous effects according to the type of employment (sectors and occupations) and the education level of population, before analyzing the temporal scope of the RER effects across time periods.

3.1 *Average metropolitan effects*

We begin by analyzing the impact of RER and other transportation on local (municipal) growth, both in terms of employment and population. Since it might take years for firms and households to relocate in response to transportation improvements, we estimate 'traditional' growth equations, in which a 'growth' dependent variable (between years t and $t-1$) is regressed on a set of explanatory variables measured in the initial year $t-1$. Here, we focus on the 1968–2010 period and estimate the following equation:

$$\begin{aligned}
1968\text{--}2010 \Delta \ln(\text{density}) = & \mu_0 + \mu_1 \times 2010 f(\text{distance to RER station}) \\
& + \mu_2 \times 1968 f(\text{distance to non-RER stations \& ramps}) \\
& + \mu_3 \times 1968 \ln(\text{densities}) \\
& + \mu_4 \times f(\text{distance to the nearest 1968 empl. center}) \quad (2) \\
& + \sum_i (\mu_{5,i} \times \text{geography}_i) + \sum_i (\mu_{6,i} \times \text{history}_i) \\
& + \sum_i (\mu_{7,i} \times 1968 \text{ socioeconomy}_i).
\end{aligned}$$

It is important to point out that, since there were no RER stations in 1968, our main explanatory variable is the 2010 distance to the nearest RER station (or its logarithm depending on specifications). On the other hand, since there were other railroads and highways in the initial year, we include the distance to the nearest access to a non-RER infrastructure (station or ramp) in 1968 (or its logarithm). We also control for characteristics related to the initial urban spatial structure of the Paris metropolitan area, *i.e.* the 1968 employment and population densities, the (log) distance to the nearest 1968 employment center, and geography and history variables. We additionally control for the 1962 population size, and 1968 socioeconomic characteristics at the municipal level: unemployment rate; share of employment in manufacturing, in construction, and in services, used as proxies for economic specialization; share of executives and professional workers, to account for average income level; and share of population with university degree, as a proxy for the level of human capital. Here again, we run two-stage least square regressions using distance to the nearest 1870 railroad and distance to the nearest Roman road as instruments to correct for endogeneity (see Appendix C for further details).

Table 3 reports our main TSLS results for employment (columns 1 to 3) and for population (columns 4 to 6) when using direct distances (Panel A) and logged distances (Panel B). In order to understand the determinants of employment growth, we start by including

the 2010 distance to RER station only in column 1; then, we only include the 1968 distance to non-RER transportation in column 2; and we eventually we include both transportation variables in column 3 (all specifications additionally include the other control variables). The corresponding results for population growth are reported in columns 4, 5 and 6, respectively. We find negative and significant effects for RER and non-RER transportation, revealing that employment and population growth increase the closer a municipality is to a railroad station (RER and non-RER) or a highway ramp.

Table 3: The effect of RER on municipality growth, TSLS: Average metropolitan effects

Dependent variable:	1968–2010 $\Delta \ln(\text{Employment density})$			1968–2010 $\Delta \ln(\text{Population density})$		
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A: Direct distances						
2010 Distance to RER station	-0.018 ^a (0.005)		-0.023 ^a (0.005)	-0.011 ^a (0.003)		-0.014 ^a (0.004)
1968 Distance to non-RER stat/ramp		-0.068 ^a (0.023)	-0.089 ^a (0.022)		-0.048 ^a (0.015)	-0.061 ^a (0.015)
1968 $\ln(\text{Employment density})$	-0.543 ^a (0.078)	-0.574 ^a (0.080)	-0.543 ^a (0.076)	-0.011 (0.058)	-0.030 (0.058)	-0.011 (0.057)
1968 $\ln(\text{Population density})$	0.629 ^a (0.089)	0.653 ^a (0.075)	0.463 ^a (0.088)	0.038 (0.063)	0.044 (0.050)	-0.075 (0.061)
First-stage statistic	164.05	61.72	30.22	164.05	61.72	30.22
Panel B: Logged distances						
2010 $\ln(\text{Distance to RER station})$	-0.212 ^a (0.059)		-0.245 ^a (0.059)	-0.132 ^a (0.038)		-0.156 ^a (0.040)
1968 $\ln(\text{Distance to non-RER stat/ramp})$		-0.436 ^a (0.137)	-0.508 ^a (0.128)		-0.310 ^a (0.092)	-0.356 ^a (0.087)
1968 $\ln(\text{Employment density})$	-0.553 ^a (0.077)	-0.609 ^a (0.083)	-0.583 ^a (0.077)	-0.014 (0.057)	-0.052 (0.061)	-0.036 (0.058)
1968 $\ln(\text{Population density})$	0.626 ^a (0.090)	0.630 ^a (0.078)	0.443 ^a (0.090)	0.036 (0.064)	0.027 (0.051)	-0.092 (0.062)
First-stage statistic	154.04	41.71	37.92	154.04	41.71	37.92
Distance to the nearest 1968 center	Y	Y	Y	Y	Y	Y
Geography	Y	Y	Y	Y	Y	Y
History	Y	Y	Y	Y	Y	Y
1968 Socioeconomy	Y	Y	Y	Y	Y	Y
Instrument:						
Distance to the nearest 1870 railroad line	✓	✓	✓	✓	✓	✓
Distance to the nearest Roman road	✓	✓	✓	✓	✓	✓

Notes: 1300 observations for each regression. Geography variables are land area, altitude, index of terrain ruggedness, and elevation range. History variables are the population level in 1962 and dummy variables for municipalities (1) that were Roman settlements (based on DARMC maps), (2) that were major towns between the 10th and the 15th centuries (based on DARMC maps), (3) that were major towns between the 16th and the 19th centuries (based on Bairoch, 1988), (4) with a monastery built between the 12th and 16th centuries (based on DARMC maps), and (5) that hosted important fairs between the 10th and the 16th centuries (based on DARMC maps). Socioeconomic variables are the 1968 unemployment rate, the 1968 shares of employment in Manufacturing, in Construction, and in Services, the 1968 share of executives and professionals, and the 1968 share of population with university degree. Robust standard errors are in parentheses. ^a, ^b, and ^c indicates significant at 1, 5, and 10 percent level, respectively.

More precisely, the specifications including distances to both RER and non-RER infrastructures (columns 3 and 6) reveal that each additional kilometer closer to the nearest RER station increases employment and population growth by about 2 percent and 1 percent respectively. The corresponding elasticities are of 2 percent and 1.5 percent for a 10 percent increase in distance to the closest RER station. Yet, the effects are higher for the 1968 non-RER transportation, which is found to increase employment and population growth by 9 percent and 6 percent, respectively; the corresponding elasticities being of 5 percent and 4 percent respectively.⁷ Finally, it is important to notice that the coefficients for both distances are not statistically different when they are individually, as opposed to jointly, estimated (columns 1 and 2 vs. 3 for employment, columns 4 and 5 vs. 6 for population). We take advantage of this feature in our last empirical analysis.

A closer look at the estimates for initial employment and population densities also provides interesting insights to understand the dynamic of employment growth (columns 1 to 3). The positive and significant coefficients for the 1968 log of population density reveal that employment growth is higher in municipalities that were initially more densely populated. On the other hand, the negatively significant coefficients for the 1968 log of employment density show that employment growth is lower in municipalities which initially had a larger employment density. Combined together, these two results tell us that employment follows population.

3.2 *Proximity matters!*

Admittedly, the results discussed above show limited growth effects. This is not surprising since, as we have already noticed, these effects are *average* effects estimated for the 1,300 municipalities that make up the Paris metropolitan area. In this subsection, we explore whether these effects are uniform or rather heterogeneous across space.

⁷We also conducted the estimations with the 2010 (log of) distance to the nearest non-RER transportation in specifications corresponding to those in columns 2, 3, 5 and 6. The estimated coefficients for the (log) 2010 non-RER distance are statistically indistinguishable from those for the (log) 1968 distance reported in Table 3.

Table 4: The effect of RER on municipality growth, TSLS: The closest municipalities

Dependent variable:	1968–2010 $\Delta \ln(\text{Employment density})$					1968–2010 $\Delta \ln(\text{Population density})$				
	dist to RER					dist to RER				
	≤ 1 km	≤ 2 km	≤ 3 km	≤ 4 km	≤ 5 km	≤ 1 km	≤ 2 km	≤ 3 km	≤ 4 km	≤ 5 km
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Panel A: Direct distances										
2010 Distance to RER	-0.081 ^a (0.023)	-0.063 ^a (0.015)	-0.056 ^a (0.015)	-0.042 ^a (0.014)	-0.035 ^b (0.014)	-0.062 ^a (0.017)	-0.040 ^a (0.012)	-0.031 ^a (0.012)	-0.027 ^a (0.010)	-0.024 ^b (0.009)
First-stage statistic	20.12	38.25	43.61	54.05	59.20	20.12	38.25	43.61	54.05	59.20
Panel B: Logged distances										
2010 $\ln(\text{Dist to RER})$	-0.642 ^a (0.206)	-0.754 ^a (0.196)	-0.799 ^a (0.194)	-0.643 ^a (0.174)	-0.606 ^a (0.167)	-0.522 ^a (0.157)	-0.475 ^a (0.144)	-0.473 ^a (0.137)	-0.413 ^a (0.112)	-0.394 ^a (0.108)
First-stage statistic	21.56	26.42	31.82	44.68	50.33	21.56	26.42	31.82	44.68	50.33
1968 Dist to non-RER	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
1968 $\ln(\text{Densities})$	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Dist to 1968 center	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Geography	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
History	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
1968 Socioeconomy	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Instrument:										
Dist to 1870 railroad	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Dist to Roman road	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations:	188	321	421	476	530	188	321	421	476	530

Notes: Geography variables are land area, altitude, index of terrain ruggedness, and elevation range. History variables are the population level in 1962 and dummy variables for municipalities (1) that were Roman settlements (based on DARMC maps), (2) that were major towns between the 10th and the 15th centuries (based on DARMC maps), (3) that were major towns between the 16th and the 19th centuries (based on Bairoch, 1988), (4) with a monastery built between the 12th and 16th centuries (based on DARMC maps), and (5) that hosted important fairs between the 10th and the 16th centuries (based on DARMC maps). Socioeconomic variables are the 1968 unemployment rate, the 1968 shares of employment in Manufacturing, in Construction, and in Services, the 1968 share of executives and professionals, and the 1968 share of population with university degree. Robust standard errors are in parentheses. ^a, ^b, and ^c indicates significant at 1, 5, and 10 percent level, respectively.

We start by grouping municipalities according to their proximity to a RER station. Since the average commuting distance in metropolitan Paris is around 11-13 km (Aguilera and Mignot, 2004; Aguilera, 2005) and walking and biking are key to rail use, we focus on the closest municipalities and consider five different concentric rings each 1 kilometer larger than the previous, starting from a minimum distance of 1 kilometer from the nearest RER station to as far as 5 kilometers. Following equation (2), we thus estimate the effect on employment and population growth of the distance to a RER station for five different subset of municipalities: 188 municipalities located less than 1 km from a RER station; 321 located less than 2 km away; 421 located less than 3 km away; 476 located less than 4 km away; and 530 located less than 5 km away. The corresponding results are displayed in

Table 4, where density gradients are reported in Panel A and elasticities in Panel B. The estimated effects on employment (resp. population) are reported in columns 1 to 5 (resp. 6 to 10), each column corresponding to a different subset of municipalities (from less than 1 km away in columns 1 and 6 to less than 5 km away in columns 5 and 10). These results confirm that the effect decreases with the distance from a RER station, and clearly reveal that it is particularly stronger in municipalities which are closer to a RER station: from 8 percent in the 1-km ring to 3 percent in the 5-km ring for employment growth, and from 6 percent to 2 percent for population growth. By contrast, the corresponding results over all municipalities were of about 2 percent for employment and 1 percent for population (columns 3 and 6 of Table 3, Panel A).

Table 5: The effect of RER on municipality growth, TSLS: Core vs. periphery

Dependent variable:	1968–2010 $\Delta \ln(\text{Employment density})$				1968–2010 $\Delta \ln(\text{Population density})$			
	dist to centers				dist to centers			
	≤ 5 km	≤ 5 km (no centers)	> 5 km	> 5 km (no centers)	≤ 5 km	≤ 5 km (no centers)	> 5 km	> 5 km (no centers)
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Panel A: Direct distances								
2010 Distance to RER	-0.074 ^a (0.028)	-0.125 ^a (0.046)	-0.019 ^a (0.005)	-0.019 ^a (0.005)	-0.063 ^a (0.018)	-0.091 ^a (0.029)	-0.012 ^a (0.003)	-0.012 ^a (0.003)
First-stage statistic	4.57	3.21	31.56	30.40	4.57	3.21	31.56	30.40
Panel B: Logged distances								
2010 $\ln(\text{Dist to RER})$	-0.773 ^b (0.324)	-0.935 ^b (0.380)	-0.198 ^a (0.059)	-0.203 ^a (0.059)	-0.655 ^a (0.222)	-0.654 ^a (0.244)	-0.123 ^a (0.041)	-0.127 ^a (0.040)
First-stage statistic	2.95	4.06	33.45	31.94	2.95	4.06	33.45	31.94
1968 Dist to non-RER	Y	Y	Y	Y	Y	Y	Y	Y
1968 $\ln(\text{Densities})$	Y	Y	Y	Y	Y	Y	Y	Y
Dist to 1968 center	Y	Y	Y	Y	Y	Y	Y	Y
Geography	Y	Y	Y	Y	Y	Y	Y	Y
History	Y	Y	Y	Y	Y	Y	Y	Y
1968 Socioeconomy	Y	Y	Y	Y	Y	Y	Y	Y
Instrument:								
Dist to 1870 railroad	✓	✓	✓	✓	✓	✓	✓	✓
Dist to Roman road	✓	✓	✓	✓	✓	✓	✓	✓
Observations:	221	159	1079	1033	221	159	1079	1033

Notes: Geography variables are land area, altitude, index of terrain ruggedness, and elevation range. History variables are the population level in 1962 and dummy variables for municipalities (1) that were Roman settlements (based on DARMC maps), (2) that were major towns between the 10th and the 15th centuries (based on DARMC maps), (3) that were major towns between the 16th and the 19th centuries (based on Bairoch, 1988), (4) with a monastery built between the 12th and 16th centuries (based on DARMC maps), and (5) that hosted important fairs between the 10th and the 16th centuries (based on DARMC maps). Socioeconomic variables are the 1968 unemployment rate, the 1968 shares of employment in Manufacturing, in Construction, and in Services, the 1968 share of executives and professionals, and the 1968 share of population with university degree. Robust standard errors are in parentheses. ^a, ^b, and ^c indicates significant at 1, 5, and 10 percent level, respectively.

We then group municipalities according to their proximity to employment (sub)centers. The idea is to compare municipalities in denser areas to those in peripheral areas because they might be different in their commuting patterns, types of employment and/or population, housing stock and tenure, among other characteristics. Table 5 reports results of estimating equation (2) for employment (columns 1 to 4) and population (columns 5 to 8) when using direct distances (Panel A) and logged distances (Panel B). We split the 1,300 municipalities between the 221 located less than 5 km from an employment center (columns 1-2 and 5-6), and the remaining 1,079 located beyond 5 km (columns 3-4 and 7-8). Additionally, in some regressions we also drop municipalities that are part of (sub)centers (columns 2, 4, 6 and 8). In general, all results reinforce the previous idea that the effect of proximity to RER is higher in municipalities in denser areas than in peripheral ones: 7-6 percent (columns 1-2 and 5-6) vs. 2-1 percent (columns 3-4 and 7-8). Furthermore, in denser areas the effect is higher in non-central municipalities: the estimated coefficients increase when we drop (sub)centers from the sample (columns 1 and 5 vs. columns 2 and 6).

In summary, although the results reported in Table 3 show limited *average* growth effects, we can clearly see that these effects are heterogenous across space: the previous analysis clearly shows that growth effects are *locally higher* close to RER stations and employment (sub)centers. In other words, proximity matters!

3.3 *Employment and population types matter!*

The Paris metropolitan area has undergone important demographic and socioeconomic changes between 1968 and 2010. Some of these changes are related to the type of employment and population. In terms of sector of activity, the share of total employment decreased from 24 to 6 percent in agriculture, from 14 to 11 percent in construction and from 22 to 11 percent in manufacturing, while it increased from 40 to 72 percent in services.

A similar trend can be observed when we look at occupations: the share of total jobs decreased from 12 to 4 percent for farmers, from 11 to 10 percent for craftsmen and from 43 to 21 percent for factory workers, while it increased from 3 to 14 percent for executives and professionals, from 11 to 24 percent for intermediate occupations and from 19 to 26 percent employees.

Regarding the education level as well, we observe a decrease in the share of non-educated individuals from 47 to 29 percent and in the share of low educated individuals from 45 to 24 percent, while the share of population with high school degree and university degree increased from 5 to 18 percent and 3 to 29 percent, respectively.

As a result, we now investigate the existence of heterogeneous effects according to the type of employment (sectors and occupations) and education level of population.

Table 6 reports TSLS results for employment across sectors (agriculture; construction; manufacturing; services in columns 1 to 4), and occupations (farmers; craftsmen; executives and professionals; intermediates; employees; factory workers in columns 5 to 10), and for population across education levels (non-educated; low educated; high school degree; university degree in columns 11 to 14). As before, Panel A displays results obtained using direct distances and Panel B those using logged distances.

A first look at the results reveals that the presence of a RER station has heterogeneous effects in this respect as well. First, while proximity to RER does not affect employment growth in agriculture and construction, getting closer to a RER station increases employment growth in manufacturing and in services especially. Second, only executives and professional workers, intermediate workers, employees and factory workers benefit from proximity to a RER station. Finally, being closer to a RER station seems to systematically benefit population growth, with a slightly more marked effect for population with high school degree.

Table 6: The effect of RER on municipality growth, TSLS: Sectors, occupations and education

Dependent variable:	1968–2010 $\Delta \ln(\text{Employment density})$										1968–2010 $\Delta \ln(\text{Population density})$			
	Sectors					Occupations					Education level			
	Agr	Con	Man	Ser	Farm	Craft	Exec	Int	Empl	Fact	None	Low	HSch	Uni
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]
Panel A: Direct distances														
2010 Distance to RER	0.003 (0.006)	-0.011 (0.007)	-0.018 ^b (0.008)	-0.028 ^a (0.006)	0.009 ^c (0.006)	-0.007 (0.006)	-0.025 ^a (0.007)	-0.025 ^a (0.007)	-0.022 ^a (0.006)	-0.026 ^a (0.007)	-0.013 ^a (0.004)	-0.019 ^a (0.003)	-0.028 ^a (0.004)	-0.023 ^a (0.006)
First-stage statistic	29.47	30.32	30.61	30.71	31.10	30.79	30.61	29.83	30.33	30.48	35.75	32.77	38.92	36.31
Panel B: Logged distances														
2010 $\ln(\text{Dist to RER})$	0.019 (0.070)	-0.100 (0.079)	-0.174 ^c (0.093)	-0.307 ^a (0.066)	0.109 (0.069)	-0.064 (0.071)	-0.257 ^a (0.087)	-0.271 ^a (0.078)	-0.231 ^a (0.072)	-0.292 ^a (0.075)	-0.134 ^a (0.038)	-0.213 ^a (0.035)	-0.298 ^a (0.047)	-0.232 ^a (0.061)
First-stage statistic	38.28	38.35	38.77	38.40	39.92	38.78	38.64	37.97	38.06	38.61	43.66	38.15	45.00	44.05
1968 Dist to non-RER	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
1968 $\ln(\text{Densities})$	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Dist to 1968 center	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Geography	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
History	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
1968 Socioeconomy	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Instrument:														
Dist to 1870 railroad	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Dist to Roman road	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: 1300 observations for each regression. Geography variables are land area, altitude, index of terrain ruggedness, and elevation range. History variables are the population level in 1962 and dummy variables for municipalities (1) that were Roman settlements (based on DARMC maps), (2) that were major towns between the 10th and the 15th centuries (based on DARMC maps), (3) that were major towns between the 16th and the 19th centuries (based on Bairoch, 1988), (4) with a monastery built between the 12th and 16th centuries (based on DARMC maps), and (5) that hosted important fairs between the 10th and the 16th centuries (based on DARMC maps). Socioeconomic variables are the 1968 unemployment rate, the 1968 shares of employment in Manufacturing, in Construction, and in Services, the 1968 share of executives and professionals, and the 1968 share of population with university degree. Robust standard errors are in parentheses. ^a, ^b, and ^c indicates significant at 1, 5, and 10 percent level, respectively.

3.4 Time matters!

We eventually investigate the potentially heterogenous effect of RER proximity over time. The purpose is threefold. First, to estimate a version of equation (2) in which our main explanatory variable, the distance to the nearest RER station, also uses values in the initial year. To consider this ‘more traditional’ growth equation, we focus on the 1975–2010 period, which witnessed the advent and expansion of the RER network. This is a way of testing the robustness of the estimated effect of the RER. Second, we also investigate the temporal scope of the RER effects. As mentioned above, employment and population responses to transportation improvements might take years. Since the length of this delay is unclear, we explore it by regressing growth equations for the 1975–1990 and 1990–2010 sub-periods. Finally, as shown in Table B.1, the different RER lines were built in separate periods: while in 1975 there was only 1 line, in 1990 there were 3 additional lines. As a result, this temporal analysis allow us to compare the effect of the whole network in 2010 with the effects of the RER lines existing in 1975 and 1990.⁸

To this aim, we regress the employment and population growth between *year t-1* and *t* on the distance to the nearest RER station in *year t-1* (or *t-2*), conditional on employment and population densities in *t-1*, and on distance to the nearest employment center, geography, history and socioeconomic variables in *t-1*:

$$\begin{aligned}
 \text{year } t-1 - \text{year } t \Delta \ln(\text{density}) = & \eta_0 + \eta_1 \times \text{year } t-1 \text{ or } t-2 f(\text{distance to RER station}) \\
 & + \eta_2 \times \text{year } t-1 \ln(\text{densities}) \\
 & + \eta_3 \times f(\text{dist to the nearest 1968 empl. center}) \\
 & + \sum_i (\eta_{4,i} \times \text{geography}_i) + \sum_i (\eta_{5,i} \times \text{history}_i) \\
 & + \sum_i (\eta_{6,i} \times \text{year } t-1 \text{ socioeconomy}_i).
 \end{aligned} \tag{3}$$

Compared to equation (2), equation (3) omits the distance to other transportation infras-

⁸Similar results, available upon request, are obtained with a cut off in 1982.

structures since we observed (at the end of Section 2.2) that it does not significantly affect the coefficients of interest. This empirical strategy also allows us to overcome a problem with one of our instruments, the distance to the nearest Roman road, which is not relevant in some periods. As a result, equation (3) is estimated by TSLS using the distance to the nearest 1870 railroad as the unique instrument.

Table 7 reports the results for employment (columns 1 to 4) and population (columns 5 to 8), alternatively using direct distances (Panel A) and logged distances (Panel B): columns 1 and 5 display estimates over the shortened 1975–2010 period, while the estimates in columns 2 and 6 are obtained on the first sub-period (1975–1990) and those in columns 3, 4, 7 and 8 correspond to the latest sub-period (1990–2010). Finally, while most specifications follow the traditional growth equation with the *year t-1* RER distance (columns 1 to 3 and 5 to 7), in columns 4 and 9 we use the *year t-2* RER distance (1975). Results for the whole 1975–2010 period (columns 1 and 5) are consistent with those obtained previously in Table 3: proximity to a RER station in 1975 increases employment and population growth. While RER effects when using direct distances are similar to their counterparts in Table 3 columns 3 and 6, the effects for logged distances are higher.

More interestingly, the analysis by sub-period shows different time responses to RER improvements for employment and population. For the case of employment, we do not find a significant RER effect on the 1975–1990 period (column 2). By contrast, the RER effect appears in the 1990–2010 period when we use the 1990 distance in column 3 (which includes both the 1975 and 1975–1990 RER networks) and it is clearer when we use the 1975 distance in column 4. Therefore, it seems that the RER effect emerges after a certain time lag. As for population, the RER effect turns out to be much more rapid, increasing population growth in the 1975–1990 period (column 6). However, at the same time, the RER effect tends to decrease over time: increasing population growth only by 0.3 percent (column 7) and 0.5 percent (column 8) in the 1990–2010 period.

To sum up, investigating the RER effects across different time periods confirms the

robustness of our previous results based on the 1968–2010 period. Most importantly, this analysis reveals that these effects are heterogenous across time periods and differ between firms and households: while there is a lagged response by firms which increases with time, the response in terms of residential location is more rapid, but decreases with time.

Table 7: The effect of RER on municipality growth, TSLS: Effects by periods

Dependent variable: Period	$\Delta \ln(\text{Employment density})$				$\Delta \ln(\text{Population density})$			
	1975–2010	1975–1990	1990–2010		1975–2010	1975–1990	1990–2010	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Panel A: Direct distances								
1975 Dist to RER station	-0.027 ^a (0.008)	-0.006 (0.007)		-0.022 ^a (0.008)	-0.018 ^a (0.004)	-0.014 ^a (0.004)		-0.005 ^c (0.003)
1990 Dist to RER station			-0.012 ^a (0.004)				-0.003 ^c (0.001)	
First-stage statistic	50.43	50.43	76.19	27.07	50.43	50.43	76.19	27.07
Panel B: Logged distances								
1975 Dist to RER station	-1.215 ^a (0.458)	-0.242 (0.327)		-1.301 ^b (0.622)	-0.857 ^a (0.261)	-0.644 ^a (0.220)		-0.344 ^c (0.182)
1990 Dist to RER station			-0.222 ^b (0.087)				-0.050 ^c (0.029)	
First-stage statistic	17.84	17.84	43.09	6.91	17.84	17.84	43.09	6.91
1975 or 1990 $\ln(\text{Densities})$	Y	Y	Y	Y	Y	Y	Y	Y
Dist to 1968 center	Y	Y	Y	Y	Y	Y	Y	Y
Geography	Y	Y	Y	Y	Y	Y	Y	Y
History	Y	Y	Y	Y	Y	Y	Y	Y
1975 or 1990 Socioeconomy	Y	Y	Y	Y	Y	Y	Y	Y
Instrument:								
Dist to 1870 railroads	✓	✓	✓	✓	✓	✓	✓	✓

Notes: 1300 observations for each regression. Geography variables are land area, altitude, index of terrain ruggedness, and elevation range. History variables are population levels in 1962 and 1968 (all columns) and 1975 and 1982 (columns 3 and 6); employment levels in 1968 (all columns) and 1975 and 1982 (columns 3 and 6); and dummy variables for municipalities (1) that were Roman settlements (based on DARMC maps), (2) that were major towns between the 10th and the 15th centuries (based on DARMC maps), (3) that were major towns between the 16th and the 19th centuries (based on Bairoch, 1988), (4) with a monastery built between the 12th and 16th centuries (based on DARMC maps), and (5) that hosted important fairs between the 10th and the 16th centuries (based on DARMC maps). Socioeconomic variables are the unemployment rate, the shares of employment in Manufacturing, in Construction, and in Services, the share of executives and professionals, and the share of population with university degree with their values in 1975 (columns 1–2 and 4–6) or in 1990 (columns 3 and 6). Robust standard errors are in parentheses. ^a, ^b, and ^c indicates significant at 1, 5, and 10 percent level, respectively.

4 Conclusions

In this paper, we investigated the effect of the RER expansion on employment and population growth in the municipalities of the Paris metropolitan area between 1968 and 2010, controlling for all other transportation modes. The main results indicate that the RER

network together with the other transportation networks have a positive and significant effect on the location of employment and population: we find that each kilometer closer to a RER station increases employment and population growth by 2 percent and 1 percent, respectively. Analyzing the heterogeneity of these results across space, employment and population types shows that these effects are higher (a) for municipalities which are the closest to a RER station and the closest to an employment (sub)center, (b) for services, for factory workers, and for population with high school and university degrees. Finally, looking at the heterogeneity over time reveals that there is no impact of the RER expansion on employment growth during the first part of the covered period, while the impact on population growth was sizable much earlier but declined over time.

This paper's contribution is non-negligible as it provides much-needed evidence from an analysis conducted at the intrametropolitan level in one of the largest metropolitan areas in Europe. Furthermore, our results for the Paris metropolitan area complement those obtained by [Mayer and Tréviën \(2015\)](#) using a different empirical strategy with a restricted sample of municipalities from the Paris metropolitan area. It is also important to note that some of our suburban and intrametropolitan results verify the theoretical predictions we discussed. First, we confirm that railroad and highway effects are heterogeneous in terms of distance to CBD. Second, we also provide evidence that the suburbanized population and employment are not evenly distributed across the suburbs: on the contrary, the population spreads out along the highways in the first stage of infrastructure development, while in this same stage, employment follows population.

A better understanding of the relationship between improvements in transport infrastructure, on the one hand, and city structure and city growth, on the other, is important, in general, for transport planners, urban planners, and policy makers and, in particular, it is crucial for making correct transport forecasts. Here, we have examined the impact of the initial stages in the development of the RER rail network on growth. Our results show that railroad investment has a major impact in these early years on population growth but

not on employment growth. Further research is, however, required to determine whether this effect is weakened as the network becomes denser.

References

- Aguilera, A. (2005). Growth in commuting distances in French polycentric metropolitan areas: Paris, Lyon and Marseille. *Urban Studies*, 42(9):1537–1547.
- Aguilera, A. and Mignot, D. (2004). Urban sprawl, polycentrism and commuting. A comparison of seven French urban areas. *Urban Public Economics Review*, 1:93–114.
- Alonso, W. (1964). *Location and Land Use. Toward a General Theory of Land Rent*. Cambridge, MA: Harvard University Press.
- Anas, A. and Moses, L. N. (1979). Mode choice, transport structure and urban land use. *Journal of Urban Economics*, 6(2):228–246.
- Bairoch, P. (1988). *Cities and Economic Development: From the Dawn of History to the Present*. University of Chicago Press.
- Banerjee, A., Duflo, E., and Qian, N. (2012). On the road: Access to transportation infrastructure and economic growth in China. Working Paper 17897, National Bureau of Economic Research.
- Baum-Snow, N. (2007a). Did highways cause suburbanization? *The Quarterly Journal of Economics*, 122(2):775–805.
- Baum-Snow, N. (2007b). Did highways cause suburbanization? *The Quarterly Journal of Economics*, 122(2):775–805.
- Baum-Snow, N. and Kahn, M. E. (2000). The effects of new public projects to expand urban rail transit. *Journal of Public Economics*, 77(2):241–263.

- Baum-Snow, N. and Kahn, M. E. (2005). Effects of urban rail transit expansions: Evidence from sixteen cities, 1970-2000. *Brookings-Wharton Papers on Urban Affairs*, pages 147–206.
- Bowes, D. R. and Ihlanfeldt, K. R. (2001). Identifying the impacts of rail transit stations on residential property values. *Journal of Urban Economics*, 50(1):1 – 25.
- Carlino, G. A. and Saiz, A. (2008). City beautiful. Working Paper 08–22, Federal Reserve Bank of Philadelphia.
- Chatman, D. G. and Noland, R. B. (2014). Transit service, physical agglomeration and productivity in us metropolitan areas. *Urban Studies*, 51(5):917–937.
- Chen, Y. and Whalley, A. (2012). Green infrastructure: The effects of urban rail transit on air quality. *American Economic Journal: Economic Policy*, 4(1):58–97.
- Combes, P.-P., Duranton, G., Gobillon, L., Puga, D., and Roux, S. (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, 80(6):2543–2594.
- Craig, S. and Ng, P. T. (2001). Using quantile smoothing splines to identify employment subcenters in a multicentric urban area. *Journal of Urban Economics*, 49(1):100–120.
- Craig, S. G., Kohlhase, J. E., and Perdue, A. W. (2016). Empirical polycentricity: The complex relationship between employment centers. *Journal of Regional Science*, 56(1):25–52.
- Diao, M., Zhu, Y., and Zhu, J. (2016). Intra-city access to inter-city transport nodes: The implications of high-speed-rail station locations for the urban development of Chinese cities. *Urban Studies*, Forthcoming.
- Donaldson, D. (2015). Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review*, forthcoming.

- Duranton, G. and Puga, D. (2015). Urban land use. In Duranton, G., Henderson, J. V., and Strange, W. C., editors, *Handbook of Regional and Urban Economics Volume 5*, pages 467–560. Amsterdam: North-Holland.
- Duranton, G. and Turner, M. A. (2011). The fundamental law of road congestion: Evidence from US cities. *American Economic Review*, 101(6):2616–52.
- Duranton, G. and Turner, M. A. (2012). Urban growth and transportation. *The Review of Economic Studies*, 79(4):1407–1440.
- Faber, B. (2014). Trade integration, market size, and industrialization: Evidence from China’s National Trunk Highway System. *Review of Economic Studies*, 81:1046–1070.
- Garcia-López, M.-A. (2010). Population suburbanization in Barcelona, 1991-2005: Is its spatial structure changing? *Journal of Housing Economics*, 19(2):119–132.
- Garcia-López, M.-A., Holl, A., and Viladecans-Marsal, E. (2015a). Suburbanization and highways in Spain when the Romans and the Bourbons still shape its cities. *Journal of Urban Economics*, 85(0):52 – 67.
- Garcia-López, M.-A., Pasidis, I., and Viladecans-Marsal, E. (2015b). Express delivery to the suburbs? Transportation effects in heterogeneous European cities. Mimeo.
- Garcia-López, M.-A. (2012). Urban spatial structure, suburbanization and transportation in Barcelona. *Journal of Urban Economics*, 72(2–3):176–190.
- Gibbons, S. and Machin, S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics*, 57(1):148–169.
- Giuliano, G., Redfearn, C., Agarwal, A., Li, C., and Zhuang, D. (2007). Employment concentrations in Los Angeles, 1980?2000. *Environment and Planning A*, 39(12):2935–2957.

- Giuliano, G. and Small, K. A. (1991). Subcenters in the Los Angeles region. *Regional Science and Urban Economics*, 21(2):163 – 182.
- Glaeser, E. L. and Gottlieb, J. D. (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the United States. *Journal of Economic Literature*, 47(4):983–1028.
- Glaeser, E. L., Kolko, J., and Saiz, A. (2001). Consumer city. *Journal of Economic Geography*, 1(1):27–50.
- Grube-Cavers, A. and Patterson, Z. (2015). Urban rapid rail transit and gentrification in canadian urban centres: A survival analysis approach. *Urban Studies*, 52(1):178–194.
- Hsu, W.-T. and Zhang, H. (2014). The fundamental law of highway congestion revisited: Evidence from national expressways in Japan. *Journal of Urban Economics*, 81(C):65–76.
- Hurst, N. B. and West, S. E. (2014). Public transit and urban redevelopment: The effect of light rail transit on land use in Minneapolis, Minnesota. *Regional Science and Urban Economics*, 46(C):57–72.
- Jiwattanakulpaisarn, P., Noland, R. B., Graham, D. J., and Polak, J. W. (2009). Highway infrastructure investment and county employment growth: A dynamic panel regression analysis. *Journal of Regional Science*, 49:263–286.
- Kotavaara, O., Antikainen, H., and Rusanen, J. (2011). Population change and accessibility by road and rail networks: GIS and statistical approach to finland 1970–2007. *Journal of Transport Geography*, 19(4):926–935.
- Li, S., Yang, J., Qin, P., and Chonabayashi, S. (2016). Wheels of fortune: Subway expansion and property values in Beijing. *Journal of Regional Science*, Forthcoming.
- Martí-Henneberg, J. (2013). European integration and national models for railway networks (1840–2010). *Journal of Transport Geography*, 26(0):126–138.

- Mayer, T. and Trévien, C. (2015). The impacts of urban public transportation: Evidence from the Paris metropolitan area. Working Paper G2015/03, INSEE.
- McDonald, J. F. (1987). The identification of urban employment subcenters. *Journal of Urban Economics*, 21(2):242–258.
- McDonald, J. F. and Prather, P. J. (1994). Suburban employment centers: The case of Chicago. *Urban Studies*, 31(2):201–218.
- McMillen, D. P. (2001). Nonparametric employment subcenter identification. *Journal of Urban Economics*, 50(3):448 – 473.
- McMillen, D. P. (2003). Identifying sub-centres using contiguity matrices. *Urban Studies*, 40(1):57–69.
- McMillen, D. P. and Redfearn, C. L. (2010). Estimation And Hypothesis Testing For Non-parametric Hedonic House Price Functions. *Journal of Regional Science*, 50(3):712–733.
- 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.
- Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *American Economic Review*, 57:197–210.
- Moretti, E. (2004). Workers’ education, spillovers, and productivity: Evidence from plant-level production functions. *American Economic Review*, 94(3):656–690.
- Muniz, I., Galindo, A., and García, M.-À. (2003). Cubic spline population density functions and satellite city delimitation: The case of Barcelona. *Urban Studies*, 40(7):1303–1321.
- Muth, R. F. (1969). *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*. Chicago, IL: University of Chicago Press.

- Muñiz, I., Garcia-López, M.-À., and Galindo, A. (2008). The effect of employment sub-centres on population density in Barcelona. *Urban Studies*, 45(3):627–649.
- Redfearn, C. L. (2007). The topography of metropolitan employment: Identifying centers of employment in a polycentric urban area. *Journal of Urban Economics*, 61(3):519–541.
- Redfearn, C. L. (2009). How informative are average effects? Hedonic regression and amenity capitalization in complex urban housing markets. *Regional Science and Urban Economics*, 39(3):297–306.
- Rosenthal, S. and Strange, W. (2008). The attenuation of human capital spillovers. *Journal of Urban Economics*, 64(2):373–389.
- Ryan, S. (2005). The value of access to highways and light rail transit: Evidence for industrial and office firms. *Urban Studies*, 42(4):751–764.
- Schuetz, J. (2015). Do rail transit stations encourage neighbourhood retail activity? *Urban Studies*, 52(14):2699–2723.
- Sivitanidou, R. (1996). Do office-commercial firms value access to service employment centers? A hedonic value analysis within polycentric Los Angeles. *Journal of Urban Economics*, 40(2):125–149.
- Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear IV regression. In Andrews, D. W. and Stock, J. H., editors, *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, pages 80–108. Cambridge: Cambridge University Press.

APPENDIX

A Employment subcenters in Paris metropolitan area

We identify employment subcenters using the method first developed by [McDonald and Prather \(1994\)](#) and improved by [McMillen \(2001\)](#). The principal idea is to estimate densities following a monocentric spatial pattern. The predicted densities obtained are subtracted from the corresponding real densities. From these residuals, those that are positive are chosen, and from these, those that are statistically significant are selected.

While [McDonald and Prather \(1994\)](#) estimate by OLS a two-dimensional density function, the log of employment density vs. the distance to CBD, [McMillen \(2001\)](#) proposes a three-dimensional density function, the log of employment density versus the north-south and the east-west distances to CBD, and uses a nonparametric estimation technique, known as locally weighted regression (LWR). Both improvements allow us to take into account geographical differences, which, in terms of the spatial pattern of densities, can occur in any direction from the CBD (e.g., steeper density gradients on the north side than on the south side of the city). Furthermore, they also allow us to define any type of monocentric spatial pattern: concave, convex or linear ([McMillen, 2001](#)).

As a result, we first estimate the following employment density equation (A.1) through LWR with a window size or bandwidth of 0.5, i.e., based on a tricube function, the nearest of the 50 percent observations receive weight ([McMillen, 2001](#)):

$$\begin{aligned} \ln(\text{Employment density}) = & \alpha_0 + \alpha_2 \times \text{north-south distance to CBD} \\ & + \alpha_3 \times \text{east-west distance to CBD,} \end{aligned} \tag{A.1}$$

where density is measured as jobs (respectively, inhabitants) per hectare, and distances are in kilometers. The CBD is defined as the 20 *arrondissements* that make up the city of Paris. Distance to CBD is the distance to the centroid of the 4th *arrondissement* (*de l'Hôtel-de-Ville*).

Second, for each site i we compute the residual as the difference between real employ-

ment density and estimated employment density, and select the ones that are significantly greater than 0 at the 10 percent level:

$$\frac{\ln(\text{Employment density}) - \ln(\widehat{\text{Employment density}})}{\widehat{\text{Standard error}_i}} > 1.64. \quad (\text{A.2})$$

Finally, we group the selected sites in subcenters when they are contiguous. We use a 'queen' criterion for contiguity: two sites (municipalities) are contiguous if they share at least one point in their boundaries. See [McMillen \(2001, 2003\)](#) and [Garcia-López \(2010\)](#) for further details on this procedure.

We apply this methodology to identify subcenters in 1968 and in 2010 in the Paris metropolitan area. In 1968, we identify 21 employment subcenters, comprising 88 municipalities, in addition to the CBD of Paris; in 2010, we identify 34 subcenters comprising 89 municipalities surrounding the CBD. At both dates, about two thirds of the subcenters identified constitute in fact a single municipality: there were six subcenters including several municipalities in 1968 and eleven in 2010. It is however worth noting that the inner suburbs of Paris (the municipalities immediately surrounding the CBD) constituted a macro-subcenter containing 63 municipalities in 1968 and 37 in 2010. Overall, 117 municipalities of the metropolitan area belonged to an employment subcenter at least once in 1968 or 2010. Among these, 28 belonged to a subcenter in 1968 only, 29 in 2010 only, and the remaining 60 were part of a subcenter at both dates.

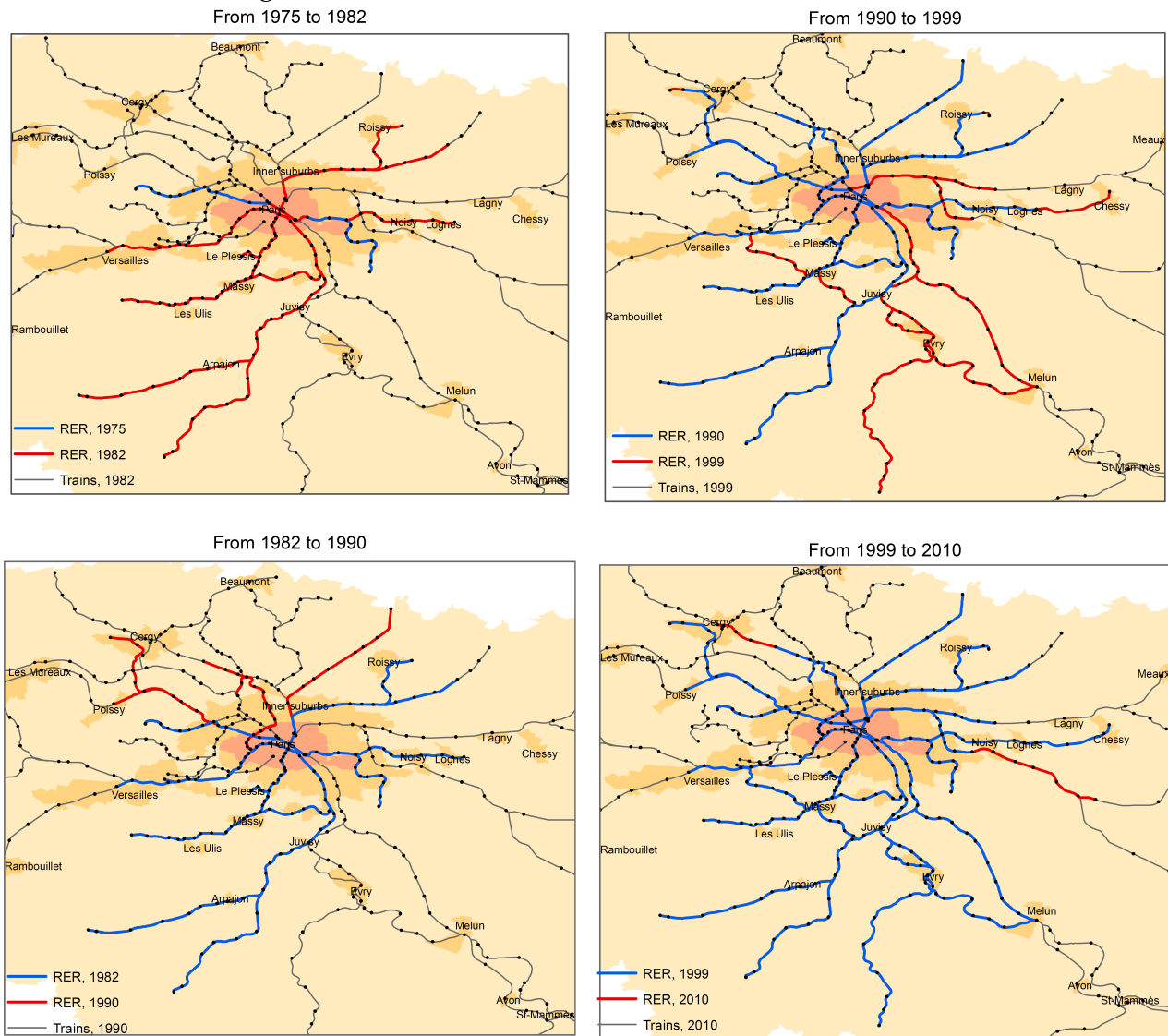
B Transportation infrastructures in metropolitan Paris

Table B.1: The evolution of transportation infrastructures in metropolitan Paris, 1968–2010

	Year [1]	Stations/Ramps [2]	Stations x Lines [3]	Lines [4]	Municipalities [5]	Length (km) [6]
Panel A: Train	1968	277	281	5	234	870
	1975	274	278	5	233	860
	1982	272	276	5	232	873
	1990	265	269	5	225	584
	1999	231	240	5	198	779
	2010	231	239	5	196	788
Panel B: RER	1968	0	0	0	0	0
	1975	22	22	1	16	39
	1982	126	129	4	84	266
	1990	159	165	4	107	358
	1999	231	240	5	158	562
	2010	243	252	5	167	587
Panel C: Métro	1968	265	338	15	33	164
	1975	273	348	15	36	173
	1982	285	360	15	41	188
	1990	291	366	15	44	196
	1999	296	376	17	46	204
	2010	299	380	17	46	208
Panel D: Tramway	1968	0	0	0	0	0
	1975	0	0	0	0	0
	1982	0	0	0	0	0
	1990	0	0	0	0	0
	1999	34	34	2	9	20
	2010	70	71	4	19	40
Panel E: Highway	1968	46		11	40	229
	1975	86		24	73	418
	1982	111		27	91	549
	1990	129		31	101	633
	1999	161		40	127	792
	2010	168		41	133	821

Notes: For railway infrastructures (Panels A to D), Column 1 reports the total number of stations in which the corresponding type of train stops, Column 2 reports the number of stations weighted by the number of lines (if two lines go through the same station then the station counts as 2), Column 3 reports the total number of lines composing the corresponding network, Column 4 reports the number of municipalities in which there is at least one station of the corresponding network, and Column 5 reports the length of the corresponding railway network (note that if the same railway is used for several lines, its length is counted only once). Information for highways, or more precisely for the main roads (including some roads smaller than highways) is reported in Panel E as follows. Column 1: number of ramps to access the highway. Column 3: total number of roads with a different label composing the highway network. Column 4: number of municipalities in which there is at least one ramp. Column 5: total length of the highway network.

Figure B.1: Evolution of the RER network, 1975–2010

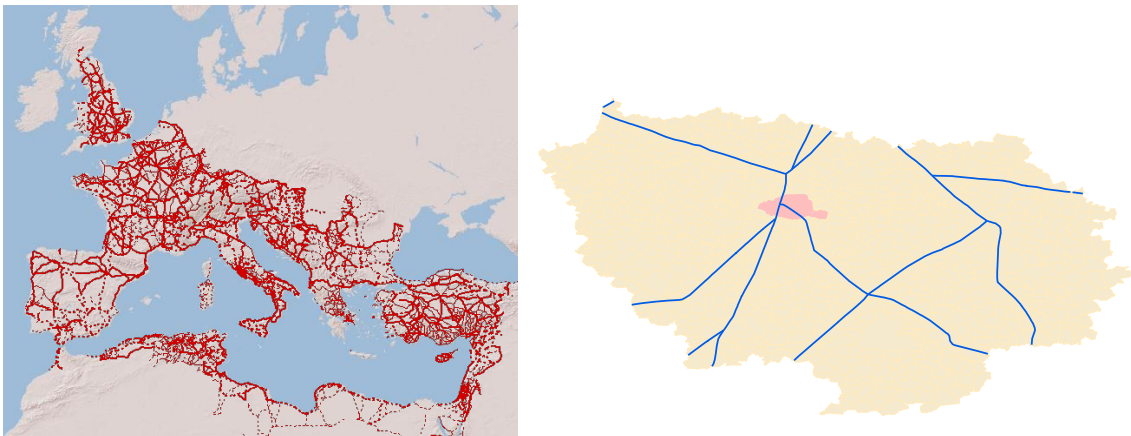


C When past infrastructures shape modern infrastructures

We explore the validity in terms of exogeneity and relevance of three candidates to instrument highways and railroads in Paris metropolitan area: Roman roads, 1810 post routes, and 1870 railroads. We do not consider more modern infrastructures such as railroads and roads in the immediate wake of World War II because, despite being more relevant, they hardly comply with the exogeneity condition and, in particular, with the exclusion restriction. Indeed, while the relevance of historical instruments decreases with time (because, for example, some Roman roads disappeared or were rebuilt as medieval roads or as post routes or railroads during the 19th century), their exogeneity (conditional on historical and geographical factors) increases with it.

The first candidate we consider consists of Roman roads. The two main Roman roads passing through Paris (Lutetia at the time) were built around 115 B.C. As a whole, our first candidate, the Roman network around Paris, was based on 526 km of roads (Figure C.1).

Figure C.1: The Roman roads

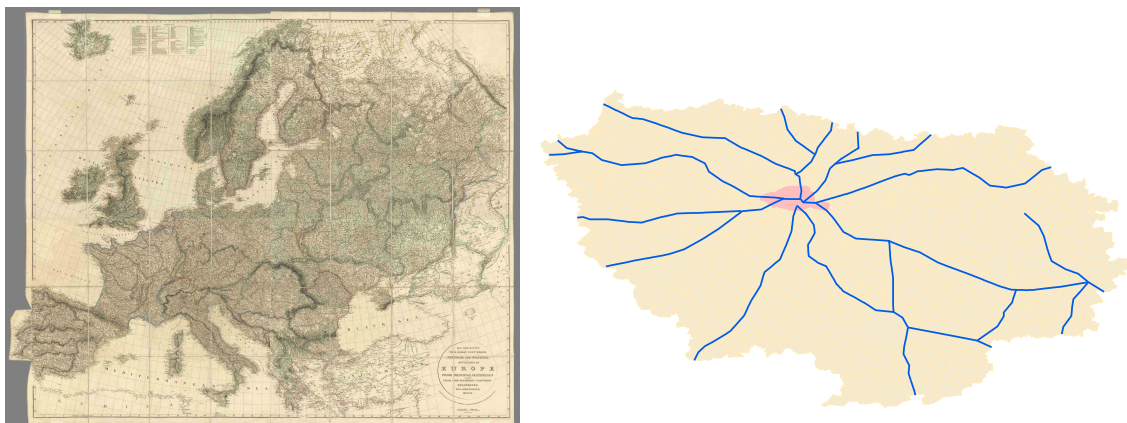


Source: Own elaboration based on the [Digital Atlas of Roman and Medieval Civilizations](#).

Our second proposed instrument is the 1810 post routes. The first fixed relay posts were established in France at the beginning of the 16th century. Over the 18th century, French engineers of roads and bridges built paved roads that connected the various relay posts. Because the primary purpose of this network was royal communication, it was

star-shaped over all the territory, with Paris at the center, and much denser in the region around Paris than in the rest of France. As a whole, this network was based on 768 km of post roads crossing the modern Paris metropolitan area (Figure C.2).

Figure C.2: The 1810 post routes



Source: Own elaboration based on digital images of an 1810 Aaron Arrowsmith map from the [David Rumsey Historical Map Collection](#).

Figure C.3: The 1870 railroads



Source: Own elaboration based on [Martí-Henneberg \(2013\)](#) maps.

The third candidate is simply the railroad network as it existed in 1870 in the region around Paris. The first line connecting Paris to a city located 18 km away (Saint-Germain) started operating in 1837. Due to the high levels of centralization in France, the 1870

railroad network (Figure C.3) also had a star-shaped form centered around Paris and was based on 698 km of railroad lines.

The fact that modern roads and railroads were built following the routes marked out by past infrastructure has been frequently pointed out in the literature. Common sense would suggest that in France as well, past infrastructures shaped current ones for the same practical reasons: it was easier and cheaper when building new transportation infrastructure to improve existing infrastructure, or to build it nearby.

In order to test the credibility of this assumption empirically in the context of the Paris metropolitan area, we regress the distance to the nearest transportation infrastructure in 2010 on the distance to the nearest historical transportation infrastructure, controlling for the urban spatial structure (with the distance to the nearest 2010 employment center),⁹ and, most importantly, geography and history:

$$\begin{aligned} \text{Dist to 2010 transportation} = & \gamma_0 + \gamma_1 \times \text{dist to historical transportation} \\ & + \gamma_2 \times \text{dist to the nearest 2010 empl center} \quad (\text{C.1}) \\ & + \sum_i (\gamma_{3,i} \times \text{geography}_i) + \sum_i (\gamma_{4,i} \times \text{history}_i) \end{aligned}$$

The inclusion of these control variables is key to our identification strategy. Although ancient transportation infrastructure may be exogenous (because of the significant changes undergone by society and economy over the long period since it was built, and because this infrastructure was not built to anticipate employment and population growth in a distant future), other factors such as the area's geography are likely to have influenced the construction and location of both ancient and modern transportation infrastructures on the grounds of the feasibility and convenience of construction. It is therefore crucial to include geographical characteristics such as land area, altitude, index of terrain ruggedness,

⁹We use this variable instead of separate distances for CBD and subcenters because the latter are highly correlated with the different transportation distances. In particular, partial correlations between distance to the nearest RER station and distance to CBD and to the nearest 2010 employment subcenters are 85 percent and 81 percent, respectively.

and elevation range as controls to comply with the exogeneity condition. It is equally important to control for the historical context, since this may explain both the presence of former infrastructure and the economic importance of today's municipalities. In order to fulfill the exclusion restriction, and because there are no historical employment and population data at the municipal level prior to 1962, we control for history by including dummy variables indicating (1) whether municipalities were Roman settlements, (2) whether they used to be major towns between the 10th and the 15th centuries and (3) between the 16th and the 19th centuries, (4) whether they had a monastery built between the 12th and 16th centuries, and (5) whether they hosted important fairs between the 10th and the 16th centuries.¹⁰ We thus assume *conditional* exogeneity of the proposed instruments, as suggested by [Duranton and Turner \(2012\)](#).

The first-stage results corresponding to equation (C.1) are reported in Table C.1. columns 1 to 4 of Panel A display the results for the RER: the distance to the nearest RER station in 2010 is very highly correlated with the distance to the nearest railroad in 1870 (column 1) and with the distance to the nearest Roman road (column 3). The values of the corresponding F-statistics confirm the strength of these instruments ([Stock and Yogo, 2005](#)). Furthermore, the F-statistic and the overidentification p-value reported in column 4 confirm that both instruments can be used simultaneously. Regarding commuter train, columns 5 to 7 of Panel A show that all three historical networks individually matter for the location of train stations. However, there is an overidentification problem when we use them simultaneously (column 8). In columns 9 to 11, we test all possible pairs of the three instruments. We find that only the 1870 railroad and Roman road combination (column 10) passes the overidentification test.

Panel B shows the results for all non-RER railroads, that is, for the distance to the nearest commuter train, métro or tramway station. Once again, we find a significant effect for each historical network (columns 5 to 7) and for the joint estimate (column 8), but

¹⁰These variables come from the [Digital Atlas of Roman and Medieval Civilizations](#), with the exception of the major cities of the 16th to 19th centuries which are identified in [Bairoch \(1988\)](#).

the latter does not pass the overidentification test. According to the F-statistics and the overidentification p-values, the distance to the nearest 1870 railroad and the distance to the nearest Roman road is our preferred combination (column 10).

Table C.1: Modern transportation as a function of past transportation, OLS

Dependent variable:	2010 Distance to the nearest										
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
Panel A: RER and commuter train											
	RER station				commuter train station						
Distance to 1870 railroads	0.961 ^a (0.079)			0.853 ^a (0.073)	0.187 ^a (0.039)			0.168 ^a (0.036)	0.182 ^a (0.038)	0.177 ^a (0.038)	
Distance to 1810 roads		-0.101 (0.105)				0.136 ^a (0.041)		0.169 ^a (0.043)	0.127 ^a (0.042)		0.178 ^a (0.042)
Distance to Roman roads			0.616 ^a (0.036)	0.574 ^a (0.036)			-0.114 ^a (0.013)	-0.124 ^a (0.013)		-0.110 ^a (0.013)	-0.129 ^a (0.013)
Adjusted R ²	0.24	0.17	0.28	0.34	0.40	0.39	0.41	0.44	0.41	0.42	0.42
First-stage statistic	147.06	0.92	300.01	231.26	23.11	10.90	74.11	48.14	16.29	54.98	50.39
Overidentification p-value				0.58				0.00	0.00	0.27	0.01
Panel B: all non-RER railroads → commuter train, subway and tramway											
	non-RER station										
Distance to 1870 railroads					0.171 ^a (0.038)			0.152 ^a (0.036)	0.165 ^a (0.037)	0.161 ^a (0.037)	
Distance to 1810 roads						0.141 ^a (0.041)		0.172 ^a (0.043)	0.133 ^a (0.041)		0.181 ^a (0.042)
Distance to Roman roads							-0.108 ^a (0.013)	-0.119 ^a (0.013)		-0.104 ^a (0.013)	-0.123 ^a (0.013)
Adjusted R ²					0.42	0.42	0.43	0.45	0.43	0.44	0.44
First-stage statistic					19.75	11.77	66.92	44.06	15.08	48.98	47.00
Overidentification p-value								0.00	0.00	0.29	0.02
Panel C: Highways and combination with non-RER railroads											
	highway ramp				non-RER station or highway ramp						
Distance to 1870 railroads	0.115 ^b (0.048)			0.100 ^b (0.046)	0.149 ^a (0.038)			0.125 ^a (0.036)	0.141 ^a (0.037)	0.137 ^a (0.037)	
Distance to 1810 roads		-0.054 (0.042)				0.076 ^b (0.032)		0.185 ^a (0.041)	0.146 ^a (0.040)		0.193 ^a (0.041)
Distance to Roman roads			-0.162 ^a (0.018)	-0.159 ^a (0.018)			-0.081 ^a (0.011)	-0.130 ^a (0.012)		-0.115 ^a (0.013)	-0.134 ^a (0.013)
Adjusted R ²	0.54	0.54	0.56	0.56	0.41	0.46	0.48	0.45	0.42	0.43	0.44
First-stage statistic	5.73	1.65	80.93	46.82	15.01	5.62	50.32	51.29	13.63	55.27	59.29
Overidentification p-value				0.80				0.00	0.00	0.52	0.00
Dist to 2010 centers	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Geography	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
History	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: 1300 observations for each regression. Geography variables are land area, altitude, index of terrain ruggedness, and elevation range. History variables are dummy variables for municipalities (1) that were Roman settlements (based on DARMC maps), (2) that were major towns between the 10th and the 15th centuries (based on DARMC maps), (3) that were major towns between the 16th and the 19th centuries (based on Bairoch, 1988), (4) with a monastery built between the 12th and 16th centuries (based on DARMC maps), and (5) that hosted important fairs between the 10th and the 16th centuries (based on DARMC maps). Robust standard errors are in parentheses. ^a, ^b, and ^c indicates significant at 1, 5, and 10 percent level, respectively.

Finally, in Panel C we explore the instruments for the distance to the nearest highway ramp (columns 1 to 4) and for the distance to all non-RER station or highway ramp (columns 5 to 11). Similar to previous panels, our preferred instruments are the distances to the nearest 1870 railroad and to the nearest Roman road (columns 4 and 10).