

Trevien, Corentin; Mayer, Thierry

**Conference Paper**

## Urban Public Transportation and Firm Location Choice Evidence from the Regional Express Rail of Paris Metropolitan area

53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Trevien, Corentin; Mayer, Thierry (2013) : Urban Public Transportation and Firm Location Choice Evidence from the Regional Express Rail of Paris Metropolitan area, 53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/123954>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Urban Public Transportation and Firm Location Choice

## Evidence from the Regional Express Rail of Paris Metropolitan area

Thierry Mayer\* and Corentin Trevien†

February 15, 2013

Preliminary version, please do not cite

### **Abstract**

This paper identifies the causal impact of urban rail transport on firm location. The evaluation of transport infrastructures always faces an important issue of endogeneity since rail lines are not located randomly. We use the natural experiment offered by the Regional Express Rail progressively opened between 1970s and 2000s in the Paris metropolitan region. We show that ignoring the endogeneity issue doubles or triples the estimated impact of rail on firm location. The regional market share of a municipality in firm location increases between 4 and 6 % with a station. However, the total number of firms grows durably only if the travel time decreases. The estimated impact is stronger on foreign firms and varies across industry. It is significant only for manufacturing, business and household service firms.

JEL Codes: *D04, H43, R42*

---

\*SciencesPo., 28 rue des Saints-Pères, 75007 Paris, France, thierry.mayer@sciences-po.fr

†Insee-SciencesPo.-Crest, 15 bd Gabriel Péri, 92240 Malakoff, France, corentin.trevien@insee.fr

# 1 Introduction

The evaluation of transportation infrastructures appears to be a major topic as the construction of roads, railway lines or airports represents a substantial public spending. At the same time, many transportation investments are planned, for example, the "Crossrail" project in London, the "Grand Paris" in France or the Obama plan for high speed rail in the United-States. Policymakers generally argue that such infrastructures are priority steps for the economic development of their region or country. This assertion has to be evaluated in the light of rigorous empirical works. This article aims to estimate the causal impact of Paris urban transportation systems on firm location. It proposes an original identification strategy based on a natural experiment provided by the implementation of a fast commuter rail system, the Regional Express Rail (RER)<sup>1</sup>. The RER has been progressively opened from the 1970s to the 1990s (see figure 1). It has currently reached 587 km all over the Paris metropolitan region. Finally, it provides empirical evidence on the way transports influence firm behavior.

The evaluation of transport systems faces an unavoidable problem of endogeneity. In fact, new infrastructures are obviously not randomly assigned. For example, urban planners are likely to connect in priority economically promising areas. They can also use transport infrastructure as a tool for urban renewal and link deprived areas. The consequence is that a naive estimation of the effect of a new infrastructure which would compare directly connected to unconnected areas would be biased. The sign of this bias remains unknown as the intentions and anticipations of the urban planners are not observed.

The literature offers some example of identification strategies that can address this major endogeneity issue and provide the unbiased effect of new infrastructures. Econometricians have used natural experiments to identify the causal effect of transport infrastructures. Duranton and Turner (2012) evaluate the impact of the highway network in the United-States on the local evolution of employment. They use an instrumental variable strategy, based upon a 1947 plan of the interstate highway system and a 1898 map of railroads, to address the endogeneity of the highways location in 1980. Michaels (2008) also uses the 1947 plan as an exogenous variation of road for trade issues. Donaldson (2010) shows that the Indian railway extension led to a decrease in interregional trade costs and increases both incomes and trade. To do so, he uses a natural experiment provided by 40,000 km of planned lines which were never built for exogenous reasons. Banerjee et al. (2012) find a moderate positive effect of transportation access on income growth in China. They use the fact that railroad lines were built in China to connect European concessions on the coast and inland historical cities in the 19th century. They bring out that crossed areas, which were located between these two sorts of cities were "quasi-randomly" linked to the railway network and can be compared to similar unconnected areas.

On the other hand, other papers study the determinants of firm location, including transportation, but they do not use a natural experiment to address the endogeneity issue of transport infrastructure location. All these works highlight a positive impact of roads (Coughlin and Segev, 2000; Gibbons et al., 2012; Holl, 2004a,b), rail (Cheng and Kwan, 2000; Kang and Lee, 2007) or airports (Strauss-Kahn and Vives, 2009) on firm location. Data availability and precision is a key issue to accurately estimate the impact of transportation on firms. Gibbons et al. (2012) insist on the fact that such an evaluation is more complicated in developed countries as transportation networks are already widely extended. In this conditions, it is necessary to measure not only if a

---

<sup>1</sup>Réseau express régional in French

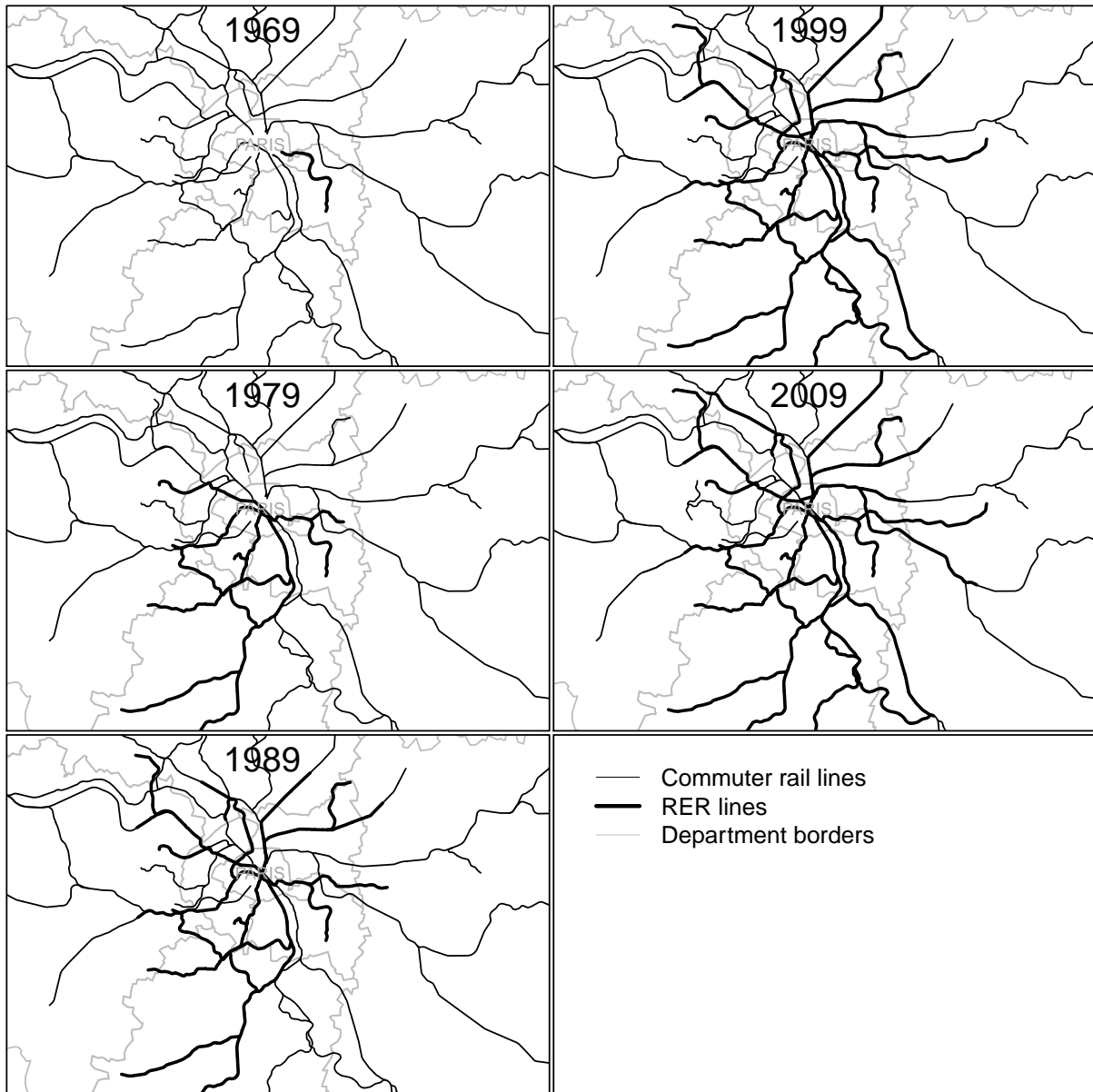


Figure 1: Progressive extension of the Regional Express Rail between 1969 and 2009

given area is linked to a network but also how well it is connected. To do so, it is necessary to have journey time data.

This article combines a rigorous approach of transport infrastructure evaluation with a standard frame of firm location choice. More concretely, we consider both domestic and foreign firm location choice across municipalities of the Paris metropolitan region. In addition to basic criterions as market potential, agglomeration, labor costs and land rent, we add the proximity to a commuter train station and an accessibility index based on travel time by public transport. We provide two strategies to address the endogeneity issue we previously reported. The first one is similar to Banerjee et al. (2012). The RER network has been implemented with the aim of connecting "new towns" located 30 km away from the historic center of the city. Commuter train crosses

municipalities located between the historical core and the five new towns of the metropolitan area. These municipalities are "quasi-randomly" treated and can be compared to similar untreated municipalities in order to estimate the causal impact of railway infrastructures. The second strategy is closer to Duranton and Turner (2012) and relies on significant differences between the initial plan presented in 1965 and the current network.

We show that ignoring the endogeneity issue quadruple the estimated impact of the infrastructure on the market share in firm location of a municipality. The impact of commuter train is significant but moderate as the opening of a RER station causes a 3 percent increase of firm location. In addition, the quality of the connection to railroads matters: the more the reduction of time associated with the opening of a new station is high the more the effect of transportation on firm location is important. The impact of the RER is very strong for foreign firm: FDI location increases by 18 % with the opening of a RER station. Finally, the effect varies across industry sector and turns out to be higher for business services and household services.

The first section of this article details the model and the estimation strategy. Then, the second section presents the datasets used for estimations. The third section sets out the results and finally, the fourth section concludes and discusses our results.

## 2 Identification and estimation

### 2.1 Theoretical grounds

The evaluation of transportation infrastructure effects on firm location is based on a simple location choice model (Head and Mayer, 2004). Each firm  $i$  computed its expected profit  $\pi_c$  in each possible location  $j$  and chooses the location which maximizes its profit. In each municipality, the profit depends of the market potential  $MP_j$  and of the fixed and variable costs  $c_j$  and  $F_j$ . The costs include wages, land rents, intermediate goods and services.

$$\pi_j = \frac{c_j^{1-\sigma}}{\sigma} MP_j - F_j,$$

The market potential designates the sale level the firm can reach in the location  $j$  and depends on the consumption spending  $E_j$  in the other locations  $k$ , the accessibility  $\phi_{jk}$  of the locations  $k$  from the location  $j$  and the competition  $G_j$  the firm faces in the location  $k$ .

$$MP_j = \sum_k \frac{\phi_{jk} E_k}{G_k}.$$

Transportation improvements are likely to affect this profit equation through multiple channels. On the one hand, it may facilitate access to labor force, location in municipalities where land is cheaper and access to customers and suppliers. On the other hand it should also strengthen competition. Consequently, the resulting impact of transportation remains unclear but positive forces seem to be stronger.

### 2.2 Identification strategy

Our identification strategy must address a major endogeneity issue, in order to provide the unbiased effect of the RER on firm location. The main problem of transportation system evalu-

ation is that new infrastructures are obviously not randomly located. The consequence is that a naive estimation which would directly compare connected to unconnected areas will be biased. As explained in the introduction, the sign of this bias remains unknown.

Our method relies on two strategies. The first is similar to Banerjee et al. (2012) and compares existing lines which were converted in RER lines with other unconverted lines. The second one uses the differences between the 1965 initial urban master plan and the commuter train network actually implemented as Duranton and Turner (2012). For each strategy, we exhibit a control group and a treatment group. We argue that location of transport infrastructures can be seen as "quasi-random" across groups.

In the 1960s, Paris metropolitan area turned out to be disorganized by the important urban growth subsequent to the Second World War. In 1965, the French Government presented a urban master plan to reorganize and frame the development of the agglomeration. This plan is called the "SDAURP" <sup>2</sup>. It included a reorganization of administrative boundaries, the construction of nine "new towns", 30 km away from Paris, and the building of new transport infrastructures. The RER, a rapid and high capacity commuter rail network was one of these projects. It was supposed to link the city of Paris, which is limited to the historical core of the agglomeration, to the major centers of the Paris metropolitan area, and especially the nine "new towns" and the two airports. This planned network was very ambitious and required the construction of hundreds of kilometers of news railways across the agglomeration.

Because of the 1970s crisis, the implementation of the initial urban plan has been more modest. Only five news towns were actually built and an important part of the transport infrastructures initially planned has been abandoned. Finally, the Regional Express Rail has consisted in an improvement of existing railroad lines. New line construction was rarer than initially planned (see figures 4 and 5).

A brief history of the railway network in France has to be exposed to understand properly the importance of the RER for the organization of the commuter train network. The French railways have been mostly built during the 19th century by private companies which obtained concessions of the French Government to link French cities to Paris. This organization has had a long-lasting impact on the geographical shape of the railway network. First, it is largely centralized and the majority of lines is directed towards Paris. Circular tracks are scarcer than radial ones. This observation is also valid for the Paris metropolitan area. Second, each company covered a specific geographical area and the competing networks were poorly connected. More precisely, each company built a specific dead-head station in Paris. Both commuters and mainline passengers had to change trains twice in Paris. Concretely, consider the case of someone who lives in the north and works in the south of the Paris metropolitan area. Before the RER implementation, he or she had to take a train to a first train station, then to use the metro to reach the second train station and finally take another train to his final destination. The organization of the commuter rail network was clearly suboptimal. In addition, some lines were still served by steam trains in the late 1960s.

The major improvement of the RER network has been to link isolated railway stations by four tunnels under Paris (see figure 5). It also includes the construction of new branch lines towards airports and "new towns" not connected by an existing line, the commissioning of new trains and higher frequencies. Despite limited new track segments, the RER finally led to a significant

---

<sup>2</sup>Schéma directeur d'aménagement et d'urbanisme de la Région Parisienne in French

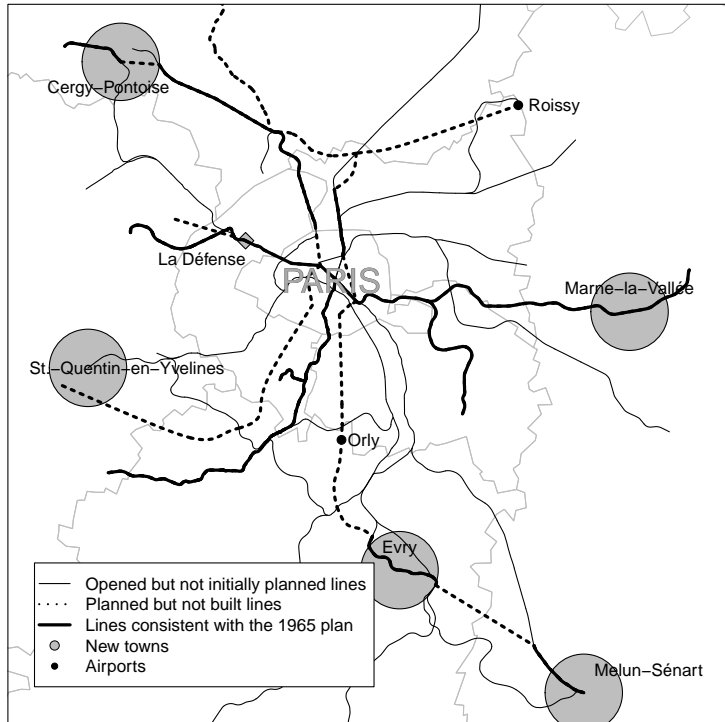


Figure 2: Comparison of the 1965 master plan and the current Regional Express Rail network

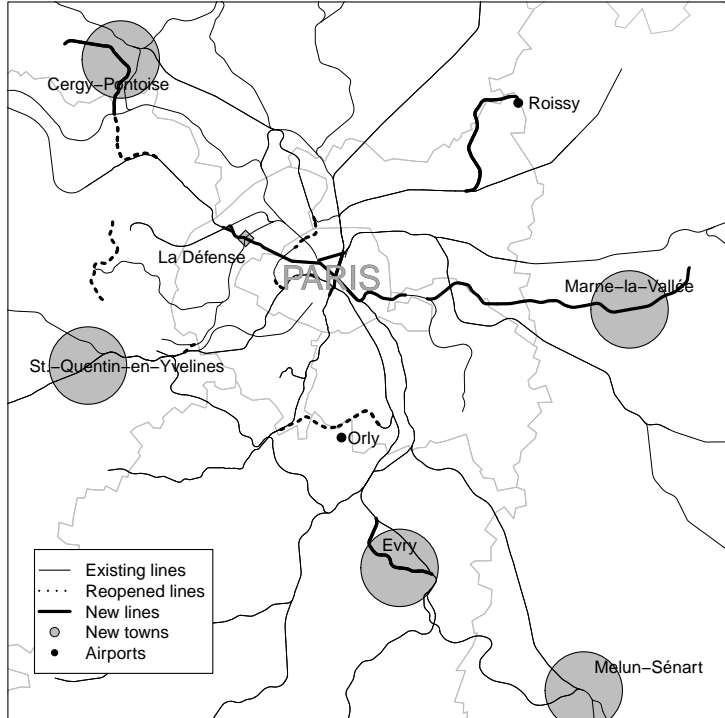


Figure 3: New and reopened lines in Paris metropolitan region between 1975 and 2005

improvement of the commuter rail network. It also reaches the goals assigned by the initial urban plan to connect the five new towns and both airports. Coming back to the previous example, the journey from home to work is easier with the Regional Express Rail. Only one train change is needed, in one of the new underground RER stations in the city center, instead of two before the opening of the RER service. The aim of our work is to determine the effect of these major improvements on firm and employment location. The RER policy offers two natural experiments that can be used to isolate the causal effect of urban transportation on firms.

The first strategy focuses on intermediate cities. Existing commuter rail stations in Paris metropolitan area have been upgraded to RER stations only if they were located between the historical city center and new economic centers (new towns, airports, business district). Other stations were left unchanged and are still served by commuter trains (see figure 1 for general case). The treatment can be considered as "quasi random" for the municipalities with intermediate RER station since the new network aimed at connecting new economic centers to Paris and not specifically these municipalities. Conversely, the control group contains municipalities with a commuter rail station out of the RER network. Termini stations are logically excluded from both groups as they are explicitly targeted by the RER policy. The treatment is clearly not exogenous in that case. By termini station, we mean the historic city of Paris and municipalities which are part of a "new town", host an airport or the business district of "La Défense". Another reason for excluding these municipalities is the impossibility to find a counterfactual. Indeed, there is no new town, airport, historic city center or business district which is not connected to the RER in the Paris metropolitan region. Municipalities connected to the underground network are also removed not to interfere with the effect of the RER. In addition, the sample is restricted to municipalities which are not further than 30 km from Paris, which is the average distance between the historical city core and the new towns. In fact, the Regional Express Rail serves a larger area than the Paris urban zone and reaches the rural fringes. We don't want to include rural municipalities in the sample because we are mainly interested in the effect of urban transportation and because rural municipalities are too different from urban municipalities.

The treatment group includes 127 municipalities and the control group contains 66 municipalities in 2004 (see figure 4 and table 1). As RER stations were gradually opened, "after" and "before" are not the same for each treated municipality. The control group tends to become smaller, while the size of the treatment group grows over time. Note that the control group municipalities are smaller in population, total job and firm number. We address this issue by computing a log-linear model with municipality fixed effects. We also present a robustness check in the next subsection based on propensity score weighting.

The second strategy relies on the differences between the 1965 initial plan and the present network. From this perspective, municipalities can be divided in three categories. The first category includes municipalities for which the 1965 plan envisaged a RER station and are currently connected to the network. The coherence between plan and realization prove they were intentionally equipped with a RER station. Consequently, they are excluded from both treatment and control groups as the RER implementation is endogenous. In the two other categories, we include, on the one hand, municipalities served by the RER but which would not be according to the 1965 master plan. On the other hand, we include municipalities which are not connected to the RER network while it would have been according to the 1965 master plan. The treatment seems more random for the two last categories: municipalities were connected for exogenous reasons. For instance, underground stations were constructed to free surface stations for mainline trains, as a consequence every commuter rail line was converted to RER; some line needed to be urgently



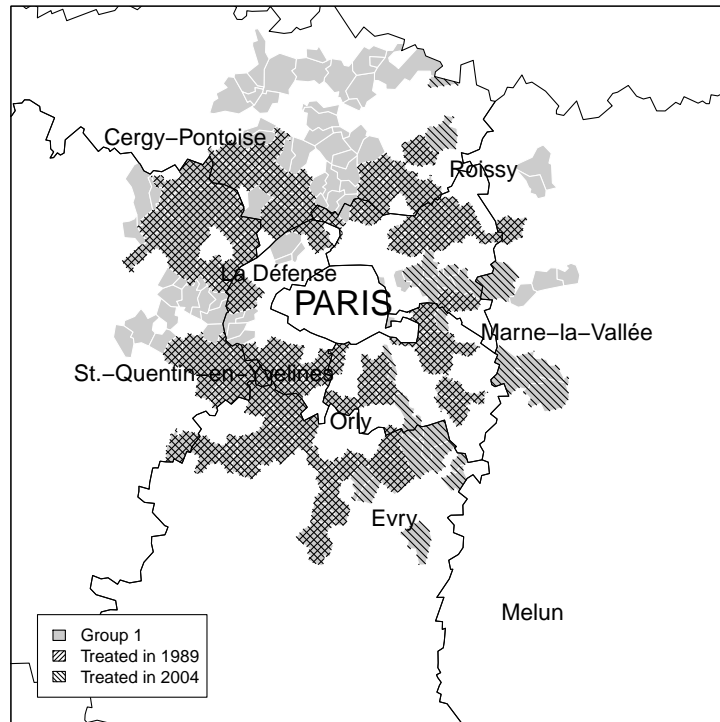


Figure 4: Municipalities selected for the first strategy

modernized, etc. (see figure 4 for general case). Note that we apply the same restriction to the sample as for the first group: only municipalities within 30 km from Paris are selected; among these, municipality served by the underground, new towns and airports are excluded. There are less municipalities selected for the second strategy (see table 1 and figure 5). The difference in population between the control and the treated group is slightly less important.

The period is split into two parts. The first period is 1974-1989, when more effective lines have been opened. The second period is 1990- 2004, when new lines brought limited service improvements. This partition is relevant in the light of our results, since the RER effect is significant before 1989 but then vanishes.

## 2.3 Estimations

The literature suggests two methods to estimate the impact of location variable on firm location choice. The first one is based on a conditional logit method and compares the probability for each firm to locate in a given location. The second one uses aggregated count data at the location level. It estimates the impact of covariates by comparing the number of firm location across locations. Our data do not contain variables varying across municipalities at the firm level. In such a case, Schmidheiny and Brülhart (2011) show that Poisson count data and conditional logit methods provide the same estimates but do not lead to the same predictions in case of a change in covariates. The consequence is that we do not need to compute estimations at the firm level. However, it has a substantial consequence on result interpretation. Conditional logit models assume a fixed number of firms. Consequently, if firm location increases in a given municipality, it necessarily implies a decrease for other municipalities. On the contrary, count data methods

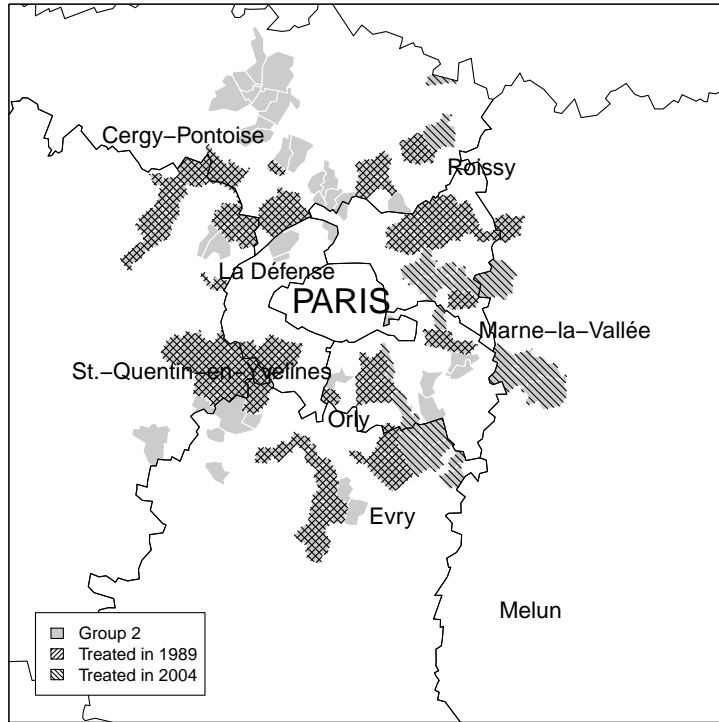


Figure 5: Municipalities selected for the second strategy

state that a firm location augmentation for a given municipality has no consequence for other municipality. It directly increases the total firm number of the region. Obviously, the reality is in between and the authors show that a nested logit model can separate the respective share of relocation and attractiveness. The estimation of such a model requires data on outside options of firms. We only have data on French firms and no information on foreign option of firms. For this reason, our method cannot isolate the effect of the relocation of firm between treated and untreated municipalities within Paris metropolitan area and the attraction of external firms which would not have located in the metropolitan area without the RER construction.

To be in line with the theoretical grounds we presented before, we estimate the effect of covariates on the probability for a firm to locate in a given municipality. The probability corresponds to the market share of a municipality, i.e. the firm location number in a given municipality, divided by the total firm location number. We estimate this model not only for all firms but also for foreign firms and break down by industry. We follow the same approach for employment. As the background model is multiplicative, we estimate it with a log-linear model. This type of model also allows to control for the various size of municipalities among groups. We add municipality fixed effects  $\alpha_i$  to eliminate the constant unobserved heterogeneity of each municipality. We add year fixed effect  $\alpha_t$  to take into account the temporal trend.

$$\log(Y_{it}) = \delta RER_{it} + \beta \log(x_{it}) + \alpha_i + \alpha_t + \epsilon_{it}$$

The Regional Express Rail improved commuter train service in many ways. That is the reason why we introduce three types of treatment variables  $RER_{i,t}$  to take into account all the improvement aspects. The first type corresponds to the presence of a Regional Express Rail station in a

Table 1: Comparison of control and treatment groups

		Strategy 1		Strategy 2	
		Control group	Treatment group	Control group	Treatment group
Group size	1974	179	14	94	.
	1989	90	103	51	43
	2004	66	127	31	63
Average number of jobs	1974	.	.	.	.
	1989	1202	3174	1462	3005
	2004	.	.	.	.
Average number of firms	1974	22	42	25	.
	1989	88	176	103	178
	2004	60	158	60	151
Average population	1974	16497	30222	17827	.
	1989	13258	25983	14740	26855
	2004	10237	26702	10853	26843

municipality. More precisely, the variable corresponds to the total surface of the municipality which is located within 500 meters of a RER station, normalized to one for a single station.<sup>3</sup> The second type of treatment variable is the average travel time by train to Paris. Between 1969 and 2009, the train travel time to Paris has been reduced in average by 4.2 minutes for municipalities connected to the RER network while it has been reduced only by 1.9 minutes for other municipalities. However, the reduction in travel time is heterogenous across municipality and may affect location choices in different ways. The third type of variable is accessibility index. It measures the accessibility  $A_{it}$  of a given municipality  $i$  at time  $t$  to a given opportunity  $X_{jt}$  (population, enterprises, etc.) in other municipalities  $j$  given the travel time by train  $d_{ij}$  between  $i$  and  $j$ . We tried three types of functions  $f$  suggested by the literature: exponential with  $f(X_{jt}, t_{ij}, \alpha) = \exp(t_{ij}^{-\alpha} X_{jt})$ , inverse  $f(X_{jt}, t_{ij}, \alpha) = t_{ij}^{-\alpha} X_{jt}$  or indicator function  $f(X_{jt}, t_{ij}, \alpha) = \mathbb{1}(d_{ij} \leq \alpha) \times X_{jt}$ . The indicator function offers a simple interpretation: it is the total amount of opportunity  $X$  reachable within  $\alpha$  minutes but the choice of  $\alpha$  is difficult. We retained the inverse function, which is the most widespread, with  $\alpha = 1$ , since it gave acceptable results.

$$A_{it} = \sum_j f(X_{jt}, t_{ij}, \alpha) \text{ with } \alpha > 0$$

As we do not have exact variables of the theoretical models, we use proxies for estimations: population density for land rent, unemployment rate for labor cost, total number of firms at January 1st of each year for agglomeration effect. The market potential is divided into two parts: the population for the internal market potential and a basic Harris function based on population of other municipalities and inter-municipality distance for the external market potential. Concretely, the second part market potential  $MP_{it}$  of a municipality  $i$  at time  $t$  is the sum of the population of all other municipalities  $j$  at time  $t$  divided by the distance as the crow flies between  $i$  and  $j$ .

$$MP_{it} = \sum_{j \neq i} \frac{P_{jt}}{d_{ij}^\alpha}$$

<sup>3</sup>The variable is divided by the surface of a disk of radius 500 meters, in such a way that it is equal to one if the whole station surroundings are located in the same municipality.

The identification of our model relies on a difference-in-differences strategy. The central assumption is that each control group would have grown in the same proportions as the treatment groups in absence of the treatment. Thanks to our identification strategies, there might be no unobserved phenomenon correlated to the treatment. In a word, the whole gap we find between control and treatment group is due to the treatment. However, the fundamental assumption of difference-in-differences could be partially rejected for another reason. In facts, the control group is likely to be affected negatively by the treatment. Concretely, many firms do not give importance to the exact location within Paris metropolitan area. They are susceptible to prefer a municipality that provides an efficient transportation system to another that does not. That is the reason why the control group could have evolved differently in absence of the commissioning of the RER. Consequently, the result we get could be overestimated.

### 3 Data description

Our datasets provide information on firm location choice, socio-demographic features of Paris metropolitan area municipalities and the evolution of urban transportation system between the 1970s and the 2000s. Municipality is the more accurate geographical scale available as more precise data are computed only since the 1990s. However, French municipalities are particularly small in comparison with other European countries and prove to be an adapted geographical scale for this type of estimations. The Paris metropolitan region contains 1300 municipalities.

Our data give information on firms between 1974 and 2004, including the industry sector, the location at a municipality level, the opening and closing years. We also have employment at the firm level between 1978 and 2003. Note that until the 1990s, foreigner investors had to register every acquisition of French firms with the treasury. A dataset of foreign direct investment (FDI) has been computed using this administrative requirement. So, we have data on shareholding including the nationality of the stakeholder and the proportion of capital held by foreigners. As this compulsory registration has been phased out in the 1990s, reliable data on FDI are available only until 1994. We use census data to know the population, the unemployment rate and the social composition of municipalities at each census year in 1968, 1975, 1982, 1990 and 1999. For the other years, we consider that variable remains stable and we add year fixed effects in our estimations to address the approximation.

We constructed new data on the transportation system of Paris metropolitan area. These data include the precise surface of each municipality within 500 meters of a train or metro station, each year between 1969 and 2009. We also computed the travel time by rail between every municipality for the same period of time. It is an essential information, as the RER conversion of existing lines led to mixed improvement of the transportation service quality. This travel time data are used to compute accessibility indexes.

### 4 Results

The table 2 shows how important it is to estimate a well-specified model not to overestimate the effect of transportation infrastructures on firms. The results of our two identification methods are reported in columns (5) and (6): a RER station opening causes an increase between 4 and 6 percent of the market share in firm location of a municipality. The two methods yield slightly different results. The first column presents the most naive way to estimate the parameter of interest.

Table 2: Effect of RER on the market share of new firm location

	Firm location					
	(1)	(2)	(3)	(4)	(5)	(6)
Presence of a RER station	0.0910*** (0.0131)	0.0800*** (0.0175)	0.0800*** (0.0167)	0.0408*** (0.0143)	0.0361* (0.0187)	0.0513*** (0.0185)
Log of unemployment rate	0.0239*** (0.00474)	0.00890* (0.00522)	0.00890 (0.00696)	-0.0127 (0.0173)	-0.00142 (0.0488)	-0.0196 (0.0351)
Log of population	-0.0113 (0.00998)					
Log of internal market potential	0.328*** (0.0131)	0.350*** (0.0200)	0.350*** (0.0688)	0.207** (0.0852)	0.103 (0.0866)	-0.0279 (0.0303)
Log of external market potential	0.247*** (0.0190)	-3.109*** (0.324)	-3.109*** (0.412)	-1.454*** (0.528)	-0.188 (0.862)	-1.015 (1.230)
Log of total number of firms	0.506*** (0.0106)	0.137*** (0.0161)	0.137*** (0.0289)	0.370*** (0.0575)	0.164** (0.0815)	0.273*** (0.101)
Observations	20800	20800	20800	8704	2400	1408
Sample	Complete	Complete	Complete	Within 30km from Paris	Group 1	Group 2
Year FE	✓	✓	✓	✓	✓	✓
Municipality FE		✓	✓	✓	✓	✓
Clustered SE			✓	✓	✓	✓
Time period	1974-1989	1974-1989	1974-1989	1974-1989	1974-1989	1974-1989

Standard errors in brackets - Significance levels: \*\*\* 1%, \*\* 5%, \* 10%

It includes all the municipalities of the Paris metropolitan region, with only year fixed effects. Ignoring endogeneity issues causes an overestimation of the interest parameter which is two or three times larger, depending on the identification strategy. Adding municipality fixed effects reduced the estimated parameter but we still obtain overestimated parameters. Finally, clustering standard errors, as recommended by Bertrand et al. (2004), slightly decreases the significance. For additional robustness checks on autocorrelation of residuals, see table ?? in appendix. We computed two alternative methods to clustered standard errors and take into account autocorrelation of residuals.

We find no evidence of a RER station opening on the total number of firms (see table 3) for both groups. It is due to the fact that new firm locations and firm closures increase together. This result could be explained by a strengthened competition which excludes firms already located from municipalities connected to the RER. Another explanation is a lower life expectancy of newly located firms. The coefficient associated with firm closures seems even greater for the first strategy. However, the difference is not significant and the result should be interpreted cautiously as the French firm register is more reliable to register new companies than terminations of business. We find no effect of RER station opening on job creation, job destruction and total number of jobs in any groups (see table 4) with one exception. The job destruction rate seems to increase with the opening of a new station for the second identification strategy but it does not cause a significant decrease in the total number of jobs.

Table 3: Effect of RER on the market share of new firm location, firm closure and total number of firms

	Firm location		Firm closure		Total number of firms	
Presence of a RER station	0.0355* (0.0189)	0.0610*** (0.0190)	0.0466** (0.0198)	0.0426 (0.0292)	0.00162 (0.0118)	0.00741 (0.0127)
Log of unemployment rate	-0.0156 (0.0359)	-0.0167 (0.0397)	-0.0535 (0.0544)	-0.0116 (0.0353)	-0.0442 (0.0412)	0.00808 (0.0129)
Log of internal market potential	-0.0197 (0.0353)	0.320** (0.136)	-0.166*** (0.0509)	-0.103 (0.126)	0.218*** (0.0339)	0.410*** (0.0442)
Log of external market potential	0.332 (0.742)	0.0656 (1.156)	-3.730*** (1.104)	-3.546** (1.460)	1.629*** (0.476)	0.258 (0.595)
Log of total number of firms	0.293*** (0.0704)	0.0743 (0.0937)	1.257*** (0.191)	0.827*** (0.143)		
Observations	3088	1504	3088	1504	3088	1504
Sample	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2
Year FE	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓
Clustered SE	✓	✓	✓	✓	✓	✓
Time period	1974-1989	1974-1989	1974-1989	1974-1989	1974-1989	1974-1989

Standard errors in brackets - Significance levels: \*\*\* 1%, \*\* 5%, \* 10%

Table 4: Effect of RER on the market share of job creation, job destruction and total number of jobs

	Job creation		Job destruction		Total number of jobs	
Presence of a RER station	0.0473 (0.0536)	0.0617 (0.0771)	0.0963 (0.0628)	0.170** (0.0728)	0.0157 (0.0334)	0.0324 (0.0472)
Log of unemployment rate	-0.0464 (0.0908)	-0.0258 (0.0819)	-0.0699 (0.127)	0.105* (0.0556)	0.0601 (0.0983)	-0.0170 (0.0508)
Log of internal market potential	-0.139 (0.302)	-0.0524 (0.243)	0.331 (0.315)	0.459** (0.213)	0.357 (0.218)	0.424*** (0.153)
Log of external market potential	-4.805** (2.324)	-4.698 (4.184)	-0.500 (2.562)	-0.578 (3.898)	2.900* (1.722)	1.713 (2.655)
Log of total number of jobs	0.909*** (0.118)	0.901*** (0.209)	-0.233** (0.100)	-0.0884 (0.110)		
Observations	2123	1034	2123	1034	2316	1128
Sample	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2
Year FE	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓
Clustered SE	✓	✓	✓	✓	✓	✓
Time period	1978-1989	1978-1989	1978-1989	1978-1989	1978-1989	1978-1989

Standard errors in brackets - Significance levels: \*\*\* 1%, \*\* 5%, \* 10%

Table 5: Effect of RER on the market share of on the market share of new foreign firm location

	Share of capital held by foreign investors over 10%		Share of capital held by foreign investors over 50%	
Presence of a RER station	0.189*** (0.0494)	0.272*** (0.0694)	0.176*** (0.0477)	0.255*** (0.0666)
Log of unemployment rate	-0.0373 (0.0351)	0.00420 (0.0192)	-0.0303 (0.0329)	-0.000993 (0.0181)
Log of internal market potential	0.0801*** (0.0272)	0.0288 (0.108)	0.0735*** (0.0225)	0.0277 (0.105)
Log of external market potential	-4.594*** (0.961)	-4.287*** (1.522)	-4.369*** (0.895)	-3.633** (1.400)
Log of foreign firm stock (< 10%)	0.113*** (0.0258)	0.0972*** (0.0342)		
Log of foreign firm stock (< 50%)			0.113*** (0.0273)	0.104*** (0.0349)
Observations	3088	1504	3088	1504
Sample	Group 1	Group 2	Group 1	Group 2
Year FE	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓
Clustered SE	✓	✓	✓	✓
Time period	1974-1989	1974-1989	1974-1989	1974-1989

Standard errors in brackets - Significance levels: \*\*\* 1%, \*\* 5%, \* 10%

The impact of a RER station opening in a municipality is very strong on foreign firms location (see table 5). The increase in the market share location lies between 18 % and 23 % depending on the identification strategy and the share of the capital held by foreign investors.

The effect of RER is very different before and after 1990. The opening of a new station results in a market share increase of firm location only before 1990. On the contrary, we find no significant impact of RER after 1990. This result can be easily explained by the characteristics of rail lines opened after and before this date. The first two RER lines, opened in the 1970s and 1980s, (A and B) benefited from higher investments, as long tunnels in Paris, new stations, etc. In addition, they mainly serve the urban part of the Paris metropolitan region. On the other hand, the last two lines, opened in the 1990s and 2000s, (D and E) did not bring major improvements. Logically, the treatment effect is higher before 1990.

The impact of a RER station on foreign firms increases over time (see table 7). The treatment effect is immediate but is clearly stronger after five years for both groups. Concerning the other firms, the results are not clear and do not coincide between groups. According to the first strategy, the RER causes a delayed increase only five years after the opening of a station. On the contrary, the second strategy suggests that the impact of RER is transitory and vanishes after 5 years. We find no effect of RER on employment. In addition, we find no difference between control and treatment groups before the opening of a station for any variable. This result can be seen as a placebo test which validates our identification strategies.

Table 6: Effect of RER on the market share of on the market share of firm location before and after 1990

	Firm location			
Presence of a RER station	0.0355* (0.0189)	0.0610*** (0.0190)	-0.00976 (0.0140)	-0.0109 (0.0233)
Log of unemployment rate	-0.0156 (0.0359)	-0.0167 (0.0397)	0.0201 (0.0371)	0.0744 (0.0894)
Log of internal market potential	-0.0197 (0.0353)	0.320** (0.136)	-0.0436 (0.126)	-0.118 (0.266)
Log of external market potential	0.332 (0.742)	0.0656 (1.156)	-3.113** (1.417)	-0.499 (2.294)
Log of total number of firms	0.293*** (0.0704)	0.0743 (0.0937)	0.0891 (0.0641)	0.360*** (0.0710)
Observations	3088	1504	2895	1410
Sample	Group 1	Group 2	Group 1	Group 2
Year FE	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓
Clustered SE	✓	✓	✓	✓
Time period	1974-1989	1974-1989	1990-2004	1990-2004

Standard errors in brackets - Significance levels: \*\*\* 1%, \*\* 5%, \* 10%

We split the sample into two parts: inner suburbs and outer suburbs<sup>4</sup>. For both groups, we find that the impact of RER on firm location is stronger in outer suburbs but not significant for inner suburbs.

The effect of the RER strongly varies across industry (see table ?? and ??). With both strategies, we find no impact of RER on retail, health, education and community work and construction. Concerning the first two sectors, enterprises are likely to serve in priority local customers at the scale of a municipality. Consequently, it is not surprising they do not benefit from a region-wide transportation system. Construction firms are very dependant from road transport. Concretely, workers need to access building sites which can be located far away from company offices. For that reason, we cannot expect a significant impact of the RER on construction firms. The manufacturing sector is responsive to urban transportation infrastructure. Since these enterprises cannot use the RER to access neither customers nor suppliers, this result is only due to a labor market effect. It highlights that access to workers is a key issue for manufacturing enterprises. The strongest effect is obtained for business services which is consistent with other papers on firm location (Py and Hatem, 2009). The positive impact of a station opening on household services is quite unexpected.

As stated before, it is very crucial to introduce the travel time in our estimations. The opening of a station is not sufficient to assess the quality of the network connection. To address this

<sup>4</sup>The inner suburbs correspond to the three closest departments to Paris: Hauts-de-Seine, Val-de-Marne and Seine-Saint-Denis. The outer suburbs designate the four closest departments to Paris: Val-d'Oise, Yvelines, Essonne and Seine-et-Marne



Table 7: Evolution of the RER effect on the market share of new firm location and job creation over time

	Firm location		Job creation		Foreign firm location	
Between 2 and 4 years before RER	-0.0115 (0.0164)	-0.0171 (0.0190)	-0.0741 (0.0541)	0.0210 (0.0686)	-0.0604* (0.0357)	-0.0833* (0.0500)
Between 1 y before and 1 y after RER	0.0172 (0.0207)	0.0135 (0.0308)	0.0208 (0.0471)	-0.0216 (0.0456)	0.0881* (0.0493)	0.108** (0.0490)
Between 2 and 4 years after RER	0.0214 (0.0181)	0.0320 (0.0220)	-0.00849 (0.0341)	0.0471 (0.0327)	0.0623* (0.0376)	0.132*** (0.0477)
5 years and more after RER	0.0355* (0.0181)	0.0407 (0.0258)	-0.0396 (0.0375)	-0.0635* (0.0364)	0.133*** (0.0459)	0.321*** (0.0697)
Log of unemployment rate	-0.0141 (0.0362)	-0.0106 (0.0389)	-0.0456 (0.0909)	-0.0105 (0.0878)	-0.0305 (0.0365)	0.00712 (0.0197)
Log of internal market potential	-0.0195 (0.0354)	0.185 (0.124)	-0.159 (0.303)	-0.327 (0.476)	0.0787*** (0.0286)	0.0321 (0.113)
Log of external market potential	0.358 (0.746)	0.632 (1.142)	-4.890** (2.344)	-4.286 (4.211)	-4.494*** (0.978)	-4.090*** (1.547)
Log of total number of firms	0.293*** (0.0706)	0.0722 (0.105)				
Log of total number of jobs			0.910*** (0.119)	0.904*** (0.216)		
Log of foreign firm stock (< 10%)					0.118*** (0.0266)	0.0854** (0.0354)
Observations	3088	1504	2123	1034	3088	1504
Sample	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2
Year FE	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓
Clustered SE	✓	✓	✓	✓	✓	✓
Time period	1974-1989	1974-1989	1978-1989	1978-1989	1974-1989	1974-1989

Standard errors in brackets - Significance levels: \*\*\* 1%, \*\* 5%, \* 10%

question, we use data on travel time by rail and we compute two alternative treatment variables: the average travel time to Paris (see table 11) and an accessibility index to population (see table 12). Note that we do not compute estimations for the second strategy since a big part of the second control group is not connected to the rail network. In that case, travel time by rail has no meaning. For the accessibility index to population, we use a simple inverse function in order to be consistent with the functional form of the external market potential. Exponential and indicator functions provide similar estimates as well as other variables that population. We obtain different but coherent results with the two treatments. A decrease of one minute in the travel time to Paris results in increase of both new firm location and total number of firms. In contrast to previous estimations, there is no increase in firm closures. It shows that the quality of rail service matters:

Table 8: Effect of RER on the market share of new firm location by geographical area

Firm location				
Presence of a RER station	-0.00185 (0.0242)	0.0126 (0.0333)	0.0648 (0.0403)	0.0745 (0.0445)
Log of unemployment rate	0.0231 (0.0549)	-0.00451 (0.0186)	-0.0248 (0.0566)	-0.0939 (0.0575)
Log of internal market potential	-0.0757*** (0.0236)	0.226 (0.179)	0.124 (0.0909)	0.171 (0.156)
Log of external market potential	1.525 (0.978)	1.792 (1.373)	-0.872 (1.283)	0.918 (1.688)
Log of total number of firms	0.476*** (0.100)	0.429*** (0.155)	0.194* (0.107)	-0.0191 (0.153)
Observations	1152	608	1504	704
Sample	Group 1 inner suburbs	Group 2 inner suburbs	Group 1 outer suburbs	Group 2 outer suburbs
Year FE	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓
Clustered SE	✓	✓	✓	✓
Time period	1974-1989	1974-1989	1974-1989	1974-1989

Standard errors in brackets - Significance levels: \*\*\* 1%, \*\* 5%, \* 10%

the growth of new firm location results in an increase of the total number of firms only if a RER station opening reduces the travel time. Accessibility index to population has no effect on firm location, it only seems to reduce firm closures. In addition, the total number of jobs grows by 2 % when the accessibility index to population increases by 1%.

#### 4.1 Robustness checks

To guaranty the robustness of our results, we tried a set of alternative models. These models are in line with Imbens and Wooldridge (2008) and Bertrand et al. (2004). The first set of models tries to provide evidence of the robustness of the identification strategy (see table 13 page 26). First, we allow for various temporal trends according to the distance to Paris. Then, we weight regressions by propensity score. According to Imbens and Wooldridge (2008), it partially addresses a omitted variable bias by reducing the correlation between included and omitted variables. For the estimation of propensity score matching, we include distance to Paris, squared distance to Paris, density, population and unemployment rate in 1968, presence of a commuter rail station in 1969 and surface. In every cases, the parameter of interest decreases but the variance also decrease, confirming the robustness of our results.

The second type of tests tries to prove that we do not wrongly state that the RER effect on firm location is significant. Bertrand et al. (2004) point that in differences-in-differences models, residuals can be serially correlated. Ignoring this correlation might cause underestimation of the standard errors of the parameter associated with the treatment variable. Consequently, one could

Table 9: Effect of RER on the market share of new firm location by industry (manufacturing, construction and retail)

	Manufacturing		Construction		Retail	
Presence of a RER station	0.0690** (0.0300)	0.0800* (0.0439)	0.0193 (0.0268)	0.0469 (0.0328)	0.0106 (0.0276)	0.0166 (0.0322)
Log of unemployment rate	-0.0315 (0.0398)	0.0227 (0.0371)	0.0432 (0.0541)	0.00272 (0.0412)	0.0331 (0.0313)	0.0698* (0.0385)
Log of internal market potential	-0.0411 (0.0291)	-0.0309 (0.146)	-0.000956 (0.0517)	0.0105 (0.150)	-0.0532 (0.0386)	-0.00229 (0.133)
Log of external market potential	-0.809 (0.985)	0.696 (1.648)	-0.642 (1.076)	1.532 (1.789)	-0.633 (1.068)	1.299 (1.528)
Log of total number of firms	0.511*** (0.120)	0.404** (0.185)	0.137 (0.138)	0.0101 (0.169)	0.479*** (0.144)	0.397* (0.222)
Log of firm stock Manufacturing	-0.223** (0.0859)	-0.318*** (0.104)				
Log of firm stock Construction			0.00903 (0.107)	-0.142 (0.109)		
Log of firm stock Retail					-0.136 (0.122)	-0.236 (0.215)
Observations	3088	1504	3088	1504	3088	1504
Sample	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2
Year FE	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓
Clustered SE	✓	✓	✓	✓	✓	✓
Time period	1974-1989	1974-1989	1974-1989	1974-1989	1974-1989	1974-1989

Standard errors in brackets - Significance levels: \*\*\* 1%, \*\* 5%, \* 10%

wrongly conclude that a policy as a causal effect when it does not have. We compute a two-stage procedure proposed by Bertrand et al. (2004) which allows to ignore time serie information. "First, one can regress  $Y_{st}$  [the dependant variable] on state fixed effects, year dummies and any relevant covariates. One can then divide the residuals of the treatment states only into two groups: residuals from years before the laws, and residuals from years after the laws. The estimate of the laws' effect and its standard error can then be obtained from an OLS regression in this two-period panel." We also compute a block bootstrap estimation that gives standard errors very similar to clustering (see table 14 page 27). The two-stage method, also provides significant estimates. These results indicate that our model properly addresses temporal auto-correlation of residuals.

Table 10: Effect of RER on the market share of new firm location by industry (business services, household services, health, education and community work)

	Business services		Household services		Health, education, community work	
Presence of a RER station	0.150*** (0.0400)	0.204*** (0.0603)	0.0559** (0.0261)	0.0648* (0.0381)	0.0334 (0.0344)	0.0385 (0.0442)
Log of unemployment rate	-0.0962*** (0.0310)	-0.0213 (0.0492)	-0.0391 (0.0396)	-0.0453 (0.0442)	-0.0500 (0.0357)	0.00770 (0.0500)
Log of internal market potential	0.0523 (0.0461)	0.0555 (0.186)	0.0478 (0.0358)	0.169 (0.169)	-0.0295 (0.0489)	0.0585 (0.177)
Log of external market potential	-5.878*** (1.065)	-7.194*** (1.707)	-1.422 (1.016)	-0.788 (1.713)	-2.139* (1.119)	-3.822* (2.126)
Log of total number of firms	0.189 (0.115)	0.0871 (0.147)	0.424*** (0.140)	0.260 (0.191)	0.0108 (0.105)	0.0104 (0.191)
Log of firm stock Business services	0.0967* (0.0514)	-0.0317 (0.0858)				
Log of firm stock Household services			-0.453*** (0.101)	-0.568*** (0.143)		
Log of firm stock Health, education					0.106* (0.0551)	0.119 (0.0718)
Observations	3088	1504	3088	1504	3088	1504
Sample	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2
Year FE	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓
Clustered SE	✓	✓	✓	✓	✓	✓
Time period	1974-1989	1974-1989	1974-1989	1974-1989	1974-1989	1974-1989

Standard errors in brackets - Significance levels: \*\*\* 1%, \*\* 5%, \* 10%

Table 11: Effect of journey time to Paris on firm and job location

	Firm location	Firm closure	Total number of firms	Job creation	Job destruction	Total number of jobs
Journey time to Paris	-0.0110 (0.00723)	-0.0165 (0.0104)	-0.00649 (0.00465)	-0.0156 (0.0252)	-0.0262 (0.0187)	-0.00333 (0.0119)
Log of unemployment rate	-0.0138 (0.0482)	-0.0330 (0.0445)	0.00239 (0.0201)	-0.0215 (0.133)	0.0784 (0.119)	0.186** (0.0933)
Log of internal market potential	0.116 (0.0844)	-0.0419 (0.109)	0.319*** (0.0524)	-0.310 (0.374)	0.128 (0.361)	
Log of external market potential	-0.355 (0.844)	-3.948*** (1.214)	0.842* (0.499)	-3.985 (2.648)	-0.889 (3.052)	1.759 (1.999)
Log of total number of firms	0.146* (0.0835)	1.044*** (0.107)				
Log of total number of jobs				0.876*** (0.144)	-0.317** (0.122)	
Observations	2355	2355	2355	1628	1628	1774
Sample	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1
Year FE	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓
Clustered SE	✓	✓	✓	✓	✓	✓
Time period	1974-1989	1974-1989	1974-1989	1978-1989	1978-1989	1978-1989

Standard errors in brackets - Significance levels: \*\*\* 1%, \*\* 5%, \* 10%

Table 12: Effect of accessibility to Paris on firm and job location

	Firm location	Firm closure	Total number of firms	Job creation	Job destruction	Total number of jobs
Index of accessibility to population	-0.00505 (0.00580)	-0.0167*** (0.00426)	-0.00752 (0.00467)	0.367 (0.366)	0.484* (0.279)	0.0219*** (0.00485)
Log of unemployment rate	-0.00188 (0.0487)	-0.0265 (0.0448)	0.00302 (0.0201)	-0.0186 (0.132)	0.0820 (0.119)	0.188** (0.0930)
Log of internal market potential	0.0931 (0.0862)	-0.0801 (0.112)	0.311*** (0.0515)	-0.304 (0.372)	0.127 (0.359)	
Log of external market potential	-0.216 (0.858)	-4.235*** (1.233)	0.845* (0.509)	-4.484* (2.641)	-1.610 (3.035)	1.513 (2.002)
Log of total number of firms	0.165** (0.0820)	1.078*** (0.106)				
Log of total number of jobs				0.870*** (0.141)	-0.313*** (0.120)	
Observations	2400	2400	2400	1650	1650	1800
Sample	Group 1	Group 1	Group 1	Group 1	Group 1	Group 1
Year FE	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓
Clustered SE	✓	✓	✓	✓	✓	✓
Time period	1974-1989	1974-1989	1974-1989	1978-1989	1978-1989	1978-1989

Standard errors in brackets - Significance levels: \*\*\* 1%, \*\* 5%, \* 10%

## 5 Conclusion

### References

- Abhijit Banerjee, Esther Duflo, and Nancy Qian. On the road: Access to transportation infrastructure and economic growth in china. NBER Working Papers 17897, National Bureau of Economic Research, Inc, 2012.
- Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan. How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119(1):249–275, 2004.
- Maggie Xiaoyang Chen and Michael O. Moore. Location decision of heterogeneous multinational firms. *Journal of International Economics*, 80(2):188–199, 2010.
- Leonard K. Cheng and Yum K. Kwan. What are the determinants of the location of foreign direct investment? the chinese experience. *Journal of International Economics*, 51(2):379–400, 2000.
- Bernard Collardey. *Les trains de banlieue, tome II de 1938 à 1999*. Editions La Vie du Rail, 1999.
- Pierre-Philippe Combes, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux. The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica* Forthcoming, 2012.
- Cletus C. Coughlin and Eran Segev. Location determinants of new foreign-owned manufacturing plants. *Journal of Regional Science*, 40(2):323–351, 2000.
- Matthieu Crozet, Thierry Mayer, and Jean-Louis Mucchielli. How do firms agglomerate? a study of fdi in france. *Regional Science and Urban Economics*, 34(1):27–54, 2004.
- James C. Davis and J. Vernon Henderson. The agglomeration of headquarters. *Regional Science and Urban Economics*, 38(5):445–460, 2008.
- Fabrice Defever. Functional fragmentation and the location of multinational firms in the enlarged europe. *Regional Science and Urban Economics*, 36(5):658–677, 2006.
- Dave Donaldson. Railroads of the raj: Estimating the impact of transportation infrastructure. Nber working papers, National Bureau of Economic Research, Inc, 2010.
- Gilles Duranton and Diego Puga. Micro-foundations of urban agglomeration economies. In J. V. Henderson and J. F. Thisse, editors, *Handbook of Regional and Urban Economics*, volume 4, chapter 48, pages 2063–2117. 2004.
- Gilles Duranton and Matthew A. Turner. Urban growth and transportation. *Review of Economic Studies*, forthcoming, 2012.
- Stephen Gibbons, Teemu Lyytikäinen, Henry Overman, and Rosa Sanchis-Guarner. New road infrastructure: the effects on firms. SERC Discussion Papers 0117, Spatial Economics Research Centre, LSE, 2012.
- Keith Head and Thierry Mayer. Market potential and the location of japanese investment in the european union. *The Review of Economics and Statistics*, 86(4):959–972, 2004.
- Adelheid Holl. Manufacturing location and impacts of road transport infrastructure: empirical evidence from spain. *Regional Science and Urban Economics*, 34(3):341–363, 2004a.

- Adelheid Holl. Transport infrastructure, agglomeration economies, and firm birth: Empirical evidence from Portugal. *Journal of Regional Science*, 44(4):693–712, 2004b.
- Guido M. Imbens and Jeffrey M. Wooldridge. Recent developments in the econometrics of program evaluation. NBER working papers, National Bureau of Economic Research, Inc, 2008.
- Sung Jin Kang and Hong Shik Lee. The determinants of location choice of South Korean FDI in China. *Japan and the World Economy*, 19(4):441–460, 2007.
- Xuepeng Liu, Mary E. Lovely, and Jan Ondrich. The location decisions of foreign investors in China: Untangling the effect of wages using a control function approach. *The Review of Economics and Statistics*, 92(1):160–166, 2010.
- Philippe Martin and Carol Ann Rogers. Industrial location and public infrastructure. *Journal of International Economics*, 39(3-4):335–351, 1995.
- Thierry Mayer, Isabelle Mejean, and Benjamin Nefussi. The location of domestic and foreign production affiliates by French multinational firms. *Journal of Urban Economics*, 68(2):115–128, 2010.
- Guy Michaels. 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, 2008.
- Loriane Py and Fabrice Hatem. Internationalisation et localisation des services : une analyse sectorielle et fonctionnelle appliquée aux firmes multinationales en Europe. *Économie et Statistique*, 426(1):67–95, 2009.
- Kurt Schmidheiny and Marius Brülhart. On the equivalence of location choice models: Conditional logit, nested logit and Poisson. *Journal of Urban Economics*, 69(2):214–222, 2011.
- Vanessa Strauss-Kahn and Xavier Vives. Why and where do headquarters move? *Regional Science and Urban Economics*, 39(2):168–186, 2009.
- Stephen Ross Yeaple. Firm heterogeneity and the structure of U.S. multinational activity. *Journal of International Economics*, 78(2):206–215, 2009.



## **A Robustness checks**

Table 13: Effect of RER on the market share of new firm location - Robustness checks

Presence of a RER	0.0316*	0.0583***	0.0316*	0.0583***	0.0305**	0.0476***
Standard error	0.0104**		0.0104**		0.0104**	0.0104**
	(0.0187)	(0.0188)	(0.0187)	(0.0188)	(0.0144)	(0.0141)
	(0.0144)					
Log of unemployment rate	-0.0128	-0.0158	-0.0128	-0.0158	-0.00464	-0.0108
Standard error	-0.0108		-0.0108		-0.0108	-0.0108
	(0.0357)	(0.0389)	(0.0357)	(0.0389)	(0.0318)	(0.0308)
	(0.0318)					
Log of internal market potential	-0.0212	0.324**	-0.0212	0.324**	-0.00536	0.241*
Standard error	0.0058		0.0058		0.0058	0.0058
	(0.0364)	(0.136)	(0.0364)	(0.136)	(0.0430)	(0.128)
	(0.0440)					
Log of external market potential	1.131	0.356	1.131	0.356	0.238	0.242
Standard error	0.55		0.55		0.55	0.55
	(0.946)	(1.303)	(0.946)	(1.303)	(0.695)	(1.190)
	(0.872)					
Log of total number of firms	0.307***	0.0889	0.307***	0.0889	0.299***	0.127
Standard error	0.005		0.005		0.005	0.005
	(0.0731)	(0.0923)	(0.0731)	(0.0923)	(0.0688)	(0.0940)
	(0.0719)					
Year × distance to Paris	0.000115	0.00130	0.000115	0.00130		
Standard error	0.000124		0.000124			
	(0.00119)	(0.00209)	(0.00119)	(0.00209)		
	(0.00111)					
Year × (distance to Paris)	-0.0000143	-0.0000419	-0.0000143	-0.0000419		
Standard error	0.0000853		0.0000853			
	(0.0000206)	(0.0000592)	(0.0000327)	(0.0000592)		
	(0.0000312)					
Observations	3088	1504	3088	1504	3088	1504
	3088					
Sample	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2
Year FE	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓
Clustered SE	✓	✓	✓	✓	✓	✓
Weighted by propensity score			✓	✓	✓	✓

Table 14: Effect of RER on the market share of new firm location - Robustness checks

Presence of a RER station	0.0355* (0.0189)	0.0613*** (0.0195)	0.0355* (0.0197)	0.0613*** (0.0190)	0.0129* (0.00781)	0.0255** (0.0105)
Observations	3088	1504	3088	1504	1424	704
Sample	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2
Year FE	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓
Clustered SE	✓	✓				
Bock Bootstrap SE			✓	✓		
Two-stage estimation					✓	✓