WORKING

paper

BANQUE DE FRANCE

EUROSYSTÈME

Communication Costs and the Internal Organization of Multi-Plant Businesses: Evidence from the Impact of the French High-Speed Rail

Pauline Charnoz[1], Claire Lelarge[2] & Corentin Trevien[3]

August 2017, WP 635

ABSTRACT

We document the impact of travel time between headquarters and affiliates of geographically dispersed corporate groups on the management of such business organizations. Theory suggests that the easier circulation of managers might facilitate the transmission of information between production plants and headquarters, thus fostering growth and functional specialization (on production activities) at remote affiliates and decreasing operational costs at the group level. We test these predictions on the population of French corporate groups, using the expansion of the High Speed Rail network as a shock on internal travel times. We estimate that HSR induced the creation of one production job for the average affiliate in the service industries (against 0.2 job in retail, trade or manufacturing industries), and the shift of around one managerial job from affiliate to HQ. At the group level, descriptive regressions suggest that the impact on the operational profit margin is around 0.5 percentage points in most industries.[4]

Keywords: Communication costs, headquarters, firm organization, public infrastructure, high-speed rail.

JEL classification: R30 L22 R40

[1] DARES, Crest and RITM, Université Paris Sud /Paris Saclay (France), pauline.charnoz@ensae.fr
[2] Banque de France (France) and CEPR, claire.lelarge@banque-france.fr
[3] Insee, Crest and Sciences-Po. (France), corentin.trevien@ensae.fr

NON-TECHNICAL SUMMARY

Large corporations operating multiple affiliates located in different areas are prevalent and account for a disproportionate fraction of output and employment. The Figure below shows in particular that a very large fraction of employment throughout the entire French metropolitan territory is managed at arm's length by headquarters located in Paris. The splitting and implantation decision of such corporations is well understood: they involve a trade-off between the gains to access remote markets and the managerial costs to operate distant affiliates. Yet, the precise nature of these managerial costs remains largely a black box. In this paper, we contribute to fill this gap and provide detailed evidence about the nature of the higher managerial costs implied by geographic dispersion.

Sphere of Corporate Influence of 4 of the Largest French Cities
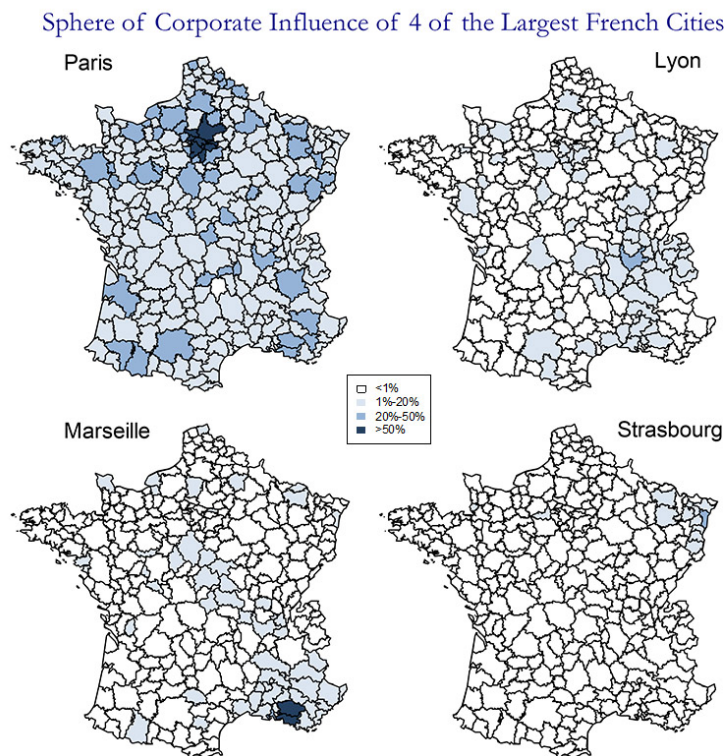
Note: authors' computations using matched DADS and LIFI data, as of 2011. The figure shows the share of private employment in each commuting zone that is under control of HQs located in the respective cities, as a percentage of total private employment.

We take advantage of large scale administrative and survey data allowing us to describe the structure of the workforce of all French corporate groups over 19 years, from 1993 to 2011. We also rely on the expansion of the French high-speed rail (HSR) network over that period to assess the organizational impact of decreases in communication costs between headquarters and affiliates of groups benefiting from the new infrastructure. Our empirical investigation is guided by theoretical predictions. The literature in corporate finance (Giroud, 2013; Giroud and Mueller, 2015) suggests that the geographical dispersion of affiliates is likely to hamper information gathering and monitoring by managers of the group's headquarters. This amplifies moral hazard problems and decreases the incentives to invest in remote affiliates, thus negatively affecting their size. The literature in economic geography (Duranton and Puga, 2005) suggests in turn that these information flows are to be understood as transfers of HQ services from

headquarters to affiliates. Rendering these transfers more costly increases the need to locate the corresponding support activities at remote affiliates, which lowers the functional specialization of affiliates on their production activities.

Our regressions show that the impact of reduced travel times is positive on both affiliate size (production capacity) and functional specialization, as predicted by theory. Furthermore, we obtain that this impact is highest in the service industries, where information to be transmitted is arguably softer (Petersen and Rajan, 2002). Results are also significant in the trade and manufacturing industries, but point estimates are lower. To gauge the orders of magnitude, we propose to quantify the impact of the HSR network as a whole on the management structure and size of remote affiliates benefiting from it as of 2011. We obtain that for an affiliate benefiting from the average decrease in travel time to its headquarters, the availability of the HSR infrastructure induced the shift of roughly one job from administrative to production activities in service industries against 20% of a job in other industries (retail, trade or manufacturing). Affiliates operating in the manufacturing and business services industries also experienced decreases in production labor costs, of around half the cost of a production job on average. At the group level, our regressions suggest that the impact on the operational profit margin ranges from 0.5 to 1.5 percentage points depending on the industry. Our empirical analysis can also be used to provide estimates of the overall cost of geographical dispersion for multi-site businesses, although in a partial equilibrium setting. For example, our estimates imply that remote affiliates in the personnel service industries would operate with roughly 1.5 additional production jobs on average if geographical distance could be fully abolished by a perfect communication technology. The figure would rise to 4 production jobs on average in business services industries, and to 2 production jobs in manufacturing industries. The latter quantifications are to be interpreted as upper bounds for the productivity effect that can be expected from internal communication technologies.

Le Lien entre Coûts de Communication et Organisation Interne des Entreprises Multi-Sites : l'Impact du TGV

Résumé

Nous documentons l'impact des temps de parcours entre quartiers généraux et filiales de groupes d'entreprises géographiquement dispersés sur la gestion de ces structures. Le raisonnement théorique suggère que les managers, en se déplaçant plus aisément, facilitent la transmission d'information entre QG et unités de production, ce qui incite à la spécialisation fonctionnelle de ces dernières. Nous testons cette prédiction sur la population des groupes d'entreprises françaises, en utilisant l'expansion du réseau ferroviaire à grande vitesse comme un choc sur les temps de parcours internes. Le TGV aurait en moyenne induit la création d'un emploi de production dans les filiales de services, de 0,2 emploi dans les filiales industrielles ou commerciales, ainsi que le transfert d'un emploi de gestion vers les QG. Au niveau du groupe, les régressions descriptives suggèrent que l'impact sur la marge opérationnelle brute est d'environ 0,5 point de pourcentage dans la plupart des secteurs.

Mots-clés : Coûts de communication, quartier général et organisation d'entreprise, infrastructure publique, Train à Grande Vitesse.

1 Introduction

Large corporations operating multiple affiliates located in different areas are prevalent and account for a disproportionate fraction of output and employment. Figure 1 shows that in France, such geographically dispersed corporate groups account for around 40% (6 million workers) of total employment in the for-profit sector in 2011,[1] and even more when taking account of groups headquartered from abroad. The splitting and implantation decisions of such corporations has been the object of study of a large literature in economic geography (e.g. Aarland et al., 2007; Davis and Henderson, 2008; Henderson and Ono, 2008; Strauss-Kahn and Vives, 2009) as well as in trade (see Mugele and Schnitzer, 2008 or the recent survey about the prominent role of multinationals in Antràs and Yeaple, 2014). They involve a trade-off between the gains to access remote markets and the managerial costs to operate distant affiliates. Yet, the precise nature of these managerial costs remains largely a black box. More generally, very little is known about the way business organizations that are geographically dispersed are managed in practice. In this paper, we contribute to fill this gap and provide detailed evidence about the nature of the higher managerial costs implied by geographic dispersion.

Figure 1: The Prevalence of Geographically Dispersed Business Organizations
Break-down of total employment in the for-profit sector
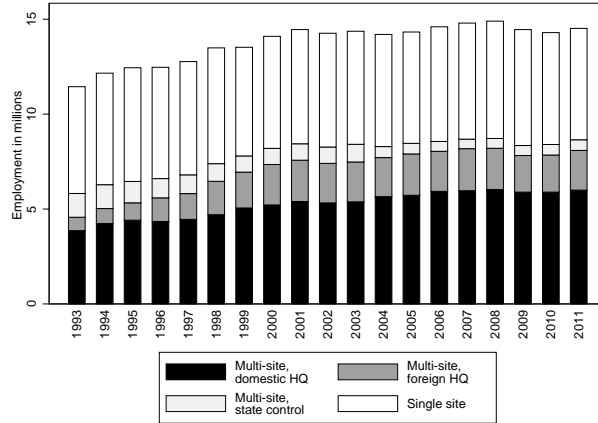
Sources: Matched DADS files and LIFI survey, covering the for-profit sector (except agricultural activities and workers of the personnel service industries directly employed by households). See section 3 for further details.
Note: Employment is measured in terms of days of work, normalized by 365 (to be comparable with headcounts in full time equivalents).

To that end, we take advantage of large scale administrative and survey data allowing us to describe the structure of the workforce of all French corporate groups, over 19 years, from 1993 to 2011. We also rely on the expansion of the French high-speed rail (HSR) network over that period to assess the organizational impact of decreases in communication costs between headquarters and affiliates of groups benefiting from the new infrastructure.

Our empirical investigation is guided by theoretical predictions. The literature in corporate finance (Giroud, 2013; Giroud and Mueller, 2015) suggests that the geographical dispersion of affiliates is likely

---

[1] Excluding workers directly hired by households.

to hamper information gathering and monitoring by managers of the group's headquarters. This amplifies moral hazard problems and decreases the incentives to invest in remote affiliates, thus negatively affecting their size. A prediction of this literature is therefore that higher travel times impede information flows between headquarters and affiliates and affect affiliates size negatively. The literature in economic geography (Duranton and Puga, 2005) suggests in turn that these information flows are to be understood as transfers of "headquarters services" from headquarters to affiliates. This "functional" approach therefore provides the complementary prediction that increasing the cost of these transfers increases the need to locate the corresponding support activities at remote affiliates, which decreases the functional specialization of affiliates on their production activities. Last, we also test whether workers at affiliates are differentially impacted by changes in internal communication costs, depending on their skills, as suggested in Garicano (2000) or Garicano and Rossi-Hansberg (2006).[2]

We test these predictions in the French data. We address the problem of the endogenous placement of the HSR infrastructure by taking advantage of the dyadic nature of our setting, where observations are "pairs" of affiliates (in one location) and headquarters (located elsewhere). This feature allows for the introduction of high dimensional location specific time fixed effects as in Giroud (2013). These fixed effects control accurately for the demand shocks and local price effects (e.g.land prices) potentially induced by the new infrastructure. Identification in our setting therefore comes from differential variation of communication costs across affiliates located in the same area, but belonging to differently structured multi-unit corporate groups. We also take advantage of alternative (or complementary) identification strategies to check the robustness of our baseline results. First, we take advantage of the information about un-realized lines as in Donaldson (2014) to construct an estimation sample that is selected, but arguably less likely to suffer from endogeneity concerns. Alternatively, we focus on affiliates located at intermediate places, which benefited from HSR only because they are on the way to a large city as in Chandra and Thompson (2000) and Banerjee et al. (2012). Our regressions show that the impact of reduced travel times is positive on both affiliate size (production capacity) and functional specialization, as predicted by theory. Furthermore, we obtain that this impact is highest in the service industries, where information to be transmitted is arguably softer (Petersen and Rajan, 2002). Results are also significant in the trade and manufacturing industries, but point estimates are lower. To gauge the orders of magnitude, we propose to quantify the impact of the HSR network as a whole on the management structure and size of remote affiliates benefiting from it as of 2011. We obtain that for an affiliate benefiting from the average decrease in travel time to its headquarters, the availability of the HSR infrastructure induced the shift of roughly one job from administrative to production activities in service industries against 20% of a job in other industries (retail, trade or manufacturing). Affiliates operating in the manufacturing and business services industries also experienced decreases in production labour costs, of around half the cost of a production job on average. At the group level, our regressions suggest that the impact on the operational profit margin ranges from 0.5 to 1.5 percentage points depending on the industry. Our empirical analysis can also be used to provide estimates of the overall cost of geographical dispersion for multi-site businesses, although in a partial equilibrium setting. For example, our estimates imply that remote affiliates in the personnel service industries would operate with roughly 1.5 additional production jobs on average if geographical distance could be fully abolished by a perfect communication technology. The figure would rise to 4 production jobs on average in business services industries, and to 2 production

---

[2]Related to this aspect, Bassanini et al. (2017) and Landier et al. (2009) provide an in-depth analysis of labour adjustments on the extensive margin (dismissals) at remote affiliates depending on the distance to their headquarters.

jobs in manufacturing industries. The latter quantifications are to be interpreted as upper bounds for the "productivity" effect that can be expected from internal communication technologies.[3]

Our paper is related to two main strands of the literature. First, it contributes to the literature that seeks to understand productivity differences across firms. Several articles investigate the productivity advantage of large firms or groups. Most closely related to our work are the articles by Bloom et al. (2012a) and Bloom et al. (2012b) who focus on the productivity advantage of multinationals, while Atalay et al. (2014) focus on vertically integrated corporate groups. Interestingly, the latter emphasize the importance of internal transfers of "intangible inputs" such as "managerial talent" to explain the value of vertical integration, and ultimately the productivity premium of such corporations. Their strategy to describe the impact of changes of ownerships to shed light on the management of multi-unit groups complements our own approach, which rather relies on shocks on internal communication costs. Most interestingly for us, they obtain that newly integrated firms tend to decrease the number of their "local" managers (the share of non-production workers drops), which are replaced with the managerial resources of headquarters. This result parallels our own findings, first because headquarters' managers appear to be substitutes for local managers, and second because they appear to be the privileged managerial resource within complex business organizations (at the expense of managers located in remote affiliates).

Our paper is also related to the large literature investigating the economic impact of transport infrastructures. Most papers in this field focus on standard rail or road infrastructure, which essentially generate a reduction in trade barriers pertaining mainly to the circulation of *goods* (Donaldson and Hornbeck, 2016). The impact of such shocks are well understood (Melitz, 2003): they induce increases in the global volume of trade as well as in firm selection, and ultimately aggregate productivity (Michaels, 2008; Datta, 2012; Banerjee et al., 2012; Donaldson, 2014; Faber, 2014; Ghani et al., 2016). In contrast, high-speed rail is a transportation technology that is almost only accessible to *people*.[4] As we argue above, the mechanisms involved are likely to be very different (Bernard et al., 2015; Nunn, 2007; Cristea, 2011).

The remainder of the article is organized as follows. Section 2 proposes a synthesis of the theoretical predictions relating internal communication costs to the managerial organization of corporate groups and their overall performance. In section 3, we describe our data as well as the French HSR network. We also provide a comprehensive picture of the geographical dispersion of French corporate groups. Our empirical strategy is explained in section 4 and the main regressions results at the affiliate level are discussed in section 5. Section 6 provides additional descriptive results at the (entire) group level and section 7 concludes.

---

[3]In this respect, our results also show (by revealed preferences) that face-to-face interactions made easier by the HSR technology remain crucial, in spite of the development of other means of communication (phone, e-mail, visio-conference) over the same period (Storper and Venables, 2004). The difficult identification issues underlying this statement are discussed in detail in section 4.3.

[4] In the case of the French High Speed Rail program, most of the infrastructure is not even accessible to freight (at the notable exception of mail) because HSR tracks are too steep for the weight of freight trains. More generally, in this paper, we only focus on the "productivity" effect of HSR and do not consider the potential market creation impact of HSR. This aspect could be particularly relevant in industries related to tourism, but the analysis of this dimension would require an entirely different identification strategy.

2 The Management of Multi-Plant Businesses: A Review of Theoretical Predictions

2.1 Geographical Dispersion and Affiliate Size

The literature in corporate finance (in particular Giroud, 2013, Giroud and Mueller, 2015 or Kalnins and Lafontaine, 2013) relates travel times to information transmission between headquarters and affiliates in settings characterized by information asymmetries and moral hazard problems.

More specifically, these contributions focus on the two-layer managerial structure of corporate groups, with both managers at headquarters ("principals"), who are endowed with the ultimate decision rights, and managers at remote affiliates ("agents") who are not, but who have an informational advantage over managers at headquarters about the profitability of local investment projects. The management of such business organizations features a moral hazard problem when the interests of managers at affiliates are not fully aligned with the interests of managers at headquarters. Whether managers at affiliates over-invest (over-hire) or under-invest (under-hire) when decision rights are delegated to them depends on whether they have preferences for (local) "empire building" strategies, or conversely if they prefer an excessively "quiet life".[5] In both cases, resources allocated to affiliates might be misallocated, sometimes diverted. To limit the inefficiencies and informational rents arising in such settings, headquarters have an incentive to limit the resources allocated to remote affiliates below what would be optimal in the absence of such information asymmetries.

Higher travel times between headquarters and remote affiliates render monitoring more difficult and are therefore associated with lower investment (Giroud, 2013) and lower complementary employment (Giroud and Mueller, 2015) at affiliates, i.e. smaller affiliate size.[6] To test the relevance of such mechanisms in the French data, we replicate the analyses initially performed by Giroud (2013) and Giroud and Mueller (2015) on US multi-plant firms and investigate the correlations between employment at affiliates and travel time to their headquarters, with an expected negative sign. We also examine the link between investment and travel time. However, investment is not observable at the affiliate level in the French data, such that we are only able to estimate these regressions at the group level.[7]

2.2 Geographical Dispersion and the Functional Specialization of Affiliates

A complementary insight stems from the literature in economic geography, suggesting that the previous type of mechanism might be also relevant to analyze the articulation of the different "functions" within a firm or within a corporate group. In Duranton and Puga (2005) for example, firms are considered as bundles of two broad types of functions: "headquarters services" on one hand, and "production activities" on the other hand. These two functions can be either pooled in the same location or split into different plants. Splitting is costly because of the agency problems outlined in section 2.1: Duranton and Puga (2005) model such mechanisms in a rather reduced form, as a fraction of managers' time that is lost in travels to visit

---

[5]Bertrand and Mullainathan (2003) actually show that the second case is more frequent among US managers, which implies that managers at affiliates are likely to under-invest when investment decisions are delegated to them.

[6]To ease exposure, the entire reasoning in section 2.1 implicitly assumes that investment decisions for affiliates are delegated to local managers. This needs not be the case: the prediction of a negative relationship between travel time and affiliate size holds (in expectation) whatever the organizational choice if HQ managers are risk averse.

[7]The identification issues arising from the aggregation across all affiliates at the group level are explained in section 6.

remote affiliates.[8] However, there are gains to split firms when there are agglomeration economies that are differentiated for each function: for example, the possibility to outsource certain activities to local suppliers that might be specifically appropriate, the optimization of labor costs across local labour markets (depending on the characteristics of local labour supply), or simply easier access to local markets for final goods.

A distinctive prediction of this literature therefore relates travel times between headquarters and affiliates to the optimal mix of functions present in each implantation. More precisely, as shown more formally in Appendix A, higher travel times increase the cost of transferring headquarters services to remote affiliates, thus decreasing the incentives to specialize each site by function.[9] We therefore expect affiliates that are located further away to be relatively less focused on their production activities because of the need to complete locally some of the administrative tasks that are too costly to complete from headquarters.

This prediction can be tested by regressing the share of employment at affiliates that is devoted to production activities (as opposed to managerial activities) against travel time between affiliates and headquarters. We expect the sign of the corresponding coefficient to be negative. Conversely, we expect that the share of managerial functions that are completed at headquarters is low when geographical dispersion is high. We test this prediction by estimating the correlation, at the group level, between the share of managerial jobs that are located at headquarters and the average travel time to their affiliates. We expect this correlation to be negative. In complementary regressions, we also investigate whether production workers at remote affiliates might be differentially affected by decreases in communication costs with headquarters, depending on their skills. Following Garicano (2000) and Garicano and Rossi-Hansberg (2006),[10], we hypothesize that higher communication costs with headquarters in the form of higher travel times might increase the need for skills at remote affiliates, thus generating a positive correlation between travel times and the wages paid out at remote affiliates.

2.3 Geographical Dispersion and Operational Profit at the Group Level

In sections 2.1 and 2.2, higher travel times between headquarters and affiliates are always associated with higher overall costs of operating large, multi-plant businesses. We therefore expect a negative correlation at the entire group level between the operating profit margin and average travel time to affiliates. Furthermore, corporate groups which are more expensive to operate should be relatively smaller, other things equal, because the resources that are dissipated in communication costs cannot be invested in alternative projects, at either existing or new units. First, remote affiliates are unambiguously expected to be smaller (section 2.1) because of the higher costs of operating them. Second, whether the other units are also harmed by the highest

---

[8]Refer to Acemoglu et al. (2007), for example, for a detailed description of the underlying trade-offs involved in principal-agent framework for the optimal organizational choice between delegation of authority to a local manager or centralized decision taking at headquarters. Typically, the gain to rely on the local manager's superior information is compared with the risk that he could use his informational advantage to make choices that are not in the best interest of the group as a whole. In such a framework, shorter travel times ease information acquisition by principals, decrease the informational advantage of local managers, and shift the trade-off in favour of centralized control at headquarters.

[9]In Duranton and Puga (2005), this result is a general equilibrium outcome: a decrease in the cost of managing remote affiliates (if sufficiently large) shifts the entire economy from an equilibrium where no firm is geographically dispersed and cities specialize by sector, to an equilibrium where all firms adopt a multi-site organizational form and cities specialize by function. The authors suggest that a "smoother" result would hold in an augmented version of the model incorporating some firm level heterogeneity, together with (sufficiently large) adjustment costs of reorganization. Such additional features would explain why all firms would not split instantaneously and relocate all of their units along the new HSR lines as they open.

[10] See also Garicano and Rossi-Hansberg (2012); Antràs et al. (2006) and Antràs et al. (2008).

overall operating costs depends on the financing or managerial resource constraints of the group. Typically, *constrained* groups will tend to reallocate the scarce resources from relatively expensive units to relatively cheaper ones. In this case, higher travel times to one affiliate might foster the reallocation of resources to closer affiliates that are cheaper to operate, thus positively affecting them (Stein, 1997).[11] In *unconstrained* groups in contrast, no such externality arises, such that no other unit is affected by the cost of operating one specific unit. We test these predictions by investigating the correlations between affiliates' size and variations in travel time *at other units*. We expect increases in travel times at other units to positively affect the amount of resources allocated to a considered affiliate in constrained groups only.

3   Data

3.1   The Geographical Organization of French Corporate Groups

We rely on several matched datasets to recover the geographical organization of French corporate groups, as well as the structure of their internal workforce and various accounting indicators, most importantly their investment decisions and profit margin.

As in Combes et al. (2005), we first use the LIFI[12] files containing the structure of ownership of French firms to recover the structure of French corporate groups. Prior to 1999, the LIFI files only covered companies of the private sector whose size was above at least one of three different thresholds, defined in terms of financial stakes in other firms (higher than 1.2 million euros), sales (60 million euros) or employment (500 workers). From 1999 onwards, these files are complemented with the Diane-Amadeus (Bureau Van Dijk) dataset, which is constructed from commercial court records and covers smaller business groups. While coverage is a little less comprehensive in the first five years of our period of analysis,[13] our file contains almost exhaustive information about corporate groups operating in France from 1999 onwards.

This file allows us to identify the various plants within any French corporate group (or independent firm) as in Combes et al. (2005). We then borrow from the property right approach to introduce the distinction between the "headquarters" of business groups and their "affiliates". As in Bolton and Scharfstein (1998), the role of headquarters in our setting is to bring control on the allocation of resources inside corporate groups. We consider that managers in the entity ultimately owning the assets of any affiliate have the residual control rights over its assets, and can decide upon how these are to be used.[14] The corresponding entity, ie. the site having the actual ultimate control over all assets in the group *via* the direct or indirect ownership of more than half of the equity in any of the group affiliates (La Porta et al., 1999), is considered as headquarters in

---

[11]However this potentially positive indirect impact is necessarily smaller, in absolute terms, than the direct, negative one on the remote affiliate.

[12]The acronym "LIFI" stands for "LIaisons FInancières" (financial linkages). See e.g. Boutin et al. (2013) or Cestone et al. (2016) for alternative uses of this dataset.

[13]This implies that we do not detect some of the multi-site groups in the early period, which are therefore excluded from our estimation sample (see section 4). This mainly affects our estimates for the opening of the Northern line in 1993/1994, with a loss in the statistical power of our setting. Undetected multi-site groups are the smallest ones, and typically less likely to contribute to identification in our setting, which in the most stringent specification includes affiliate and times × group fixed effects (this requires the group to own several, differentially affected sites, as explained in section 4). Therefore, the bias potentially generated by under-coverage of the small groups is likely to be small.

[14]This implies that they also have the ultimate control over hiring decisions, although on a daily basis, these decisions might be delegated to local managers at affiliates (or even, divisions within affiliates). See Bassanini et al. (2017) for an empirical investigation of the managerial practices "at the bottom", at the level of multi-site firms or companies.

our dataset.[15] All other plants are considered as affiliates.

However, it is well-known that plants within groups (and even companies within groups) might be created, terminated and replaced for reasons that are uncorrelated with the human resource management practices we want to focus on (Aubert and Sillard, 2005; Picart, 2004). For example, firm and plant identifiers change when the legal status of companies evolves, most often because of regulatory constraints (e.g. upper bounds on the admissible number of shareholders for certain legal forms, *etc.*) or for fiscal and administrative reasons which are entirely orthogonal to the mechanisms described in section 2. We choose to abstract from such phenomena by aggregating the information across all plants of a same group, having the same activity (at the 1 digit level) and located in the same municipality (French "commune") into a single "affiliate" unit.[16] This aggregation is almost transparent since all variations of travel time induced by HSR line openings are homogeneous within municipalities - and even within commuting zones as we explain below. Conversely, the main benefit of this procedure is to increase a little bit the statistical power of our setting by increasing the number of years an affiliate unit is observed - 3.2 years on average in our final estimation sample.

This dataset is complemented with exhaustive fiscal information sourced from the BRN/RSI files,[17] providing complementary accounting information such as investment and operational profit. It is also matched with exhaustive employment data (DADS), which we describe in detail in section 3.2. Together, these files allow us to provide a comprehensive picture of remote control in France. Figure 1 shows that between 1993 and 2011, roughly 50% of the domestic workforce is employed in multi-site business organizations, be they managed by domestic headquarters, foreign headquarters or by the French State. In our regression analyzes however, we exclude the last two categories: we exclude affiliates headquartered from abroad because we miss the precise inforamtion about the location of their headquarters and are not able to compute travel times. We also exclude affiliates ultimately controlled by the State because the mechanisms presented in section 2 might not apply to their specific case.[18]

Figure 2 describes the geography of remote control, as of 2011. In this figure, we introduce the geographical meshing, at the commuting zone level, we rely on in our analysis. Commuting zones allow us to take account of the fact that workers (or more specifically, managers) who are asked to travel for professional reasons might depart from home rather than from their workplace. In the absence of information about the precise location of their departure point, commuting zones are by definition the best statistical guess we can get about it.[19] Panel (A) of figure 2 provides for each commuting zone, the share of businesses that are under the control

---

[15]In cases where these headquarters are non-employer holdings, we rather choose the employer company that is most directly related to the holding, in terms of rank of control, and in cases where several companies meet this criterion, we select those having the largest share of executives (see section 3.2 below). The goal of this procedure is to locate (probabilistically) the upper management team of the corporate group.

[16]Our aggregation procedure also enables to abstract from plant transfers occurring on very short distances (within the same municipality), for reasons related to building capacity. There are as many as 36,000 municipalities in France, such that the average surface is lower than 18 km^2.

[17]See Boutin et al. (2013) for a previous use of these files in a similar setting. Note that affiliates of corporate groups are not allowed to report to the fiscal administration with the simplified scheme (RSI) such that in practice we rely almost solely on the BRN data.

[18]In such organizations, profit maximization might not be the main criterion at the headquarters level.

[19]Commuting zones are defined "as the geographical area within which most of the labour force lives and works, and in which establishments hire most of their workforce". The empirical analysis is in fact not very sensitive to this choice: earlier versions of this paper were based on a geographical meshing constructed from the train station catchment areas and provided the same qualitative and quantitative results. Appendix D discusses these alternative methodological choices in greater detail.

of headquarters located outside the commuting zone, i.e. further away than the standard daily commuting patterns. This share is higher than 15% in most commuting zones, and above 20% in a number of zones located in the northern half of the country, in the "sphere of corporate influence" of the capital city (Paris).[20] The employment weighted version of the same indicator is presented in panel (B) and shows that in most of the country, more than 30% of employment is managed at arm's length. This share rises to rates above 50% in a significant number of commuting zones, mostly located in the northern part of the country (near Paris).

Figure 2: Share of Affiliates Under Control of a HQ Located Outside the Considered Commuting Zone
In percentages, by 2011
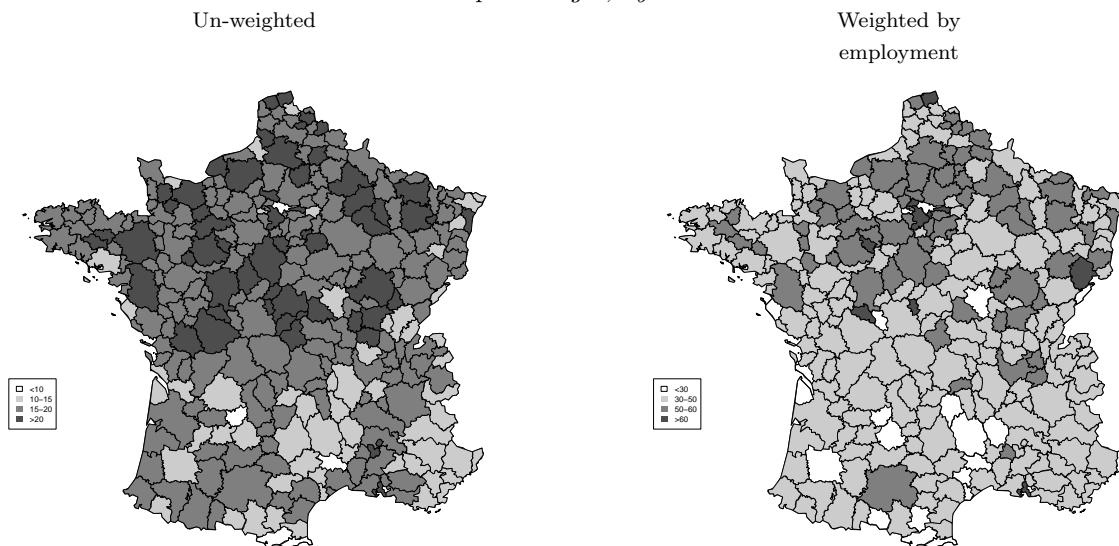
Un-weighted

Weighted by
employment

Sources: Matched DADS files and LIFI survey, covering the for-profit sector (except agricultural activities and workers of the personnel service industries directly employed by households).
Notes: The left panel describes the number of affiliates in each commuting zone that are under the control of a HQ located outside the zone, as a share of the total number of businesses (affiliates and HQs) located in each zone. The right panel describes the share of private employment in each commuting zone that is under control of an external HQ. This indicator can be interpreted as an employment weighted version of the previous.

3.2 The Organization of the Workforce within French Corporate Groups

We complement the previous files with exhaustive worker level information sourced from the DADS[21] files. These files are available from 1993 onwards, and include roughly 14 million workers per year in the recent period. They allow to track economically active plants within each group and provide a rich description of their internal workforce and wage structure.

Most importantly, the classification of occupations in the DADS files allows us to contrast the workforce allocated to production activities with the workforce allocated to managerial activities, both at headquarters and affiliates. This distinction is in particular useful to test the empirical predictions of section 2.2. We

---

[20]See appendix B for a comparison of the "spheres of corporate influence" of different French cities showing the disproportionate weight of Paris, as compared with any other French city (e.g. Lyon, Marseille, Strasbourg).
[21]The acronym "DADS" stands for "Déclarations Annuelles de Données Sociales". See e.g. Caliendo et al. (2015) for a previous use of these files in a similar setting.

Figure 3: Structure of the Workforce at Affiliates vs. Headquarters, 1993 - 2011

Workforce at affiliates   Workforce at headquarters
(A) Functions related to management
As a share of total employment

(B) Functions related to production
As a share of total employment
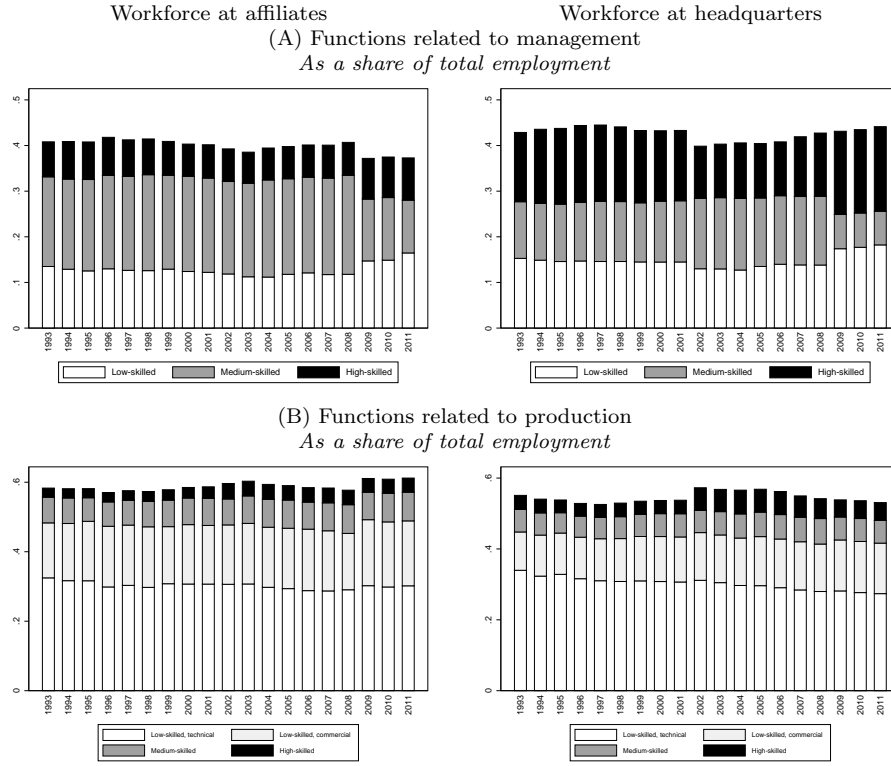
Sources: Matched DADS files and LIFI survey; units which are part of geographically dispersed business organizations.
Notes: Employment is measured in days. The breaks in the series in 2002 and 2009 were generated by a change in the codification procedure for occupations in the DADS files.

interact this typology of functions with the indicators of hierarchical layers proposed in Caliendo et al. (2015) to get the following partition (codes of the French classification are indicated in parentheses):

- *Functions related to management:*

    – Low-skilled: office workers and clerks (54),
    – Medium-skilled: mid-level manager, mid-level professionals and related (42-46),
    – High-skilled: heads of businesses (2), top managers and professionals (35, 37).
      This last category will capture the concept of "managers" introduced in section 2.

- *Functions related to production:*

    – Low-skilled: skilled industrial and manual workers (62 and 63), drivers (64), skilled transport and wholesale workers (65), unskilled industrial workers (67),[22]
    – Medium-skilled: technicians (47), supervisors and foremen (48),
    – High-skilled: science and educational professionals (34), technical managers and engineers (38).

---

[22]Notice that we allocate commercial low-skilled workers to production activities, which is of particular relevance for the retail and trade industries. For medium and high-skilled workers, unfortunately, the classification available in our file does not allow to distinguish between managerial and commercial activities.

Figure 3 describes the structure of the workforce in terms of these definitions, both at affiliates and headquarters. Employment in each occupation is measured in days of work (between start and end of the labour contract of each worker) in order to measure employment in full time equivalents. Unfortunately, two methodological changes in the coding of occupations, which were introduced by the Statistical Institute in 2002 and 2009, produce two breaks in the series.[23] Abstracting from this difficulty, panel (A) of figure 3 shows that *managerial functions* are equally important, on average, at headquarters and affiliates. What differentiates HQs from affiliates sharply is not the weight of these functions, but the structure of skills *within* them: headquarters employ 15 to 20% of their workforce in higher management positions, against 5 to 10% in the case of affiliates. Conversely, affiliates employ around twice as many middle managers (ca. 20%) as headquarters. In contrast, the structure of the workforce allocated to *production activities* (panel B in figure 3) is not highly contrasted between headquarters and affiliates: headquarters only tend to employ slightly more high-skilled production workers.

Unreported complementary descriptive statistics unsurprisingly confirm that the structure of the workforce at affiliates is highly differentiated across industries. For example, skilled production and managerial workers represent 11% of the workforce in manufacturing industries and even 21% in business services. In contrast, the share ranges between 6 and 8% only in the personnel services, retail and trade, or transport industries. More specifically, the share of "managers" (high-skill managerial workers) is around 4% in all industries, except in the business services where it reaches 10%.[24]

3.3   HSR Network and Rail Travel Times

The last ingredient that is needed for our empirical analysis is the information about travel times between affiliates and headquarters. We approximate this information with rail travel times. This first required to collect detailed information from the archives of the French national rail company, its open-data platform and complementary technical publications to recover the expansion of the HSR network over time. The outcome of this first task is represented on figure 4, which describes the evolution of the HSR since 1981 and more specifically, during our period of analysis (1993 to 2011). The first HSR segment was opened in 1981 on the track between the two French largest cities, Paris and Lyon. This line was subsequently extended in 1994 and 2001 to ultimately reach Marseille on the Mediterranean Coast. The network was also extended towards the Atlantic coast in 1989-1990, towards Lille and London in 1993 and 1994, and towards Strasbourg and Frankfurt (in Germany) in 2007.[25]

The exact procedure that was implemented to compute rail travel times from the previous information and historical time tables is explained in Appendix D.[26] High-speed trains operate at twice the maximum standard rail speed (up to 320km/h on the dedicated infrastructure), which explains why the expansion of the HSR network had a huge impact on rail travel times across the territory. Figure 5 describes reductions in

---

[23]Our empirical strategy, which saturates each regressions with rich temporal fixed effects, is relatively immune to this measurement issue (see section 4).

[24]In the latter case, part of them are likely to be in fact allocated to production (but non "technical") activities, which might unfortunately generate some attenuation bias in our empirical analysis.

[25]As of today, high-speed rail service also includes cross-border services to UK, Belgium, Netherlands, Germany, Switzerland, Italy and Spain. However, this feature of the network in not taken into account in the present paper because we are not able to locate foreign headquarters (or affiliates) with our data.

[26]This appendix also contains precise descriptive statistics about reductions in travel times in our estimation sample at the dates of the main HSR line openings.

Figure 4: Evolution of the French HSR Network between 1981 (resp. 1993) and 2011
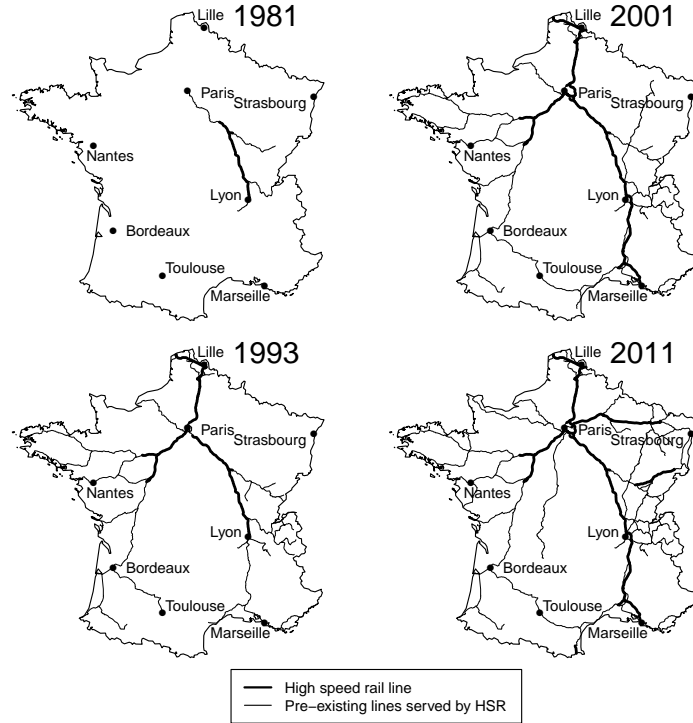
Sources: Archives and open-data platform of the French national rail company; rail fan web sites. Authors' computations.

travel times between 1993 and 2011 for selected destinations: Paris, Marseille, Lyon and Strasbourg. Between 1993 and 2011, new line openings mainly benefited the Eastern and Southern regions, which experienced the largest accessibility gains, both in terms of time and geographical range. Paris experienced the symmetric gains towards the Mediterranean zone and the Eastern zone. A few examples enable to gauge orders of magnitude: rail travel time between Marseille and Paris decreased from 6h40 to 4h40 in 1982, to 4h18 in 1994 and ultimately to 3h00 in 2001. Between Strasbourg and Paris, travel time decreased from 3h55 to 2h20 in 2007, when the Eastern line opened. Figure 5 also shows that the entire territory benefited from HSR line openings. This explains that we can always find at least one affiliate experiencing a reduction in travel time to its headquarters in any commuting zone, as shown by the maps in Appendix C.

3.4 Sample Descriptive Statistics

Table 1 provides a comprehensive description of our final sample. About 40% of the affiliates of geographically dispersed corporate groups benefit from HSR on the track to their headquarters. This high share is explained by the fact that the first HSR lines, opened between 1981 and 1993, connected particularly dense areas (Paris, Lyon, Lille). Only 4% of affiliates of geographically dispersed corporate groups benefited from the rail travel time reductions induced by the later network expansion (between 1993 and 2011). In total, 7% of domestically headquartered corporate groups were affected via at least one of their affiliates.[27] The

---

[27]Figures 10 and 8 in appendix C complement table 1 and provide a full description of the distribution of travel time reductions at the dates of the main line openings, as well as the precise geographical location of affiliates which benefited from them.

Figure 5: Reduction in Rail Travel Times to Selected Destinations
in minutes, between 1993 and 2011
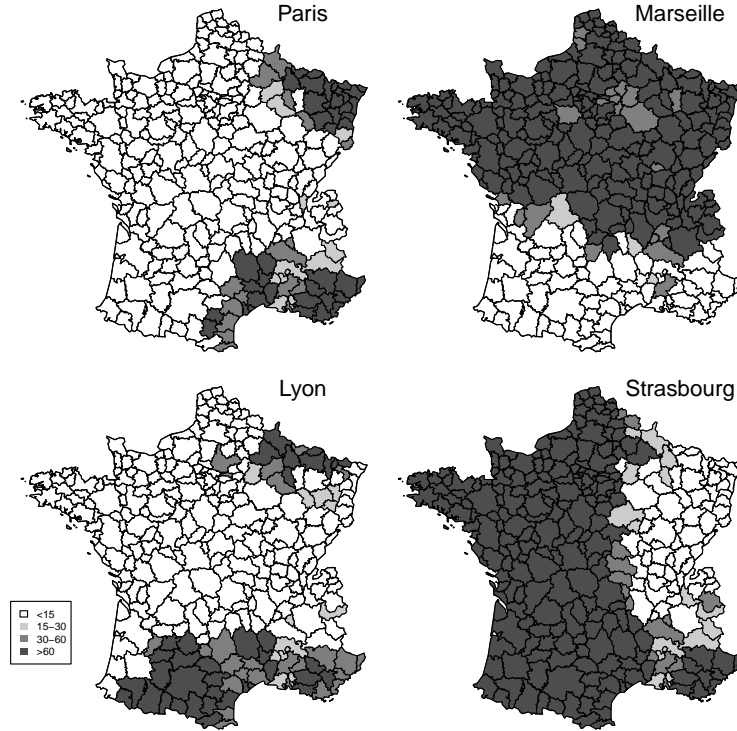
Sources: Archives and open-data platform of the French national rail company; rail fan web sites. Authors' computations.

benefiting affiliates are on average located further away from their respective headquarters than affiliates which did not experience any rail travel time reduction: this is due to the fact that HSR is typically a long distance mean of transportation (see section 3.3 and appendix E). Their headquarters are also more often located in Paris: this feature is explained by the shape of the HSR network, which is typically organized as a spider web centered on the capital city. The average rail travel speed between an affiliate and its headquarters is around 80 km/h when the HSR technology is not available, and around 110-120km/h when it is available on at least part of the track.[28] The latter value is much lower than the HSR commercial speed (320km/h), which indicates that most affiliates only benefit from HSR on a small portion of the track to their headquarters.

Table 1 also provides a breakdown of the industry structure of our sample. It shows that the sample of affiliates which benefited from the 1994 line opening was somewhat specific and presented a higher share (than the population average) of affiliates active in manufacturing and transport industries. This specificity is explained by the industrial specialization of the areas which became connected to the HSR network at that date, and translates into somewhat differentiated structural characteristics: larger affiliate sizes, and somewhat lower shares of production workers.[29]

---

[28] These travel speeds are computed as the ratio between distance "as the crow flies" between headquarters and affiliates and rail travel times.

[29] These differences might be a signal of the severity of the endogeneity problems discussed in section 4 below. Our empirical strategy, also described in section 4, is designed to address them.

Table 1: Summary statistics
Affiliates of multi-location firms, 1993-2011

	Affiliates experiencing reductions in travel time between 1993 and 2011				Affiliates never affected
	1993-2011	At main years of change:			
		1993/ 1994	2000/ 2001	2006/ 2007	
Geography:					
Distance to HQ (km)	402	475	482	326	195
HQ in Paris	40%	46%	37%	37%	34%
Travel times (hours):					
Travel time [variation $t-1/t$]	3.84 [-]	4.13 [-0.21]	4.11 [-0.69]	3.10 [-0.88]	2.06 [0]
Speed [variation $t-1/t$]	-	115 [+6]	117 [+17]	105 [+23]	95 [0]
HSR user [variation $t-1/t$, ppt]	82% [-]	100% [+7]	100% [+15]	100% [+64]	41% [0]
Workforce and wages at affiliates:					
Employment	38	40	31	31	32
Share of production workers	65%	61%	63%	66%	66%
Annual wage of prod. workers (k€)	15.5	14.8	14.5	16.9	15.9
Share of managers	6%	8%	7%	6%	7%
Industry affiliation:					
Manufacturing industries	19%	24%	17%	16%	19%
Personnel services	9%	12%	9%	7%	11%
Retail and trade	48%	35%	46%	51%	46%
Business services	17%	20%	20%	19%	18%
Transport	7%	10%	8%	6%	6%
Counts of business units:					
Nb observations (aff. × year)	198,088 [9%]	5,802	9,728	7,253	2,109,167 [91%]
Nb affiliates	28,207 [4%]	//	//	//	699,928 [96%]
Nb HQs (groups)	9,114 [7%]	2,161 [2%]	3,568 [3%]	2,429 [2%]	126,745 [93%]

Sources: Matched DADS files and LIFI survey; business units (HQs or affiliates - see section 3.1 for definition) which are part of multi-location groups between 1993 and 2011. Employment is measured in days of work but is reported in this table as headcount equivalents (days divided by 360).
Notes: The main "years of change" correspond to the opening dates of the Northern line (1993/1994), of the connection to Marseille (2000/2001) or of the Eastern line (2006/2007).

4 Empirical Strategy

4.1 Baseline Specification

Our equations of main interest are estimated at the affiliate level and describe the patterns of correlations between travel time and the various outcomes discussed in section 2, typically: affiliate employment, the share of employment allocated to production activities, or the wages of those production workers. They take the following generic specification:

$$y_{ijlt} = \beta.T_{ijlt} + \varepsilon_{ijlt} \tag{1}$$

where subscript i denotes the affiliate, j its headquarters, l denotes the commuting zone where the affiliate is located and t denotes time. Last, T_{ijlt} denotes travel time between the affiliate and its headquarters and y_{ijlt} is the outcome of interest.

We first notice that travel times between headquarters and affiliates are likely to be endogenous in these specifications, since the location of affiliates (relative to their headquarters) is probably endogenously determined by unobserved characteristics that are also correlated with the outcome variables. For example, a higher specificity of the production of the affiliate might be correlated with both a higher proximity of the considered affiliate to its headquarters, and higher wages paid out to its production workers. To address such concerns, we estimate all regressions industry by industry. We also insert affiliate level fixed effects α_{ijl} into our regression framework to control for all unobserved and non-time varying heterogeneity. Therefore, in all regressions, the relation between travel time and the various outcomes of interest is identified by *changes* in travel times (rather than by their levels), namely those generated by the expansion of the HSR network over time.

However, this regression framework remains affected by the standard but difficult problem that the placement of the HSR infrastructure might be endogenous, thus rendering variations in travel times also endogenous in our framework. One major concern is that there might exist some local unobserved heterogeneity, e.g. local growth potential, which could have determined the governmental decision to build the HSR infrastructure, and which might be also correlated with our outcome variables, thus biasing our results.[30] Furthermore, the HSR infrastructure itself might have boosted local growth,[31] a phenomenon we want to control for in our regressions. We address these two issues using an identification strategy similar to Giroud (2013) and Giroud and Mueller (2015), by introducing large sets of commuting zone × time "fixed" effects (α_{lt}), on top of the affiliate level fixed effects (α_{ijl}):

$$y_{ijlt} = \alpha_{ijl} + \alpha_{lt} + \beta.T_{ijt} + \gamma.X_{jt} + \varepsilon_{ijlt} \tag{2}$$

These many dummy variables capture the time varying *local* heterogeneity (in particular, shocks to local demand) which could generate the above endogeneity problems. Notice that all of these fixed effects are identifiable in our setting, because the impact of variation in travel times on the outcomes of interest is identified *jointly* from the locations of affiliates *and* from the location of their headquarters, i.e. travel time

---

[30] We go back to the governmental decision criteria in full details in section 4.2 below.

[31] It might also boost local prices, such as rents and the price of land. Our methodology enables to control for this aspect very accurately, especially in regressions about affiliates' creation and destruction in Appendix G in which it is a major concern.

is a *dyadic* variable. More intuitively, our baseline identification strategy amounts to use as a control group for a given affiliate, the set of local affiliates located in the *same* commuting zone, which are are therefore exposed to the same local shocks, but which are unaffected (or differently affected) by travel time reductions because of a different location of their headquarters[32]. We also provide complementary regressions specified in long differences (at either three or five years), which check that our results are not driven by further spurious short term demand shocks which could be indirectly correlated with the main explanatory variables of interest.

In terms of the estimation method, the inclusion of several sets of high-dimensional fixed effects renders estimation non-trivial, despite the fact that equation 2 is fully linear. We choose to apply the iterative procedure proposed by Guimaraes and Portugal (2010) and implemented in STATA by Correia (2014). Its principle is to iterate on sets of normal equations that are conveniently defined. The only practical constraint in our implementation of their methodology is to use *continuous* empirical proxies for the explanatory variables of main interest, in order to insure identification (see appendix F for full details).

4.2   Complementary and Alternative Identification Strategies

Our baseline identification strategy therefore takes advantage of the dyadic nature of our setting to control in a powerful way for the local shocks potentially affecting the activity of affiliates. It addresses the main endogeneity concerns while preserving a reasonable amount of statistical power. Still, a number of important threats to identification remain, which we discuss here, together with ways of addressing them.

First, our empirical setting is not robust to shocks affecting either the headquarters or the group as a whole, in cases where these shocks are temporally correlated with travel time reductions. This could happen for purely spurious reasons: in this case, it is enough to insert controls for these shocks at the group level. All of our regressions therefore include controls for group-level market conditions X_{jt}, both on the domestic and the international markets. In robustness checks, we also experiment with even more demanding specifications in which we insert a third set of high-dimensional fixed effects, either at the time $\times$ headquarters' commuting zone level, or even at the time $\times$ group level - although sample size and (the lack of) statistical power become an issue in such specifications. The estimated equations thus write:

$$y_{ijlt} = \alpha_{ijl} + \alpha_{lt} + \begin{matrix} \alpha_{l't} \\ \alpha_{jt} \end{matrix} + \beta.T_{ijt} + \gamma.X_{jt} + \varepsilon_{ijlt} \tag{3}$$

where l' denotes the commuting zone were the affiliate's headquarters are located. Equation 3 differs from equation 2 in that a third series of fixed effects is introduced into the specification: either $\alpha_{l't}$, which are the HQ commuting zone $\times$ time "fixed" effects, or α_{jt}, which denote group level $\times$ time "fixed" effects.[33]

In cases where the correlation between the group level (or headquarters level) shocks and the opening of the HSR lines are not simply spurious but endogenous, then the previous strategies are inadequate. This corresponds to the most severe threat to our identification strategy, where the set of dyads experiencing de-

---

[32]In the reported regressions in section 5, we exclude affiliates which are not part of a multi-site group, mainly for practical reasons: discarding these observations reduces drastically our file size and the associated computing times. However, in principle, they could serve as additional controls for local shocks. Un-reported regressions show that results are basically unaffected by this choice.

[33]Groups are identified by the same subscript as their headquarters (j).

creases in travel times is endogenously selected. This problem is more or less difficult to solve, depending on the causes of this endogenous selection. In the more favourable case, it is driven by only a few observations: this would happen, for example, in cases where some corporate groups had lobbied in favour of the HSR line openings and extensions that were most favourable to them (Giroud, 2013). This behaviour would produce an amplification bias in our regressions if the impact of travel time was heterogeneous across French corporate groups, and highest among those potential "lobbyists".[34] To mitigate such concern, we estimate specifications where we remove the most "suspicious" observations: either the largest affiliates in each commuting zone, or affiliates benefiting from HSR for a large portion of the rail track to their headquarters.[35] Alternatively, we also estimate specifications where we remove all groups headquartered in Paris, i.e. those that are closest to the location where governmental decisions about HSR extensions are taken.

Again, these strategies are however not sufficient in cases where the endogeneity problem affects a larger set of observations. To evaluate this risk, it is useful to analyse more precisely the decision criteria which prevailed at the governmental level when selecting *which* HSR line to build, and *when*. To that end, we rely on a very useful governmental document produced in 1991 (Ministère de l'Équipement, 1991) and endorsed by the Prime Minister of the time, which described the different lines that were envisioned at that time, together with a description of the decision criteria. Four different aspects were put forward: the rate of return of the HSR line for the national rail company (taking account of its financing capacity), the return for the entire society (which encompasses the previous criterion), the impact of the project on aspects related to spatial planning, and issues related to international policy. The latter criterion induced that projects implying a connection with neighbouring countries were prioritized. Table 2 synthesizes the first two criteria for all projects, as forecast in 1991, while figure 6 provides the corresponding maps: panel (A) shows the historical official document and panel (B) provides a more precise breakdown of the actually implemented projects (as of 2011) and the un-implemented segments. All of the actually implemented HSR lines were included in the 1991 project, but many lines foreseen in 1991 have not been constructed: for example, the lines connecting Paris to the "Centre" of the country, lines connecting the city of Bordeaux to Toulouse, Paris or Spain, *etc*. In most cases, standard rail tracks have been upgraded to become amenable to HSR trains, but at a reduced speed. Strikingly, except for the case of the three lines reported at the bottom of table 2 (TGV Auvergne, Limousin and Normandie), which exhibit exceptionally low expected profitability, actual implementation looks only weakly correlated with the rates of return that were expected in 1991. We interpret this as evidence for the major role played by the non-quantitative criteria, such as the above mentioned local and European policy concerns, together with the relative frequent project phasing and rescheduling that were imposed by budgetary constraints or un-anticipated technical difficulties.[36] These factors seem to have played a major role in determining the set of actually implemented lines and the associated dates of implementation.

---

[34] Conversely, there is no bias if the probability to be a successful "lobbyist" is uncorrelated with the magnitude of the gains expected from travel time decreases. This would happen, for example, if the chances of being a successful lobbyist are correlated with size (or the location of headquarters in Paris), but not the gains from travel time decreases (e.g. in cases where these gains are homogeneous across all corporate groups irrespective of their size, or somewhat larger in small organizations).

[35] This second strategy is very similar to what is implemented in Giroud (2013), in specifications where he only considers indirect flights where either the last leg of the flight (involving the plant's home airport) or the first leg of the flight (involving headquarters' home airport) remains unchanged.

[36] For example, in some cases, expensive art works or deviations were required by local authorities or lobbyists to preserve the environment (e.g. protected areas, wineyards, *etc.*) - see Zembri (1997).

We first propose to consider that conditionally on being part of the 1991 plan, the latter factors that have determined actual implementation are likely to be less endogenous than factors which have determined overall inclusion into the 1991 project. In other words, the above developments suggest that "by-scheme" variation (whether built or not, and the timing of the construction, conditionally on being part of the 1991 plan) is less likely to be endogenous than "on-plan" variation (whether in the original 1991 project or not). This feature can be exploited as in Donaldson (2014), by estimating our baseline specification on the sub-sample restricted to affiliates which would have benefited from the 1991 project (excluding the three least profitable lines:[37] Auvergne, Limousin and Normandie), had it been fully implemented. Alternatively, we can make the weaker assumption that only the *timing of opening* of actually built HSR lines is exogenous in our setting. This suggests to estimate our regression only on the sub-sample of affiliates benefiting at some point from decreases in travel times over our estimation period, ie. between 1993 and 2011. Last, we also propose specifications relying on an even weaker assumption: that travel time reductions were exogenous to affiliates benefiting from HSR *via* intermediate HSR stations only. Indeed, these affiliates only benefited from travel time reductions because they were located at intermediate places, which were served by HSR only because they were on the way to a larger city, as in Chandra and Thompson (2000) and Banerjee et al. (2012).[38]

4.3   Further Identification Concerns

A first concern is that the new HSR infrastructure could have been utilized for alternative usages beyond the transport of people, such that our results might be driven by mechanisms that are different from those depicted in section 2. However, this is unlikely to be a strong concern in practice. Indeed, a technical feature of the dedicated HSR infrastructure is that dedicated HSR tracks are often too steep to be accessible to freight trains.[39] This rules out the possibility that our results might be driven by standard market access mechanisms rather than by those of section 2.[40]

A more problematic technical feature of the HSR network is that it also served, together with the entire rail network of the national company, as a basis infrastructure for an optic fibre network leased to commu-

---

[37]Ministère de l'Équipement (1991) and table 2 suggest that there was no clear synthetic criterion of selection for the actually implemented HSR lines of the initial 1991 project. Therefore, it is not possible to use an identification strategy based on schemes that "just passed" or "just failed" the test. However, we chose to discard the three least profitable lines to ensure that the actually implemented lines had the same *ex ante* expected return as the non-implemented lines which we use as controls. The resulting procedure is not strictly speaking an instrumental variable strategy: indeed, in this specification, we do not argue that the placement of the 1991 plan was more exogenous than the placement of the actually implemented network, as would have been required for an instrumental variable. We only suggest that the factors determining actual implementation (i.e. the mapping between the 1991 plan and the actually implemented network) are likely to be more exogenous to our relation of interest. Conversely, we argue that the information about insertion into the 1991 plan captures the unobserved heterogeneity generating the main endogeneity problems.

[38] There is one last identification strategy that has been proposed in the literature in economic geography, that is based on instrumental variables computed from least cost paths between cities, either using straight lines as in Ghani et al. (2016), or minimized gradients as in Faber (2014). These strategies are less powerful and more difficult to implement in our setting, because in most cases, HSR lines have been built to minimize costs (while preserving speed), and the resulting network is very close to straight paths, as shown in figure 4. The few notable exceptions concern several short segments of the Mediterranean and Eastern lines, but they did not affect the location of HSR stations (and served areas) and had very low impact on rail travel times.

[39]In fact, the future line which will connect Nîmes to Montpellier in 2017 will the the first HSR to accommodate both freight and passenger traffic, first thanks its specific technical features, and second because of the lower frequency of planned HSR services which make it possible to operate train services at very different speed on the same line.

[40]Still, the new availability of the HSR infrastructure might have freed capacity for the transport of good on the standard rail network, thus having an indirect positive impact on market access. This would generate a demand shock which has no reason to be related to the decrease in travel times between an affiliate and its headquarters, and which will be captured by the α_{lt} fixed effects inserted in all our specifications.

Table 2: Synthesis of the 1991 Governmental Project

Planned HSR lines, boken-down by implementation status, as of 2011

| | | Approx. length (km) (1) | Passengers (mil.) (1991) (2) | Forecasts as of 1991 | | |
				Pass. increase (mil., %) (3)	Rate of return: Rail company (4)	Total (5)
(A) Ultimately implemented schemes						
TGV Est	Eastern Line (Paris to Strasbourg)	460	8.4	6.1 (73%)	**4.3**	**8.8**
Interconnexion Sud	Parisian inter-connection	49	12.6	0.8 (7%)	8.2	9.6
TGV Provence	Southern line (Paris to Marseille)	219	16.6	4.3 (26%)	**9.8**	**13.0**
TGV Rhin-Rhône	Rhin/Rhone connection	425	9.5	5.8 (61%)	5.9	10.7
(B) Un-implemented schemes, under construction as of 2011						
TGV Aquitaine	Paris/Tours to Bordeaux	480	14.7	5.4 (37%)	7.6	10.0
TGV Bretagne	Paris/Le Mans to Rennes	156	9.1	3.1 (34%)	7.4	**13.6**
TGV Grand-Sud	Marseille/Nîmes to Montpellier and further	70	3.7	1.6 (42%)	5.0	12.0
(C) Un-implemented schemes						
Liaison Transalpine	Lyon to Turino (Italy)	261	11.4	7.7 (68%)	6.0	10.0
TGV Côte d'Azur	Marseille to Nice	132	5.1	2.7 (53%)	**8.4**	11.0
TGV Languedoc -Roussillon	Montpellier to Perpignan	290	5.8	3.7 (65%)	6.1	9.0
TGV Midi-Pyrénées	Bordeaux to Toulouse	184	2.3	1 (44%)	5.5	**6.5**
TGV Pays de la Loire	Le Mans to Angers	78	7.5	0.9 (12%)	5.4	7.7
TGV Picardie	Paris to Amiens and the UK	165	13.5	0.9 (7%)	**4.8**	5.0
(D) Un-implemented schemes, excluded from controls						
TGV Auvergne	Paris to Clermont-Ferrand	130	2.9	1.0 (37%)	3.1	6.7
TGV Limousin	Paris to Limoges	174	3.3	0.7 (21%)	2.4	4.4
TGV Normandie	Paris to Caen and Rouen	169	5.5	1.6 (29%)	0.1	3.0

Sources: Ministère de l'Équipement (1991) (p. 13 to 15), available on line at:
http://temis.documentation.developpement-durable.gouv.fr/documents/Temis/0000/Temis-0000132/3034.pdf
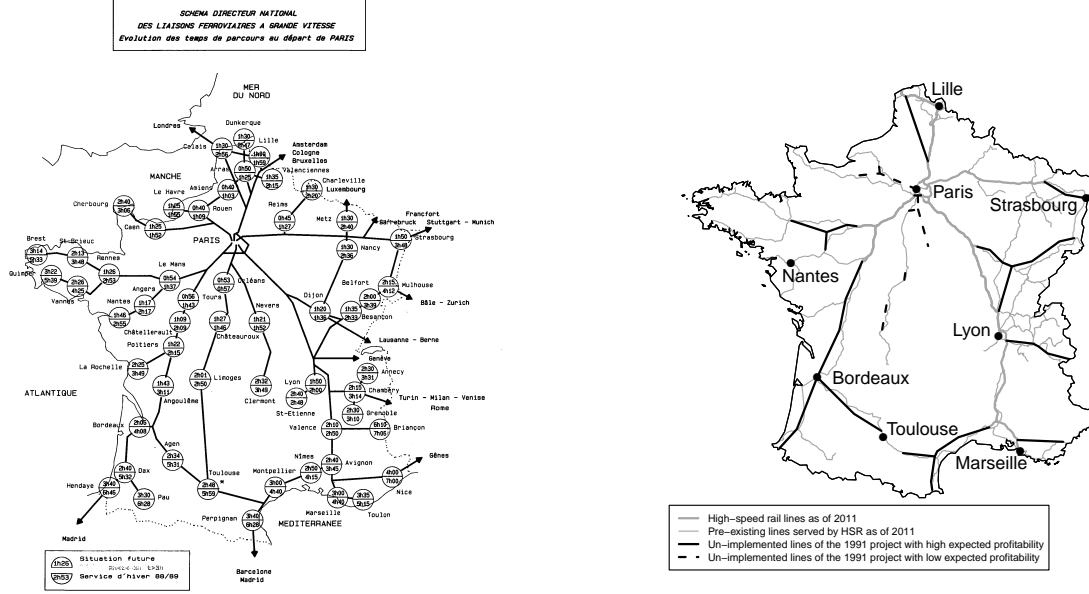
nication operators. This implies that proximity to the rail infrastructure was correlated with easier access to high-speed internet over our period of analysis, which rises the concern that our results might be driven by ICT, rather than HSR, as a communication technology. Several aspects, however, mitigate this concern. First, the HSR network extensions between 1993 and 2011 are small relative to the entire (mainly standard) rail network, which also served as a basis for the optic fiber network maintained by the rail company. Furthermore, in most cases, they duplicated a pre-existing standard rail equipment. Second, as documented in Février and Sraer (2007), several other network infrastructures actually served as alternative bases for high-speed internet: the network of the historical telecommunication operator and also waterways and highways, which cover the territory in a much denser meshing than the HSR network. This implies that the dates of opening of the HSR lines are unlikely to coincide with the new availability of ICT services.[41] These aspects will however be checked in specifications incorporating lags and leads of the travel time variable. We also propose to use specifications estimated on affiliates located relatively far away (further than 10 km) from the rail infrastructure, for which such issue of confounding factors might be somewhat less severe, or to estimate specifications restricted to the early period (between 1993 and 2003), during which high-speed internet was not yet available, as documented in Février and Sraer (2007).

---

[41]Furthermore, the new availability of ICT services has no reason to be commensurable with rail travel time reductions, as measured in hours. As in Giroud (2013), all of our results have to be interpreted as providing the additional value of face-to-face interactions in an environment where ICTs become increasingly available over time (Storper and Venables, 2004).

Figure 6: Map from the 1991 Governmental Project
Realized and un-realized HSR lines

(A) Historical document

(B) Comparison with actual lines
(as of 2011)

Sources for panel (A): Ministère de l'Équipement (1991), p. 18, available on line at:
http://temis.documentation.developpement-durable.gouv.fr/documents/Temis/0000/Temis-0000132/3034.pdf
Sources for panel (B): Archives, open-data platform of the French national rail company and complementary technical publications.

Last, a limitation of our indicator of travel time is that it relies solely on rail travel times. This might be an issue if they are not close to minimal travel times, or more generally, if managers do not travel by train, but rather by road or by air. Unfortunately, our information system does not provide us directly with this information. However, we can rely on complementary survey data (Transport and travel surveys, 1994 and 2008, Insee and SOES) to assess the magnitude of the measurement error arising from our approximation. The corresponding descriptive statistics are presented in Appendix E.[42] These analyses show, first, that rail is a popular mean of transportation for the medium and long distance travels we focus on:[43] its market share is above 50% on these segments of the transportation market. This suggests that rail travel times were of practical relevance for business trips, either directly, or at least indirectly as a benchmark.[44] Second, we inspect the expansion of the highway network over our period of analysis, together with the expansion of the airline industry for domestic travels and conclude that these two means of transportation did not generate travel time reductions that were commensurable with those induced by HSR. Altogether, the measurement

[42]This point is important, since all our specifications, which are saturated with fixed effects, require the explanatory variables to be measured accurately. It is well known that fixed effects tend to amplify the attenuation bias arising from measurement errors (Griliches and Hausman, 1986).

[43]Recall that we only consider trips between affiliates and headquarters which are located in different commuting zones. This rules out short distance travels.

[44]Notice that travel time by air (including access time to airports, which are often located in the outskirts of cities while train stations are typically located at city centres) is not lower than rail travel time for most domestic trips. See Behrens and Pels (2012) for a similar argument in the case of the London-Paris passenger market.

error generated by the fact that our indicator misses some potentially relevant travel time reductions induced by other means of transportations is likely to be relatively small, and should if anything lead to attenuation biases, preventing us from finding any impact of travel time on our outcome variables.[45]

5 Results at the Affiliate Level

5.1 Main Specification

The results for our baseline specification are reported in table 3. Columns (1) and (2) first investigate the relationship between travel time and affiliate size. The concept of size which is most consistent with the mechanisms described in section 2.1 is the production capacity of affiliates, as measured by their production workers in column (2). Consistently with the predictions of section 2.1, we obtain a strong and consistently negative correlation between the productive capacity of affiliates and travel times to their headquarters. This negative correlation holds across all industries except in the transport industry, where the obtained point estimate is incorrectly (positively) signed, but low in magnitude and statistically insignificant. Our regressions indicates that in the manufacturing industries, affiliates' size decreases by 1.9 percent when travel time to their headquarters increases by one hour. This result confirms broadly the empirical findings of Giroud (2013) and Giroud and Mueller (2015) for the case of the US manufacturing industry.[46] Similarly, we obtain that in the retail and trade industry, affiliates' size decreases by 3 percent per additional hour of travel, while the negative correlation reaches a magnitude of 5.5 percent in the personnel services industries and even 7 percent in the business services industry.

While affiliates' production capacity is negatively correlated with travel times, column (1) shows that there is no robust relation between affiliates' total employment (encompassing both production and managerial jobs) and travel times to their headquarters. This contrasting pattern is explained by the fact that production jobs and managerial jobs respond symmetrically, and with similar magnitudes, to changes in travel times, thus stabilizing affiliates' total employment. This pattern is documented in column (3), which shows that the share of production workers decreases consistently across all industries with higher travel times - while conversely, the share of managerial workers increases.[47] This is directly indicative of the relevance of the functional specialization mechanism described in section 2.2 (Duranton and Puga, 2005). Namely, affiliates which are "closer" to their headquarters and which therefore benefit more easily from the associated "headquarters" (managerial) services tend to be more specialized on their production activities. Conversely, affiliates which are located further away exhibit a lower share of production workers in their total workforce. Again, industry level heterogeneity is high: while functional specialization of affiliates on their production activities decreases by one percentage point per hour of increase in travel times in the manufacturing industries as well as in the retail and trade industries, the number reaches 1.6 percentage point in the transport industries, and around 2 percentage points in the service industries (personnel services and business services).

Strikingly, the two main empirical predictions of section 2.1 and 2.2 hold across all industries, but appear

---

[45] Appendix E discusses this aspect in detail. We also explain why we are likely to slightly over-estimate travel time variations when customers switch from air or road to HSR upon HSR line openings. This also generates attenuation biases in our setting.

[46] Giroud and Mueller (2015) adopt a specification with dummy variables which does not allow a precise comparison of estimated magnitudes.

[47] Recall that the share of managerial workers corresponds to 1 minus the share of production workers (see section 3.2).

Table 3: Travel Time to HQ and the Structure of the Workforce at Affiliates
1993-2011, affiliates of multi-site corporate groups only

	Employment (ln)		Structure of workforce (shares)		Wages (ln)	
	All workers, all skills (1)	Production workers, all skills (2)	Production workers, all skills (3)	Managerial workers, high-skilled (4)	Production workers, all skills (5)	Managerial workers, high-skilled (6)
(A) Manufacturing Industries						
Travel time	-0.013	-0.019*	-0.009***	0.002*	0.008**	-0.004
	(0.009)	(0.010)	(0.002)	(0.001)	(0.004)	(0.010)
Mean of dependent var.	60.670	48.915	0.753	0.046	18,334	45,704
Observations	426,595	408,861	426,595	426,595	408,861	191,912
(B) Personnel Services						
Travel time	0.005	-0.055***	-0.020***	0.018***	0.009	-0.065***
	(0.016)	(0.019)	(0.005)	(0.002)	(0.008)	(0.018)
Mean of dependent var.	23.089	15.835	0.733	0.047	13,204	31,628
Observations	241,846	227,936	241,846	241,846	227,936	86,833
(C) Retail and Trade						
Travel time	-0.022***	-0.030***	-0.008***	-0.002***	-0.004	-0.011**
	(0.005)	(0.006)	(0.002)	(0.001)	(0.002)	(0.005)
Mean of dependent var.	16.848	11.094	0.660	0.060	13,925	35,556
Observations	1,045,869	965,013	1,045,869	1,045,869	965,013	381,759
(D) Business Services						
Travel time	-0.009	-0.070***	-0.018***	0.000	0.013*	-0.047***
	(0.011)	(0.015)	(0.003)	(0.002)	(0.008)	(0.009)
Mean of dependent var.	45.814	29.755	0.515	0.114	19,538	44,024
Observations	401,844	329,680	401,844	401,844	329,680	210,451
(E) Transport						
Travel time	0.018	0.006	-0.016***	0.012***	0.001	0.014
	(0.016)	(0.022)	(0.005)	(0.003)	(0.009)	(0.017)
Mean of dependent var.	38.610	26.857	0.584	0.052	17,811	39,486
Observations	138,865	115,205	138,865	138,865	115,205	55,475

Sources: Matched DADS files and LIFI survey; affiliates (see definition in section 3.1) which are part of multi-site groups between 1993 and 2011.

Note: All regressions include affiliates (× headquarters) level fixed effects, as well as commuting zone × year fixed effects to control for the local market conditions. Regressions also include group level exports in total sales to capture the cycle on international markets (but all results are robust to the exclusion of these controls). Travel time is un-logged and measured in hours in all regressions. The mean of the un-logged dependent variable is reported in all columns.

to be strongest in the service industries. This finding is reminiscent of the analyses in Petersen and Rajan (2002) and Stein (2002). These articles focus on the informational problems underlying capital allocation decisions, either in the relation between banks and their borrowers (Petersen and Rajan, 2002), or in the analysis of the internal capital markets of corporate groups or multi-divisional firms (Stein, 2002). They underline the fact that the mechanisms of the type described in section 2 might be more relevant for more "informationally opaque" firms, i.e. when the information to be transmitted about the investment opportunities of remote borrowers or remote affiliates is "softer",[48] and difficult to transmit credibly, especially in the distance. Petersen and Rajan (2002) argue that this is in particular the case when the market of the

---

[48]In Petersen and Rajan (2002), smaller and younger firms are in general assumed to be more opaque, but the informational transparency of firms is more precisely proxied by the availability of detailed balance sheet and profit and losses statements, of the franchising status, the more dispersed structure of ownership of firms, and the sales region of the firm, under the assumption that the larger the firm's sales region, the more likely distant lenders will know about the firm. Only the last determinant is easily transposable in our setting.

affiliate is more "local", i.e. when market conditions are more easily assessed by local managers than by headquarters. This is precisely a feature of the service industries, in which activity furthermore relies heavily on relationship building with local clients, thus magnifying the informational advantage of local (affiliates') managers.

Column (4) of table 3 goes deeper into the analysis of affiliates' functional specialization. It shows that the share of *high-skilled* managers tends to increase disproportionately when travel times to headquarters increase. This result holds especially true in the personnel services and transport industries, where the adjustment is estimated to be 5 times larger than what would be expected if it had been homogeneous across all managerial workers (of all skills).[49] This pattern is suggestive of the fact that managerial functions which require skills and which are centralized at headquarters when affiliates are sufficiently close have to be transferred to remote affiliates when communication costs increase.

In the last two columns of table 3, we investigate the relationship between travel times and wages. Column (5) first shows the correlation between travel times and the wages of production workers. Section 2 predicted that the need for skills increases when communication costs increase. Consistently with this prediction, we obtain that the relation between the wages of production workers and travel times is most often positive, but it appears to be only significant in the manufacturing and business services industries. This feature is most likely driven by the fact that the scope for wage adjustments is limited by the regulatory minimum wage in affiliates operating in the other, rather low-skill industries, thus flattening the relation between wages paid out and travel time. Affiliates operating in the manufacturing and business services industries have a wider margin to adjust wages, simply because they are more skill-intensive (section 3.2) and pay higher wages on average. Associated magnitudes are relatively large: the average geographical dispersion within French corporate groups, requiring two hours of travel time between affiliates and headquarters, would translate into a 2% increase in production costs driven by the higher wages paid out. Last, column (6) also documents that the managers who are employed at remote affiliates are more numerous, but that they are paid less on average, thus suggesting that either the local supply of high-skilled managers is not large enough to fulfill the needs of remote affiliates, or that the content of administrative activities changes.

5.2 Economic Magnitudes: Two Illustrative Quantification Exercises

To further investigate the economic magnitude of the mechanisms underlying our regression results in table 3, we propose two different exercises. In table 4, we first propose to compute the adjustments that would be implied by a "perfect" transportation (or communication) technology which could drive all travel times to zero. These adjustments are computed in the population of affiliates that were in operation in the last year of our sample, ie in 2011. Of course, this simple partial equilibrium exercise relies on the assumption that labour supply on every local labour market is infinitely elastic, which might be more or less realistic depending onthe case (see the discussion above on the case of high-skilled managers). We first note that geographical dispersion implies an average of 2 hours of travel time between affiliates and headquarters in our sample of multi-site corporate groups, with very limited industry level heterogeneity. Given the industry

---

[49]This benchmark, "proportional" case is computed as the adjustment expected if the adjustment affecting high-skilled managers in percentage points had been proportional to their share in the managerial workforce. For example, in the personnel service industries, we get: $0.020 \times \frac{0.047}{1-0.733} \approx 0.0035$. This value has to be compared with the value of the coefficient in column (4), 0.018.

level heterogeneity across point estimates in table 3, abolishing this geographical dispersion translates into differentiated relative impact on production jobs: the average production capacity would increase by around 4 percent in the personnel services and business services industries, as well as in the transport industry, and by around 2 percent in the manufacturing, retail and trade industries. Given the heterogeneity of affiliate size, the associated impact in terms of jobs is somewhat more differentiated. It reaches a maximum of around 1.5 job per affiliate in the business services and transport industries, around 80% of a (full time) job in the manufacturing and personnel services industries, and only 25% of a job in the retail and trade industries. Overall, aggregate production jobs across all affiliates of multi-site groups would increase by 3.7% if a perfect transportation technology existed which would fully abolish travel times. This figure can be interpreted as the maximum aggregate productivity impact that can be expected from such a technology.

Table 4: Overall Cost of Geographical Dispersion as of 2011
2011, affiliates of multi-site corporate groups only

Per affiliate:		Manuf. Indus. (1)	Person. serv. (2)	Retail, trade (3)	Bus. Serv. (4)	Transp. (5)
Number of affiliates		24,445	20,134	69,634	30,040	9,147
Employment	av.	53	23	16	45	43
Share production jobs		0.783	0.712	0.709	0.528	0.587
Travel time to HQ	av. (hours)	2.096	1.981	2.152	2.281	2.302
Production jobs	jobs	0.881***	0.832***	0.253***	1.524***	1.371***
	ppt	0.018***	0.040***	0.017***	0.040***	0.036***
Manager jobs	jobs	-0.209*	-0.759***	0.076***	-0.024	-1.008***
(high-skill)	ppt	-0.004*	-0.036***	0.005***	-0.001	-0.027***
Total wage bill of	jobs	-18,193**	-4,157	1,459	-18,900*	-2,164
production workers	ppt	-0.018**	-0.018	0.008	-0.030*	-0.003

Sources: Matched DADS files and LIFI survey; affiliates (see definition in section 3.1) which are part of multi-site groups in 2011. Magnitudes are computed using estimates reported in table 3.

Table 5 then provides the estimated impact of HSR, as considered as one such communication technology. Two counterfactual exercises are proposed: the first one compares the rail network in 1980 with the rail network in 2011 and therefore computes adjustments associated with the entire existing HSR network (as of 2011). The second only considers the impact of the smaller set of lines which opened during our estimation period, i.e. between 1993 and 2011. This simple exercise remains in partial equilibrium, and it furthermore neglects the potential response of competing modes of transportation that would inevitably take place if the rail network was downgraded to its 1980 configuration. We obtain that in 2011, depending on the industry, between 40% and 50% of remote affiliates actually benefit from the HSR technology. These affiliates are typically located at more than 3.3 hours from their headquarters, and the HSR technology has allowed to save roughly 1.5 hour of travel time on average. Furthermore, and unsurprisingly given the negative patterns of correlation between affiliates size and travel times documented in section 5.1, these affiliates also appear to be slightly smaller than industry average, especially in the business services and transport industries. Overall, we obtain that in the business services industry, benefiting affiliates would discard 1.2 production jobs if the HSR network "disappeared". The magnitude is similar in the transport industry, but partially compensated by an increase in high-skilled managerial jobs by 0.75 jobs, which renders the functional shift even larger than in business services industries. Estimated results in the personnel services industry are smaller because average affiliate size is also smaller. According to our regression results, the availability of the HSR network

as a whole would have induced the transfer of roughly 0.7 jobs from administrative to production activities in these industries. In the manufacturing, retail and trade industries, and given the differences in the estimated coefficients in table 3, the magnitude of the functional shift associated with HSR ranges between 20% and 60% of a job only. Last, we obtain that total savings in terms of production labour costs induced by the HSR network are of the order of magnitude of half the cost of a production job for the average affiliate in the manufacturing and business services industries.

Overall, the implied aggregate impact of the availability of the HSR network is quite reduced, because of the relatively small proportion of affiliates benefiting from it. More precisely, our estimates imply that the impact of HSR on aggregate production jobs (of affiliates operating in multi-site groups) is around 3.1% across benefiting affiliates, but only 1.4% overall, when taking account of the fact that more than 50% of affiliates do not have access to the technology. This figure would be halved again if we took account of the single site firms and corporate groups, which by definition do not benefit from HSR for their internal communications. In other terms, the "pure productivity" impact of HSR would lead to an increase in aggregate production jobs in the for-profit sector of roughly half a percentage point.

Table 5: Organizational Impact of HSR as of 2011
2011, affiliates of multi-location corporate groups benefiting from HSR only

Per *benefiting* affiliate:			Manuf. indus. (1)	Person. serv. (2)	Retail, trade (3)	Bus. serv. (4)	Transp. (5)
Affiliates	number		10,698	8,322	30,274	15,566	4,489
benefiting from HSR	share		0.438	0.413	0.435	0.518	0.491
	employment		53	22	17	41	40
Travel time to HQ	in 2011		3.307	3.315	3.347	3.351	3.502
(average)	with 1993 network		3.721	3.732	3.775	3.792	3.956
	with 1980 network		4.673	4.801	4.714	4.861	4.996
Production jobs	2011 / 1993	jobs	0.196***	0.224***	0.057***	0.342***	0.288***
		ppt	0.004***	0.008***	0.003***	0.008***	0.007***
	2011 / 1980	jobs	0.633***	0.787***	0.212***	1.168***	1.020***
		ppt	0.012***	0.030***	0.011***	0.026***	0.024***
Manager jobs	2011 / 1993	jobs	-0.046*	-0.204***	0.017***	-0.005	-0.212***
(high-skill)		ppt	-0.001*	-0.008***	0.001***	-0.000	-0.005***
	2011 / 1980	jobs	-0.150*	-0.719***	0.064***	-0.018	-0.750***
		ppt	-0.003*	-0.027***	0.003***	-0.000	-0.017***
Total wage bill of	2011 / 1993	€	-4,224**	-1,142	327	-3,978*	-438
production workers		ppt	-0.003**	-0.004	0.002	-0.006*	0.001
	2011 / 1980	€	-13,744**	-4,108	1,234	-14,687*	-1,574
		ppt	-0.011**	-0.013	0.005	-0.020*	-0.002

Sources: Matched DADS files and LIFI survey; affiliates (see definition in section 3.1) which are part of multi-site groups in 2011 and which benefit from HSR for the travel to their headquarters. Magnitudes are computed using estimates reported in table 3.

5.3 Alternative Identification Strategies

Table 6 contains a synthesis of the results obtained when relying on the alternative or complementary identification strategies discussed in sections 4.1 and 4.2. Panel (A) first replicates the baseline estimates from table 3 as a benchmark. In panels (B) and (C), we estimate our regressions with longer time differences, by using observations collected each third year or each fifth year, respectively. The cost of this strategy is obviously the drastic decrease in sample size that is induced, with the complementary side-effect that we might

introduce a non-trivial selection by focusing only on affiliates operating for more than 3 or 5 years within the same corporate group. However, table 6 shows that our main results are preserved in this experiment: all point estimates remain correctly signed and of the same order of magnitude, although the latter tend to be slightly higher when we increase the length of the time difference. This result might be driven by the fact that the corresponding specifications are estimated on a more selected set of "successful" survivors, which might be characterized by higher responsiveness to changes in internal communication costs. While the magnitude of the coefficients of interest is globally preserved, significance however becomes an issue, especially in cases where the number of observations falls below ca. 100,000. This is the "cost" of saturating specifications with our sets of high-dimensional fixed effects (columns (1) to (5)). Overall however, this first set of robustness checks rules out the hypothesis that our main results could be solely driven by spurious short term shocks affecting both travel times and our different variables of interest. Rather, the organizational responses to shocks in travel times appear to be stable both in the medium and long run.

Panels (D) and (E) consider an alternative specification with a third series of "fixed effects": either at the level of the commuting zone where the headquarters of the considered affiliate are located, interacted with year dummies in panel (D), or directly at the level of headquarters interacted with year dummies in the specification of panel (E). Notice that since there is exactly one hearquarters unit per group, the latter specification is equivalent to a specification with group level fixed effects interacted with year dummies. Regression results are virtually unchanged in panel (D), while in panel (E), the too many fixed effects tend to hinder significance in our regressions.[50] Reassuringly however, all coefficients (except one) remain correctly signed.

Next, in panels (F) to (H), we estimate the equations of main interest on the sub-samples of observations which are less likely to suffer from the endogeneity bias potentially generated by lobbying behaviour, as discussed in section 4.2. In panel (F), we estimate a specification where we remove the largest affiliates in each commuting zone, under the rationale that the latter are most likely to be at the source of such endogeneity concerns. Obtained results are identical to those of panel (A). In panel (G), we propose another experiment which is directly inspired from Giroud (2013), and where we only use affiliates which only benefit from HSR on a small portion of the track to their headquarters,[51] since these observations are less likely to have lobbyied in favour of the HSR line than those benefiting from it on the entire track to their headquarters. Again, results are fully preserved: if anything, we obtain higher point estimates than in the baseline specification of panel (A). The same applies in panel (H), where we estimate the regressions on the sub-sample of affiliates which are *not* headquartered in the Parisian commuting zones, i.e. whose headquarters are located further away from the French political centre. Overall, the results in panels (F) to (H) suggest that lobbying concerns do not seem to be very important.

---

[50]It might be useful to provide a few orders of magnitudes. In the specification of column (1) in panel (A), the commuting zones × time fixed effects generate 5,363 categories while the affiliate level fixed effects generate 71,727 categories. Only 298 of them are redundant. In the specification of column (1) in panel (D), the headquarters' commuting zones × time fixed effects generate 5,197 additional categories, of which 303 are redundant while in panel (E), the headquarters' commuting zones × time fixed effects generate 41,415 additional categories, of which 9,347 are redundant.

[51]More precisely, we restrict the sample to affiliates benefiting from the (HSR commercial) speed of 300km per hour on less than 50% of the ditance to their headquarters. This occurs either when only a small portion of the track benefits from the HSR technology, or when the rail track is not straight (i.e. distance by rail is much higher than geographical distance "as the crow flies").

This however does not imply that our regressions are freed from even more severe endogeneity concerns, as discussed in section 4.2. In panel (I), we adopt a more differentiated identification strategy and present the results obtained by restricting the sample to affiliates which would have benefited from the 1991 plan (excluding the least profitable lines), had it been fully implemented. This implies that in such specifications, all affiliates were enclosed in the cost and benefit exercise that was performed by the French government when preparing the blueprint document. Although the mechanisms we investigate in this paper are not strong enough to be likely to have contributed very significantly to the latter,[52] it is reasonable to think that the dyads of areas that were selected to be served by the new HSR technology are likely to be more homogeneous, at least according to the qualitative and quantitative selection criteria (reported in section 4.2). If anything, the estimated impact of travel time variations should be higher in the corresponding sample of affiliates, since all projected lines were selected for their higher expected social returns. Regression results presented in panel (I) tend to confirm these intuitions: while the coefficients obtained for the correlation between travel times and functional specialization in columns (6) to (10) are very similar to those obtained in the baseline specification, the estimated correlations between travel times and production capacity are consistently higher in this sample (columns (1) to (5)).

The evidence provided in panels (J) and (K) however provides a slightly different picture and tend to suggest that part of the amplification pattern in panel (I) might in fact partly be driven by residual "by scheme" endogenous selection, i.e. by the fact that affiliates benefiting from the actually implemented lines remain different from those who were on the 1991 project, but did not benefit from the actual lines. Indeed, in panel (J), we further restrict the estimation sample to dyads of commuting zones that have ultimately benefited from HSR line openings. This implies that the estimation sample in panel (J) is a sub-sample of the estimation sample in panel (I): more precisely, all affiliates that are inserted for estimation in panel (J) ultimately benefited from HSR line openings, but at different dates. Of course, sample size becomes too small to reach significance in many significations, but strikingly, point estimates go back to their baseline value of panel (A). The same applies in panel (K), where identification relies on commuting zones which have been connected to the HSR network without intention to treat in a strategy based solely on intermediate served areas that is similar to the identification strategy of Chandra and Thompson (2000) and Banerjee et al. (2012). In this specification again, point estimates are very close to those of the baseline specification, but sample size is sometimes too small to reach significance while accommodating both the restriction to intermediate areas and the two series of high-dimensional fixed effects.[53]

---

[52]This is because the main users of HSR are households, and not firms: the "Transport and travel" survey (also used in Appendix E) indicates that in 2008, more than 65% of rail travels by high speed rail have a private purpose. Among the remaining 35%, only a subset of travels are related to communication between affiliates and headquarters: alternative business purposes would for example be related to customer, supplier or investor prospection (Bernard et al., 2015; Bernstein et al., 2015).

[53]As a reminder, we keep controlling for affiliate level unobserved heterogeneity and commuting zone specific times shocks in all specifications of table 6.

Table 6: Alternative Identification Strategies

1993-2011, affiliates of multi-location corporate groups only

Dep. var.:	Production jobs (ln)					Share of production workers				
Industry:	Manuf. indus. (1)	Person. serv. (2)	Retail trade (3)	Bus. serv. (4)	Transp. (5)	Manuf. indus. (6)	Person. serv. (7)	Retail trade (8)	Bus. serv. (9)	Transp. (10)
(A) Baseline										
Coef.	-0.019*	-0.055***	-0.030***	-0.070***	-0.006	-0.009***	-0.020***	-0.008***	-0.018***	-0.016***
Std. Dev.	(0.010)	(0.019)	(0.006)	(0.015)	(0.022)	(0.002)	(0.005)	(0.002)	(0.003)	(0.005)
Obs.	408,861	227,936	965,013	329,680	115,205	426,595	241,846	1,045,869	401,844	138,865
(B) Longer time differences: 3 years										
Coef.	-0.033*	-0.002	-0.031***	-0.086***	-0.059	-0.009**	-0.009	-0.006*	-0.016**	-0.032***
Std. Dev.	(0.019)	(0.039)	(0.011)	(0.031)	(0.041)	(0.005)	(0.011)	(0.003)	(0.006)	(0.009)
Obs.	149,773	84,409	352,632	122,227	42,324	156,189	89,653	382,129	148,698	50,795
(C) Longer time differences: 5 years										
Coef.	-0.018	-0.102	-0.044***	-0.167***	-0.109	-0.008	-0.049***	-0.002	-0.035***	-0.024*
Std. Dev.	(0.030)	(0.069)	(0.016)	(0.057)	(0.073)	(0.007)	(0.018)	(0.005)	(0.012)	(0.015)
Obs.	87,255	50,428	213,530	73,114	25,314	90,812	53,552	23,0819	90,025	30,714
(D) Additional HQ commuting zone × time "fixed" effects										
Coef.	-0.023**	-0.087***	-0.020***	-0.068***	-0.003	-0.007**	-0.019***	-0.006***	-0.016***	-0.018***
Std. Dev.	(0.011)	(0.022)	(0.006)	(0.016)	(0.024)	(0.003)	(0.006)	(0.002)	(0.003)	(0.005)
Obs.	408,861	227,936	965,013	329,680	115,205	426,595	241,846	1,045,869	401,844	138,865
(E) Additional HQ × time "fixed" effects										
Coef.	-0.029*	-0.018	-0.006	-0.009	-0.061**	-0.001	-0.006	-0.001	0.002	-0.000
Std. Dev.	(0.015)	(0.032)	(0.007)	(0.021)	(0.030)	(0.003)	(0.008)	(0.002)	(0.004)	(0.006)
Obs.	408,861	227,936	965,013	329,680	115,205	426,595	241,846	1,045,869	401,844	138,865
(F) Removing largest affiliates in each commuting zone										
Coef.	-0.018*	-0.059***	-0.030***	-0.068***	-0.006	-0.008***	-0.020***	-0.008***	-0.017***	-0.016***
Std. Dev.	(0.010)	(0.019)	(0.006)	(0.015)	(0.022)	(0.002)	(0.005)	(0.002)	(0.003)	(0.005)
Obs.	405,944	225,036	962,208	326,967	112,436	423,677	238,931	1,043,053	399,065	136,052
(G) Slowest journeys only (partial HSR service)										
Coef.	-0.026*	-0.037	-0.035***	-0.130***	-0.052	-0.014***	-0.022***	-0.008***	-0.030***	-0.024***
Std. Dev.	(0.014)	(0.026)	(0.008)	(0.021)	(0.032)	(0.003)	(0.007)	(0.002)	(0.004)	(0.007)
Obs.	354,813	197,261	850,016	273,397	97,394	369,587	209,326	917,541	334,041	117,185
(H) Removing affiliates headquartered in the Paris Area (inc. La Défense)										
Coef.	-0.038***	-0.039	-0.025***	-0.103***	-0.025	-0.012***	-0.015**	-0.008***	-0.019***	-0.018***
Std. Dev.	(0.012)	(0.027)	(0.007)	(0.017)	(0.027)	(0.003)	(0.008)	(0.002)	(0.004)	(0.006)
Obs.	256,750	128,765	684,588	198,470	71,989	268,672	137,327	735,151	249,049	86,926

Sources: Matched DADS files and LIFI survey; affiliates (see definition in section 3.1) which are part of multi-location groups between 1993 and 2011.
Note: All regressions include affiliates × headquarters level fixed effects, as well as local (commuting zone) × time fixed effects to control for the local market conditions. Regressions also include group level exports in total sales to capture the cycle on international markets. In panel (B), regressions are estimated each third year between 1993 and 2011. In panel (C), regressions are estimated each fifth year between 1996 and 2011.

Table 6: Alternative Identification Strategies, continued

1993-2011, affiliates of multi-location corporate groups only

Dep. var.:	Production jobs (ln)					Share of production workers				
Industry:	Manuf. indus. (1)	Person. serv. (2)	Retail trade (3)	Bus. serv. (4)	Transp. (5)	Manuf. indus. (6)	Person. serv. (7)	Retail trade (8)	Bus. serv. (9)	Transp. (10)
	(A) Baseline									
Coef.	-0.019*	-0.055***	-0.030***	-0.070***	-0.006	-0.009***	-0.020***	-0.008***	-0.018***	-0.016***
Std. Dev.	(0.010)	(0.019)	(0.006)	(0.015)	(0.022)	(0.002)	(0.005)	(0.002)	(0.003)	(0.005)
Obs.	408,861	227,936	965,013	329,680	115,205	426,595	241,846	1,045,869	401,844	138,865
	(I) Affiliates which would have benefited from the 1991 plan only									
Coef.	-0.065***	-0.082***	-0.079***	-0.154***	-0.033	-0.019***	-0.026***	-0.007**	-0.025***	-0.015*
Std. Dev.	(0.016)	(0.028)	(0.010)	(0.025)	(0.041)	(0.004)	(0.007)	(0.003)	(0.005)	(0.008)
Obs.	201,220	110,773	429,760	169,596	55,123	209,537	116,566	464,929	202,592	65,123
	(J) Affiliates and HQs located in dyads of commuting zones which ultimately benefited from HSR line openings									
Coef.	-0.022	-0.096***	-0.014	0.033	0.010	-0.003	-0.016*	-0.002	-0.010**	-0.014**
Std. Dev.	(0.017)	(0.034)	(0.009)	(0.023)	(0.035)	(0.004)	(0.008)	(0.002)	(0.005)	(0.007)
Obs.	77,112	39,000	201,114	81,052	24,637	81,507	41,067	220,509	98,128	30,281
	(K) Affiliates located in "intermediate" areas only									
Coef.	-0.004	-0.092***	-0.045***	-0.072***	-0.040	-0.013***	-0.018**	-0.008***	-0.008	-0.009
Std. Dev.	(0.014)	(0.028)	(0.009)	(0.023)	(0.034)	(0.003)	(0.008)	(0.002)	(0.005)	(0.008)
Obs.	192,430	92,886	450,502	133,544	48,513	199,097	98,875	486,807	163,789	57,611
	(L) Affiliates located away from train station (>10km)									
Coef.	-0.036**	-0.165***	-0.028**	-0.005	0.023	-0.001	-0.039***	-0.018***	-0.017*	-0.002
Std. Dev.	(0.018)	(0.045)	(0.013)	(0.043)	(0.049)	(0.004)	(0.012)	(0.004)	(0.010)	(0.012)
Obs.	135,254	62,886	215,740	51,202	27,415	138,702	67,420	239,803	64,597	30,746
	(M) Affiliates located in small cities (<10 million inhabitants)									
Coef.	-0.016	-0.147***	-0.031***	-0.017	-0.008	-0.010***	-0.034***	-0.012***	-0.016**	-0.020***
Std. Dev.	(0.013)	(0.033)	(0.009)	(0.027)	(0.032)	(0.003)	(0.009)	(0.003)	(0.007)	(0.007)
Obs.	217,842	94,721	367,182	93,270	50,555	223,102	100,755	402,929	112,758	55,613
	(N) Early period only (1993 to 2003)									
Coef.	-0.009	-0.047	-0.063***	0.005	-0.097***	-0.010**	-0.021**	-0.003	-0.001	-0.014*
Std. Dev.	(0.018)	(0.034)	(0.013)	(0.024)	(0.037)	(0.004)	(0.008)	(0.003)	(0.005)	(0.008)
Obs.	219,378	105,760	470,390	152,454	58,066	230,507	112,060	512,683	187,734	69,904

28

Last, in panels (L) to (N), we check whether our estimation setting might be affected by confounding factors, such as the partially simultaneous expansion of the high-speed internet network, along the lines discussed in section 4.3.[54] More precisely, our main concern here is that the HSR and rail networks have served as a basis for (part of) the high-speed internet network, such that decreases in rail travel times might capture access to new ICT technologies (visio-conferences, e-mails, etc.) rather than faster travels by rail. To mitigate this risk, we estimate our regressions on the sub-sample of affiliates which are located relatively far from the train station, such that in their case, the physical connection to the potentially confounding internet network was both more costly and technically more difficult. The obtained results in panel (L) are broadly consistent with the baseline results in industries where sample size does not drop too much (i.e. not below 100,000 observations). Results are also preserved when we focus on affiliates located in small cities, which are likely to be less prioritized in terms of internet access.

Last, panel (H) provides estimation results for the early (1993 to 2003) period only, when the high speed internet technology was not yet available.[55] Unfortunately, over this period, we can only rely on two line openings (in 1994 and 2001) with very limited "pre-treatment" or "post-treatment" years, which leaves very limited identification power in our setting. In spite of these limitations, we obtain results that are in the same order of magnitudes as our baseline estimates for the manufacturing industries, personnel services industries and retail and trade industries. Results in the trade and retail industries and in the business services industries are insignificant, probably because too few affiliates operating in these sectors benefited from the 1993/1994 and 2000/2001 line openings (respectively 2.3% and 2.9% of affiliates, to be compared with 3.6% in the personnel services, 4.3% in the transport industries, and 3.1% in the manufacturing industries).[56]

Overall, the results obtained from the different identification strategies in table 6 suggest that our baseline specification with high-dimensional fixed effects controlling for unobserved affiliate level heterogeneity and for shocks at the commuting zone level seem to be broadly adequate to address the main issues posed by the endogenous placement of the HSR infrastructure. We therefore choose to keep the corresponding baseline specification for the further extensions in sections 5.4 to 5.6.

5.4 Anticipation vs. Lagged Adjustments to HSR Line Openings

In table 7, we enrich our baseline specification with indicators of travel times measured in the year preceding and following the date of the observation to investigate potential anticipation, or conversely, lagged adjustment to shocks on internal communication costs.[57] We obtain that in the services industries (personnel services and business services) as well as in the retail and trade industries, the adjustments tend to take place upon year of change or in the year after, i.e. with a short lag. Reassuringly, the obtained point estimates are also very close to those in the baseline regression. In the manufacturing and transport industries however, the

---

[54]Notice that this problem is not too severe, since a new connection to the high-speed internet network would also generate a decrease in communication costs between affiliates and headquarters, thus *a priori* generating the same adjustments as those described in section 2.

[55] Février and Sraer (2007) show that internet broadband penetration was negligible in France in 2001, and still very limited in 2002. It later caught-up with the other OECD countries in the mid-2000's.

[56]These differentiated patterns are driven by the sectoral specialization of the areas which became connected.

[57] Notice that in our setting, which is similar to a difference-in-differences setting (across affiliates experiencing reductions in travel times upon HSR line openings and those which do not), this experiment is akin to a test of the common trend (identifying) assumption.

adjustments tend to be slightly anticipated, by one year at most.[58] More precisely, in those two sectors, we obtain that the correlation between travel time and the outcome of interest is as high (and often significant) for the current as for the anticipated indicators (at t and $t + 1$), with a cumulative value which matches the baseline estimates. This suggests that the adjustments start in the year preceding the changes in rail travel time and continue during the year in which they occur. This result actually makes sense, because adjustment costs of the production capacity in these sectors are likely to be higher than in the service and trade industries: for example, the necessary investments in equipments are likely to be planned slightly in advance.

Table 7: Anticipation *vs.* Adjustment Effects
1993-2011, affiliates of multi-location corporate groups only

Dependent variable:	Production jobs (ln)			Share of production workers			Observations
	Travel time at:			Travel time at:			
	$t - 1$	t	$t + 1$	$t - 1$	t	$t + 1$	
	(1)	(2)	(3)	(4)	(5)	(6)	
Manufacturing industries	-0.004	-0.006	-0.034**	-0.001	-0.004	-0.005	209,247
	(0.015)	(0.019)	(0.016)	(0.005)	(0.006)	(0.005)	
Personnel services	-0.062**	-0.014	-0.049	0.001	-0.016	-0.005	108,148
	(0.028)	(0.037)	(0.031)	(0.009)	(0.012)	(0.010)	
Retail and trade	-0.021**	-0.016	0.012	-0.017***	0.000	-0.001	514,465
	(0.009)	(0.011)	(0.009)	(0.003)	(0.004)	(0.003)	
Business services	-0.047*	-0.068**	-0.020	-0.013**	-0.012*	-0.001	160,585
	(0.026)	(0.031)	(0.027)	(0.006)	(0.007)	(0.006)	
Transport	-0.048	-0.040	-0.039	0.007	-0.016	-0.017**	63,856
	(0.038)	(0.044)	(0.037)	(0.009)	(0.011)	(0.009)	

Sources: Matched DADS files and LIFI survey; affiliates (see definition in section 3.1) which are part of multi-location groups between 1993 and 2011 and which were in operation for at least 3 years.
Note: All regressions include affiliate × headquarter level fixed effects, as well as local (commuting zone) × time fixed effects to control for the local market conditions. Regressions also include group level exports in total sales to capture the cycle on international markets.

5.5 Non-Linearities

In table 8, we investigate whether the actual relationship between travel times and production capacity or functional specialization might be non-linear. To that end, we split our indicator of travel time into three different classes: travel times of short travels (those lasting less than 3 hours), travel times of medium length travels (lasting between 3 and 5 hours, which would still be feasible - round trip - within one day), and travel times of longer travels (lasting more than 5 hours).

We obtain that overall, the relation between travel time and production capacity or functional specialization is low and often insignificant for the shortest travels, while it is always highly significant for medium range travels or for the longest trips.[59] There is also non-negligible industry heterogeneity in these patterns. Table 8 shows in particular that the negative correlation between travel times and production jobs

---

[58]This statement is motivated by results from un-reported regressions based on a longer series of lag or lead indicators of travel time. While this approach allows to compare the point estimates obtained for a wider range of lags and leads of the indicator of travel time, the cost is the severe decrease in sample size (due to the selection of affiliates operating over a longer time period) which decreases statistical power.

[59]These findings are consistent with the fact that the market share of HSR is highest for the longest trips, as documented in table E of appendix 12.

or functional specialization strengthens when travels are longer in the manufacturing and business services industries. More precisely, in those two sectors, the correlations are low and insignificant for the shortest travels. They become highly significant for medium range travels and the point estimates (in absolute value) are respectively multiplied by two in the manufacturing industries and by three in the business services for the longest trips. In the other industries, the pattern tends to be also increasing for the indicator of production capacity, but flatter for the indicator of functional specialization.

This contrast is suggestive of the fact that the managerial structure of business organizations is differentiated across industries, and respond differently to shocks to internal communication costs. Interestingly, industries which react most non-linearly are those which rely more on high-skill (and expensive) workers, both for production and for management (see section 3.2).

Table 8: Non-linearities in the Impact of Travel Time
1993-2011, affiliates of multi-location corporate groups only

Dependent variable:	Production jobs (ln)			Share of production workers		
	Travel time at t:			Travel time at t:		
	Shorter than 3h (1)	3h to 5h (2)	longer than 5h (3)	Shorter than 3h (4)	3h to 5h (5)	longer than 5h (6)
Manufacturing industries	0.012	-0.029*	-0.055***	-0.004	-0.009**	-0.015***
	(0.014)	(0.016)	(0.018)	(0.003)	(0.004)	(0.004)
Personnel services	-0.010	-0.059**	-0.122***	-0.008	-0.030***	-0.024**
	(0.028)	(0.027)	(0.035)	(0.007)	(0.007)	(0.009)
Retail and trade	-0.017**	-0.042***	-0.038***	-0.001	-0.020***	-0.004
	(0.008)	(0.009)	(0.011)	(0.002)	(0.003)	(0.003)
Business services	-0.005	-0.054**	-0.177***	-0.008*	-0.010**	-0.038***
	(0.022)	(0.024)	(0.025)	(0.005)	(0.005)	(0.005)
Transport	-0.009	0.043	-0.080*	-0.009	-0.023***	-0.021**
	(0.028)	(0.035)	(0.044)	(0.006)	(0.007)	(0.009)

Sources: Matched DADS files and LIFI survey; affiliates (see definition in section 3.1) which are part of multi-location groups between 1993 and 2011 and which were in operation for at least 3 years.
Note: All regressions include affiliate × headquarter level fixed effects, as well as local (commuting zone) × time fixed effects to control for the local market conditions. Regressions also include group level exports in total sales to capture the cycle on international markets. Same samples and numbers of observations as in table 3.

5.6 Spillover Effects of Travel Time to Other Affiliates

Last, in table 9, we investigate whether the adjustments described in table 3 might be affected by variations in travel time at other affiliates of the same corporate group, along the lines discussed in section 8. To that end, we insert variables describing average variations in travel time at other affiliates. We also interact this variable with an indicator of the financing constraints faced by the group, in order to check whether the spillover effects of changes in communication costs at other affiliates might be differentiated in more or less constrained groups, as in Giroud and Mueller (2015). However, since most of our corporate groups are not listed, we are not able to rely on the same indicators as these authors.[60] Rather, we compute an indicator of cash holdings which was initially proposed by Boutin et al. (2013), and which is computed as the sum of all the group subsidiaries' cash holdings, normalized by total value added. We define a group as "cash poor" if

---

[60] Giroud and Mueller (2015) rely on indicators of financing constraints that are standard in the literature: the Kaplan and Zingales (1997) and Whited and Wu (2006) indices. Both are computed from Compustat (stock exchange) information.

its cash ratio is below the sample median.

Table 9: Impact of Travel Time to Other Affiliates in Group
1993-2011, affiliates of multi-location corporate groups only

	Production jobs (ln)			Share of production workers		
	Own travel time	Δtravel time at other affiliates	Δ TT at others × cash poor group	Own travel time	Δtravel time at other affiliates	Δ TT at others × cash poor group
	(1)	(2)	(3)	(4)	(5)	(6)
Manufacturing industries	-0.020**	0.022***	0.003	-0.008***	-0.000	0.001**
	(0.010)	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)
Personnel services	-0.059***	0.019***	0.016***	-0.020***	-0.003**	0.003**
	(0.019)	(0.005)	(0.006)	(0.005)	(0.001)	(0.001)
Retail and trade	-0.040***	0.055***	-0.003	-0.008***	-0.002***	0.002***
	(0.006)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Business services	-0.081***	0.061***	0.003	-0.015***	-0.000	0.002**
	(0.016)	(0.005)	(0.006)	(0.003)	(0.001)	(0.001)
Transport	-0.008	0.021***	0.000	-0.015***	-0.001	-0.002
	(0.022)	(0.006)	(0.006)	(0.005)	(0.001)	(0.001)

Sources: Matched DADS files and LIFI survey; affiliates (see definition in section 3.1) which are part of multi-location groups between 1993 and 2011.
Note: All regressions include affiliate × headquarters level fixed effects, as well as local (commuting zone) × time fixed effects to control for the local market conditions. Regressions also include group level exports in total sales to capture the cycle on international markets. Same samples and numbers of observations as in table 3.

Columns (1) to (3) of table 9 first investigate potential spillover effects on affiliates production capacity. The correlations with own travel times are reported in column (1) and broadly confirm the baseline results of table 3. The only difference is that point estimates tend to be slightly higher in these enriched specifications. Column 2 collects the estimates obtained for the impact of variations in travel times at other affiliates in the least constrained groups. All coefficients appear to be positive and highly significant, which indicates that an affiliate benefits from higher travel times at other affiliates, by being allocated relatively more resources. Interestingly, the point estimates in column (2) are close to those in column (1), which suggests that the reallocation of production capacity occurs one to one. These patterns of active reallocation of resources towards affiliates that are relatively cheaper to operate are indicative of relatively high managerial resource constraints within *all* French corporate groups. Last, column (3) shows that more constrained groups reallocate even more - twice more - in the personnel services industries.

Columns (4) to (6) present the results obtained for the indicator of functional specialization. We had no theoretical prediction in terms of this indicator, but we obtain empirically that functional specialization tends to be slightly impaired by higher travel time at other affiliates, except in financially constrained groups. This finding might be driven by the fact that the headquarters' managerial resources of French corporate groups might be particularly scarce, even when they are financially unconstrained. It seems that in such financially unconstrained groups, the HQ managerial time that is required at other remote affiliates (featuring long travel times) is partially gathered from the HQ managerial time allocated to closer affiliates, which consequently have to rely more on their local managerial resources, thus dampening functional specialization.[61] In financially constrained groups in contrast, there is no such rebalancing of managerial resources: remote affiliates

---

[61] We will document such transfers of managerial resources between remote affiliates and headquarters in section 6.

are simply smaller, and this larger downwards size adjustment (documented in columns 1 to 3) relative to unconstrained groups allows them not to tap into the HQ managerial resources allocated to other affiliates. Correlatively, the latter do not increase their local managerial workforce (thus mitigating even further overall operating costs) and preserve their optimal functional specialization.

Appendix G provides complementary results about affiliates' entry and exit. They show that the decision to close or sell an affiliate is significantly and positively correlated with higher travel times, but negatively correlated with decreases in travel times at other affiliates, with the same order of magnitude. In short, resources spent in higher communication costs adversely affect the survival of *all* affiliates. Conversely, we show that gains in travel times at other affiliates within a group allow newly created (or purchased)affiliates to be located further away.

6 Descriptive Extensions at the Group Level

6.1 Specification

Lastly, several important theoretical predictions hold at the group level rather than at the affiliate level, most notably those relating travel time to overall group size and profitability (section 8). To investigate them, we complement our empirical analyses with regressions estimated at the level of the entire group. They take the following specification:

$$y_{jt} = \alpha_j + \sum_r \alpha_{rt}.\delta_{jrt} + \beta.T_{jt} + \gamma.X_{jt} + \varepsilon_{jt} \tag{4}$$

where (as in section 4.1) subscript j denotes the group and t denotes time. Notice that we have to introduce a new subscript r to denote the *regions* where the affiliates of the considered group are located: indeed, as we explain below, it is numerically impossible to work at the more disaggregated level of commuting zones (l) in such specifications. Analogously to equation 2, y_{jt} denotes the outcome of interest at the group level (typically, group size or profitability), T_{jt} denotes average travel time to the group's affiliates and X_{jt} are additional controls: the number of affiliates to be managed, and group level export intensity (as previously).

In equation 4, the impact of average travel time to affiliates T_{jt} on the outcome variables is still identified from *variations* in travel times (rather than from their levels), thanks to the inclusion of the group level fixed effects α_j. We also insert a set of dummy variables[62] α_{rt} capturing market conditions in the areas where the group operates. However, since corporate groups in our sample are typically highly geographically dispersed, these dummy variables are not mutually exclusive, such that it is not possible to implement the same estimation procedure as previously (appendix F). To render the estimation computationally tractable with standard OLS routines, we have to aggregate the circa 300 commuting zones into 22 French metropolitan regions.[63] Lastly, the controls X_{jt} in equation 4 include the number of sites to be managed (including the headquarters themselves, in logarithm), which implies that all results are to be interpreted as normalized

---

[62]Results are robust to the use of employment shares in each zone rather than simply dummy variables indicating where the group operates.

[63]In theory, it would possible to insert into the regressions dummy variables for different sets of affiliates implantations, interacted with time dummies (and industry dummies). In practice however, the population of French multi-site groups is too small to afford such large sets of fixed effects.

"per implantation".

Although formally very similar to equation 2, equation 4 is not as well identified. First, the aggregation of geographic fixed effects at the regional, rather than commuting zone level does not allow to control for local shocks with the same level of accuracy as in sections 4 and 5. Second and more importantly, the empirical proxy for average travel time to affiliates is not comparable either to its affiliate level counterpart. It might evolve for two different types of reasons: either because travel times decrease as the HSR network expands (as previously), or because the scope of the group evolves, i.e. because affiliates' churning. The contribution of this second component is high, and mostly unrelated to the phenomena we want to focus on - although we show in Appendix G that changes in communication costs is a significant, but small driver of affiliates' entries and exits.[64] As a first attempt to control for this aspect, we propose to split our indicator of travel time into two terms as follows:

$$
T_{jt} = \frac{\sum_{i \in \mathcal{F}_{jt}} T_{(ij)t}}{\text{Card } \mathcal{F}_{jt}} - \underbrace{\frac{\sum_{i \in \mathcal{F}_{jt_0}} T_{(ij)t_0}}{\text{Card } \mathcal{F}_{jt_0}}}_{\substack{\text{Normalization term}}}
$$

Normalization term

(factored out by group level fixed effects)

$$
= \underbrace{\left(\frac{\sum_{i \in \mathcal{F}_{jt}} T_{(ij)t}}{\text{Card } \mathcal{F}_{jt}} - \frac{\sum_{i \in \mathcal{F}_{jt}} T_{(ij)t_0}}{\text{Card } \mathcal{F}_{jt}} \right)}_{\text{(A) Change in average travel time}} + \underbrace{\left(\frac{\sum_{i \in \mathcal{F}_{jt}} T_{(ij)t_0}}{\text{Card } \mathcal{F}_{jt}} - \frac{\sum_{i \in \mathcal{F}_{jt_0}} T_{(ij)t_0}}{\text{Card } \mathcal{F}_{jt_0}} \right)}_{\text{(B) Change in the group's spatial dispersion}} \tag{5}
$$

where $\mathcal{F}_{jt}$ denotes the set of affiliates in group j at date t and $\mathcal{F}_{jt_0}$ the set of affiliates in group j in the first year where the group is observed in our data (which is denoted by t_0). As in section 4, T_{ijt} denotes travel time between affiliate i and the headquarters of group j at date t. Term (A) in equation 5 captures changes in travel times driven by the evolution of the transport infrastructure, while holding the organization of the group in the configuration observed at date t. This term is directly comparable to the indicator of travel time in section 5. Term (B) is a measure of the change in the geographical dispersion of the considered corporate group that is driven by affiliate churning, holding travel times constant at their initial value.[65] This decomposition has to be viewed as a first pass at disentangling the component of communication costs that is related to the technology of transportation from the component which is related to the geographical scope of the group. At this stage however, we lack a proper identification strategy to fully control for the endogeneity of the latter. The thoroughly rigorous treatment of this aspect is left for future research, while our ambition in this section is only to provide some descriptive evidence as a complement to our analysis at the affiliate level of section 5.

---

[64]Typically, group scope change drastically when large acquisitions or divestitures (of entire sub-groups) occur. These events are mostly related to the strategies governing the groups' portfolios of activities, and are not (primarily) responses to changes in communication costs arising from the availability of new infrastructures - although Appendix G suggests that the latter are small, but significant drivers of M&As.

[65]This convention (of measuring geographical dispersion with travel times at date t_0) insures that the geographical dispersion term does not vary when the set of affiliates is stable, which is a (very) desirable property of such an indicator. Notice furthermore that the choice of this initial date is virtually transparent in our setting, which is saturated with group level fixed effects. In un-reported robustness checks, we also experimented (without major change in results) with a version of these indicators where $T_{(ij)t_0}$ is set to $T_{(ij)93}$ for all firms, which fully harmonized the metric for both terms across all corporate groups.

6.2 Results

Results obtained for this exercise at the group level are reported in table 10. Panel (A) first investigates the correlations between travel times and different indicators of "performance". Columns (2) and (4) first show that our indicator of spatial dispersion (term (B) in equation 5) is consistently negatively correlated with total value added or employment. This implies that changes in group size as normalized by the number of affiliates, ie. "per affiliate", are consistently negatively correlated with changes in the group's spatial dispersion. The pattern of correlations is also broadly negative in the case of our indicator of travel time (term (A) in equation 5) except in the case of the business services and transport industries in this more fragile empirical setting.

One more robust result is however that all point estimates obtained for value added in column (1) are smaller than those obtained for employment in column (3). This implies that higher travel times are consistently associated with lower labour productivity in all industries, a pattern which emerges in column (5). This result was an important prediction of section 8, namely, that resources dissipated in communication costs lower overall productive efficiency. This finding is further confirmed by the results obtained with the indicator of gross operational profit in column (9), which also captures the overall productive efficiency of corporate groups.[66] Furthermore, corporate groups experiencing negative shocks on their (average) profit margin should invest less: this is indeed what we show in column (9), where investment (per affiliate) appears to be consistently negatively correlated with travel times. Overall, all of these findings broadly confirm the results in Giroud (2013), although the obtained coefficients are not always significantly different from zero in our sample.

In panel (B) of table 10, we complement the analysis by digging deeper into the analysis of the shifts of managerial jobs. Recall that we started documenting such shifts at the affiliate level in table 3, where we showed that remote affiliates tend to allocate a higher share of their workforce to managerial functions. Panel (B) of table 10 shows that this pattern is compensated by the symmetric adjustment at the head-quarters level: first, groups with higher travel times to their affiliates are characterized by fewer managers at headquarters (column (1)). Second, column (5) shows that the *share* of managerial jobs that are located at headquarters relative to affiliates is also lower when average travel times to affiliates are higher. This implies that headquarters absorb most of the overall employment cost induced by higher travel times, while affiliates' employment changes in terms of composition (fewer production workers, more managers) but its level is broadly uncorrelated with travel times (column 1 in table 3).[67] Overall, these patterns are suggestive of the functional shift of managerial jobs which occurs within multi-site groups, from headquarters to remote affiliates, when travel times increase. Though less well identified than the regressions at the affiliate level, these results tend to broadly comfort the findings of section 5.

---

[66]Interestingly, we obtain in contrast that increases in spatial dispersion (term (B) in equation 5), which a re likely to be driven by a number of unobserved factors, is associated with a positive labor productivity premium (column (6)). This premium however appears to be partially dissipated in the higher wages paid out, since the pattern of correlations becomes negative with the indicator of gross operational profit in column (8). Recall indeed that the main difference between the indicator of labour productivity in columns (5) and (6) and the indicator of operational profit margin in columns (7) and (8) lies in the treatment of wages: labour productivity is defined as value added per worker, while the operational profit margin is defined as gross operating revenue (what remains from value added after wages have been paid out) over value added.

[67]Notice however that the negative relation obtained in terms of the share of wages paid out in column (7) is somewhat attenuated, which is consistent with our earlier findings in column (6) of table 3: while managers at remote affiliates are more numerous when travel times are higher, then they ten to earn lower wages, such that their shares in total managerial wages paid out increases less - and conversely, the share accounted for by managers at headquarters decreases less.

Table 10: Spatial Dispersion and Performance at the Group Level

1993-2011, corporate groups which are geographically dispersed only

(A) Spatial Dispersion and Performance at the Group Level

	Value Added (ln)		Employment (ln)		Labor Productivity (ln)		Profit Margin		Investment (ln)	
	Travel time (A) (1)	Spatial disp. (B) (2)	Travel time (A) (3)	Spatial disp. (B) (4)	Travel time (A) (5)	Spatial disp. (B) (6)	Travel time (A) (7)	Spatial disp. (B) (8)	Travel time (A) (9)	Spatial disp. (B) (10)
Manufacturing ind.	-0.067***	-0.017***	-0.026*	-0.032***	-0.042***	0.015***	-0.008**	-0.003***	-0.037	-0.020***
(225,942 obs.)	(0.014)	(0.002)	(0.014)	(0.002)	(0.012)	(0.001)	(0.003)	(0.000)	(0.035)	(0.004)
Personnel services	-0.107***	-0.044***	-0.074***	-0.042***	-0.015	-0.003	-0.010*	-0.004***	-0.064	-0.003
(131,829 obs.)	(0.024)	(0.002)	(0.025)	(0.002)	(0.023)	(0.002)	(0.006)	(0.001)	(0.067)	(0.006)
Retail and trade	-0.049***	-0.017***	-0.049***	-0.028***	-0.001	0.011***	-0.011***	-0.004***	-0.146***	-0.013***
(406,310 obs.)	(0.013)	(0.001)	(0.012)	(0.001)	(0.012)	(0.001)	(0.003)	(0.000)	(0.034)	(0.004)
Business services	0.029**	-0.041***	0.082***	-0.059***	-0.045***	0.018***	-0.007**	-0.003***	-0.015	-0.066***
(249,263 obs.)	(0.014)	(0.002)	(0.015)	(0.002)	(0.012)	(0.002)	(0.003)	(0.000)	(0.030)	(0.004)
Transport	-0.017	-0.008**	0.040*	-0.025***	-0.046**	0.018***	-0.017***	-0.002***	-0.232***	-0.034***
(52,936 obs.)	(0.024)	(0.003)	(0.024)	(0.003)	(0.021)	(0.003)	(0.005)	(0.001)	(0.059)	(0.007)

(B) Spatial Dispersion and the Balance of Managerial Jobs at HQ *vs.* Affiliates

	Managerial jobs at HQ (ln)		Wages of managerial jobs at HQ (ln)		Share at HQ in total managerial jobs		Share at HQ in total managerial wages	
	Travel time (A) (1)	Spatial disp. (B) (2)	Travel time (A) (3)	Spatial disp. (B) (4)	Travel time (A) (5)	Spatial disp. (B) (6)	Travel time (A) (7)	Spatial disp. (B) (8)
Manufacturing ind.	-0.045**	0.004*	-0.058***	0.007***	-0.012***	0.004***	0.009	0.034***
(225,942 obs.)	(0.018)	(0.002)	(0.019)	(0.002)	(0.004)	(0.000)	(0.006)	(0.001)
Personnel services	-0.114***	-0.034***	-0.115***	-0.029***	-0.041***	-0.011***	-0.028**	0.019***
(131,829 obs.)	(0.038)	(0.004)	(0.042)	(0.004)	(0.008)	(0.001)	(0.013)	(0.002)
Retail and trade	-0.085***	-0.018***	-0.100***	-0.017***	-0.035***	-0.004***	-0.018***	0.031***
(406,310 obs.)	(0.016)	(0.002)	(0.018)	(0.002)	(0.004)	(0.000)	(0.005)	(0.001)
Business services	0.001	-0.009***	-0.009	-0.005**	-0.040***	-0.003***	-0.004	0.028***
(249,263 obs.)	(0.016)	(0.002)	(0.017)	(0.002)	(0.003)	(0.000)	(0.005)	(0.001)
Transport	0.010	0.008**	0.003	0.013***	-0.008	0.008***	-0.005	0.037***
(52,936 obs.)	(0.029)	(0.004)	(0.030)	(0.004)	(0.007)	(0.001)	(0.009)	(0.002)

Sources: Matched DADS files and LIFI survey; groups owning affiliates in different locations (commuting zones) for at least one year between 1993 and 2011.
Note: All regressions include group level fixed effects, as well as dummy variables describing in which region the group operates, interacted with years (to capture local market conditions). Regressions also include the logarithm of the total number of sites (HQ and affiliates) as well as group level exports in total sales to capture the business cycle on international markets.

36

As a last exercise in table 11, we replicate at the group level the two quantification exercises that were proposed in section 5.2. Panel (A) first proposes an estimate of the gains which could be expected from a perfect transportation technology driving all travel times to zero (i.e. driving term (A) in equation 5 to zero). The estimated impact for this thought experiment is positive in terms of employment, with an associated order of magnitude ranging between 10 and 15 jobs per corporate group (4 to 10% of their total employment) in the manufacturing industries, personnel services and retail and trade. Conversely, our estimates are negative in the business services and transport industries.

The estimated impact in terms of investment is also highly heterogeneous across industries, which mainly (and unsurprisingly) reflects the differences in capital intensities required by the production process in each industry. A more interesting result lies in the striking homogeneity of the estimated impact on the operational profit margin: we obtain that setting travel times to zero would translate into an increase of the profit margin by 1.2 to 1.6 percentage points across all industries, except the transport industry, where our estimate reaches a value of 3.4 percentage points.

Table 11: Magnitudes at the Group Level
Corporate groups which are geographically dispersed only, as of 2011

(A) Abolishing Internal Travel Times (in 2011)

Average gain per group		Manuf. Indus. (1)	Person. Serv. (2)	Retail, Trade (3)	Bus. Serv. (4)	Transp. (5)
Number of groups		5,912	3,915	12,488	10,163	1,648
Average Travel time to HQ		1.690	1.388	1.392	1.677	1.962
Employment	jobs	14.939*	15.490***	8.949***	-28.570***	-20.775*
	ppt	0.044*	0.103***	0.069***	-0.138***	-0.079*
Investment	k€	451	104	222***	92	10,324***
	ppt	0.063	0.089	0.204***	0,025	0.454***
Profit margin	ppt	0.013**	0.014*	0.016***	0.012**	0.034***

(B) Impact of HSR as of 2011

Average gain per group			Manuf. Indus. (1)	Person. Serv. (2)	Retail, Trade (3)	Bus. Serv. (4)	Transp. (5)
Nb groups	number		2,723	1,303	4,527	5,216	869
benef. from HSR	share		0.461	0.333	0.363	0.513	0.527
Average travel time	in 2011		2.544	2.447	2.242	2.406	2.858
to HQ	with 1993 network		2.811	2.704	2.494	2.710	3.232
	with 1980 network		3.384	3.287	2.966	3.402	3.972
Employment	2011 / 1993	jobs	3.082*	4.456***	2.134***	-6.193***	-4.031*
		ppt	0.007*	0.019***	0.012***	-0.025***	-0.015*
	2011 / 1980	jobs	11.031*	16.706***	7.777***	-20.174***	-14.062*
		ppt	0.022*	0.062***	0.036***	-0.082***	-0.045*
Investment	2011 / 1993	k€	114	27	51***	7	2,488***
		ppt	0.010	0.016	0.037***	0.004	0.087***
	2011 / 1980	k€	429	103	194***	36	5,136***
		ppt	0.031	0.054	0.106***	0.015	0.258***
Profit margin	2011 / 1993	ppt	0.002**	0.003*	0.003***	0.002**	0.006***
	2011 / 1980	ppt	0.006**	0.008*	0.008***	0.007**	0.019***

Sources: Matched DADS files and LIFI survey; groups owning affiliates in different locations (commuting zones) for at least one year between 1993 and 2011. Magnitudes are computed using estimates reported in table 10.

Panel (B) provides the estimated impact of the HSR network as a second quantification exercise. Unsurprisingly, reported estimates are lower, first because this alternative experiment implies lower travel time reductions, and second because only a subset or geographically dispersed corporate groups actually benefit from the HSR technology. We obtain that if multi-site groups active in 2011 *and* benefiting from the HSR network for the travels to their headquarters at this date could not rely on HSR any more for some reason and only used the 1980 (standard) rail network, then their profit margin would decrease by 0.5 percentage point on average (and even by 2 percentage points in the transport industry), which remains non-negligible. In the aggregate, the estimated impact on the benefiters reaches 0.8 percentage point, but the average impact, when also incorporating groups which do not benefit from HSR at all, is only 0.3 percentage point.

7   Conclusion

In this paper, we take advantage of decreases in travel times brought about by the availability of a new transportation technology, high speed rail, in order to document the impact of internal communication costs (in the form of travel times) on the management of multi-site business organizations. Our regressions show that higher internal communication costs are associated with reduced affiliate size and lower functional specialization, as predicted by theory, and the estimated impact is highest in the service industries, where information to be transmitted is arguably softer (Petersen and Rajan, 2002). These results are established in regressions that are specified at the affiliate level, for which identification relies heavily on the dyadic nature of our setting, where observations are "pairs" of affiliates (in one location) and headquarters (located elsewhere). This feature allows for the introduction of high dimensional location specific time fixed effects controlling accurately for the demand shocks potentially induced by the new infrastructure (Giroud, 2013), and this can be further combined with several identification strategies that have been proposed in the literature: the reliance on un-realized lines from an early governmental project as in Donaldson (2014), or the restriction to affiliates located at intermediate places, which benefited from HSR only because they are on the way to a large city as in Chandra and Thompson (2000) and Banerjee et al. (2012). Complementary but more "descriptive" regressions at the group level suggest that such internal communication costs have a non-negligible impact on the group level aggregate profits.

Our description of the nature of the managerial costs implied by the geographic dispersion of corporate groups could furthermore be enriched in several additional dimensions. First, our descriptive analyses of affiliate openings and closures in Appendix G suggest that the question of the relationship between communication costs and the decisions relating to corporate groups' overall geographic organization and outsourcing decisions is most likely a fruitful domain of research. Second, we focus in this paper on within group, headquarters to affiliates communication. Natural extensions of our work could consider communication between affiliates, and more importantly, communication with external stakeholders, such as suppliers and customers as in Bernard et al. (2015) or investors (financing relationships as in Bernstein et al. (2015), access to public programs, etc.). These aspects are likely to generate large productivity and profitability effects as well.

References

Aarland, Kristin, James C. Davis, J. Vernon Henderson, and Yukako Ono (2007), "Spatial organization of firms: the decision to split production and administration." *RAND Journal of Economics*, 38, 480–494.

Abowd, John M., Robert H. Creecy, and Francis Kramarz (2002), "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data." CES/U.S. Census Bureau Technical Paper 2002-06.

Acemoglu, Daron, Philippe Aghion, Claire Lelarge, John Van Reenen, and Fabrizio Zilibotti (2007), "Technology, Information, and the Decentralization of the Firm." *Quarterly Journal of Economics*, 122, 1759–1799.

Antràs, Pol, Luis Garicano, and Esteban Rossi-Hansberg (2006), "Offshoring in a Knowledge Economy." *Quarterly Journal of Economics*, 121, 31–77.

Antràs, Pol, Luis Garicano, and Esteban Rossi-Hansberg (2008), "Organizing Offshoring: Middle Managers and Communication Costs." In *The Organization of Firms in a Global Economy*, Harvard University Press.

Antràs, Pol and Stephen R Yeaple (2014), *Multinational Firms and the Structure of International Trade*, volume 4, 55–130.

Atalay, Enghin, Ali Hortasu, and Chad Syverson (2014), "Vertical Integration and Input Flows." *American Economic Review*, 104, 1120–48.

Aubert, Patrick and Patrick Sillard (2005), "Délocalisations et réductions d'effectifs dans l'industrie française." DESE Working Paper G2005/03.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian (2012), "On the Road: Access to Transportation Infrastructure and Economic Growth in China." NBER Working Papers 17897.

Bassanini, Andrea, Giorgio Brunello, and Eve Caroli (2017), "Not in My Community: Social Pressure and the Geography of Dismissals." *Journal of Labor Economics*, 35(2).

Behrens, Christian and Eric Pels (2012), "Intermodal competition in the London-Paris passenger market: High-Speed rail and air transport." *Journal of Urban Economics*, 71, 278–288.

Bernard, Andrew B., Andreas Moxnes, and Yukiko U. Saito (2015), "Production Networks, Geography and Firm Performance." NBER Working Papers 21082.

Bernstein, Shai, Xavier Giroud, and Rick Townsend (2015), "The Impact of Venture Capital Monitoring." *Journal of Finance*, forthcoming.

Bertrand, Marianne and Sendhil Mullainathan (2003), "Enjoying the Quiet Life? Corporate Governance and Managerial Preferences." *Journal of Political Economy*, 111, 1043–1075.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen (2012a), "Americans Do IT Better: US Multinationals and the Productivity Miracle." *American Economic Review*, 102, 167–201.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen (2012b), "The Organization of Firms Across Countries." *Quarterly Journal of Economics*, 127, 1663–1705.

Bolton, Patrick and David S. Scharfstein (1998), "Corporate Finance, the Theory of the Firm, and Organizations." *Journal of Economic Perspectives*, 12(4), 95–114.

Boutin, Xavier, Giacinta Cestone, Chiara Fumagalli, Giovanni Pica, and Nicolas Serrano-Velarde (2013), "The Deep-Pocket Effect of Internal Capital Markets." *Journal of Financial Economics*, 109, 122–145.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg (2015), "The Anatomy of French Production Hierarchies." *Journal of Political Economy*, 123, 809–852.

Cestone, Giacinta, Chiara Fumagalli, Francis Kramarz, and Giovanni Pica (2016), "Insurance Between Firms: The Role of Internal Labor Markets." CEPR Discussion paper 11336.

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

Charnoz, Pauline, Claire Lelarge, and Corentin Trevien (2016), "Communication Costs and the Internal Organization of Multi-Plant Businesses: Evidence from the Impact of the French High-Speed Rail." Insee Working Paper 2016/02.

Combes, Pierre-Philippe, Miren Lafourcade, and Thierry Mayer (2005), "The Trade-Creating Effects of Business and Social Networks: Evidence from France ." *Journal of International Economics*, 66, 1–29.

Correia, Sergio (2014), "REGHDFE: Stata Module to Perform Linear or Instrumental-Variable Regression Absorbing any Number of High-Dimensional Fixed Effects." Statistical Software Components, Boston College Department of Economics.

Cristea, Anca D. (2011), "Buyer-seller relationships in international trade: Evidence from U.S. States' exports and business-class travel." *Journal of International Economics*, 84, 207–220.

Datta, Saugato (2012), "The Impact of Improved Highways on Indian Firms." *Journal of Development Economics*, 99, 46–57.

Davis, James C. and Vernon J. Henderson (2008), "The Agglomeration of Headquarters." *Regional Science and Urban Economics*, 38, 445–460.

Donaldson, Dave (2014), "Raildroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review*, forthcoming.

Donaldson, David and Richard Hornbeck (2016), "Railroads and American Economic Growth: a "Market Access" Approach." *Quarterly Journal of Economics*, 131, 799–853.

Duranton, Gilles and Diego Puga (2005), "From Sectoral to Functional Urban Specialization." *Journal of Urban Economics*, 57, 343–370.

Emangard, Pierre-Henri and Francis Beaucire (1985), "Du bon et du mauvais usage des gares TGV dans les régions traversées." *Revue de géographie de Lyon*, 60, 359–373.

Faber, Benjamin (2014), "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System." *Review of Economic Studies*, 81, 1046–1070.

Février, Philippe and David Sraer (2007), "A Structural Model of Local Loop Unbundling." Technical report, CREST and Berkeley University mimeo, URL http://sticerd.lse.ac.uk/seminarpapers/ei10122007.pdf.

FNAUT (2011), "Les Perdants du TGV - Les effets pervers de l'exploitation du TGV sur la desserte des villes moyennes." Technical report, Rapport de la Fédération Nationale des Associations d'Usagers des Transports, co-financé par la DATAR.

Garicano, Luis (2000), "Hierarchies and the Organization of Knowledge in Production." *Journal of Political Economy*, 108, 874–904.

Garicano, Luis and Esteban Rossi-Hansberg (2006), "Organization and Inequality in a Knowledge Economy." *Quarterly Journal of Economics*, 121, 1383–1435.

Garicano, Luis and Esteban Rossi-Hansberg (2012), "Organizing Growth." *Journal of Economic Theory*, 147, 623–656.

Ghani, Ejaz G., Arti G. Goswami, and William R. Kerr (2016), "Hihgway to Success: the Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing." *Economic Journal*, 126, 317–357.

Giroud, Xavier (2013), "Proximity and Investment: Evidence from Plant-Level Data." *Quarterly Journal of Economics*, 128, 861–915.

Giroud, Xavier and Holger M. Mueller (2015), "Capital and Labor Reallocation Inside Firms." *Journal of Finance*, 70, 1767–1804.

Griliches, Zvi and Jerry Hausman (1986), "Errors in Variables in Panel Data." *Journal of Econometrics*, 31, 93–118.

Guimaraes, Paulo and Pedro Portugal (2010), "A simple feasible procedure to fit models with high-dimensional fixed effects." *Stata Journal*, 10, 628–649.

Henderson, Vernon J. and Yukako Ono (2008), "Where do Manufacturing Firms Locate their Headquarters?" *Journal of Urban Economics*, 63, 431–450.

Kalnins, Arturs and Francine Lafontaine (2013), "Too Far Away? The Effect of Distance to Headquarters on Business Establishment Performance." *American Economic Journal: Microeconomics*, 5(3), 157–79.

Kaplan, Steven N. and Luigi Zingales (1997), "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?" *Quarterly Journal of Economics*, 112, 169–215.

La Porta, Rafael, Florencio Lopez de Silanes, and Andrei Shleifer (1999), "Corporate Ownership Around the World." *Journal of Finance*, 54, 471–517.

Landier, Augustin, Vinay B. Nair, and Julie Wulf (2009), "Trade-offs in Staying Close: Corporate Decision Making and Geographic Dispersion." *Review of Financial Studies*, 22, 1119–1148.

Melitz, Marc J. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71, 1695–1725.

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

Ministère de l'Équipement, du Logement et des Transports (1991), "Schéma directeur national des liaisons ferroviaires à grande vitesse." URL http://temis.documentation.developpement-durable.gouv.fr/documents/Temis/0000/Temis-0000132/3034.pdf.

Mugele, Christian and Monika Schnitzer (2008), "Organization of multinational activities and ownership structure." *International Journal of Industrial Organization*, 26, 1274 – 1289.

Nunn, Nathan (2007), "Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade." *Quarterly Journal of Economics*, 122, 569–600.

Petersen, Mitchell A. and Raghuram G. Rajan (2002), "Does Distance Still Matter? The Information Revolution in Small Business Lending." *Journal of Finance*, 57, 2533–2570.

Picart, Claude (2004), "Le tissu productif : renouvellement à la base et stabilité au sommet." *Economie et Statistique*, 371, 89–108.

Smyth, Gordon K. (1996), "Partitioned Algorithms for Maximum Likelihood and Other Non-Linear Estimation." *Statistics and Computing*, 6, 201–216.

Stein, Jeremy C. (1997), "Internal Capital Markets and the Competition for Corporate Resources." *Journal of Finance*, 52, 111–133.

Stein, Jeremy C. (2002), "Information Production and Capital Allocation." *Journal of Finance*, 57, 1891–1921.

Storper, Michael and Anthony J. Venables (2004), "Buzz: Face-to-Face Contact and the Urban Economy." *Journal of Economic Geography*, 4, 351–370.

Strauss-Kahn, Vanessa and Xavier Vives (2009), "Why and Where do Headquarters Move?" *Regional Science and Urban Economics*, 39, 168–186.

Whited, Toni M. and Guojun Wu (2006), "Financial Constraints Risk." *Review of Financial Studies*, 19, 531–559.

Zembri, Pierre (1997), "Les fondements de la remise en cause du Schéma Directeur des liaisons ferroviaires grande vitesse : des faiblesses avant tout structurelles." *Annales de Géographie*, 593/594, 219–262.

Appendices (Not For Publication)

A Formal Version of the Mechanisms Underlying Section 2

In this section, we propose a more formal approach of the mechanisms described in section 2. The proposed set-up is a small generalization of the final good sector in Duranton and Puga (2005) which allows for a more precise analysis of the shifts in managerial employment across headquarters and production affiliates.

A.1 Set-up

Such producing firms have two facilities: a headquarter and a production affiliate. They can adopt one of two organizational forms: either spatially intergrated, or multi-location. The production technologies and profit functions of a multi-location firm are as follows:

$$y = M^\alpha (r.P)^\beta \qquad \text{with } \alpha \geq 0, \ \beta \geq 0, \ \alpha + \beta < 1 \tag{6}$$

$$M = \left((\rho.H)^{\frac{\epsilon-1}{\epsilon}} + A^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}}, \qquad \text{with } \epsilon > 0 \tag{7}$$

$$\pi = y - w_P.P - w_H.H - w_A.A \tag{8}$$

In this framework, y and π denote respectively output and profit (gross of potential fixed costs), P denotes production workers (located at affiliates) and M is the managerial input bundle. We assume it is a CES aggregate of managerial workers at headquarters H and managerial workers at affiliates A, with constant elasticity of substitution ϵ. The price of output is normalized to 1 without loss of generality, while w_P, w_H and w_A denote the wages of production workers at affiliates, managers at headquarters and managers at affiliates, respectively.[68]

Most importantly, the parameters $r \in [0; 1]$ and $\rho \in [0; 1]$ describe the costs implied by geographical dispersion:

- Moral hazard problems as in section 2.1 imply that a fraction r of productive inputs are lost or diverted.

- Transportations costs might imply iceberg costs (parameterized by ρ) for the headquarters' managerial inputs as in Duranton and Puga (2005).

In contrast, a spatially integrated firm does not suffer from such losses, but might be exposed to different prices (w_P, w_H, w_A) depending on its location choice.

A.2 Input Demand for a Multi-Location Firm

The proft maximization problem of a multi-location firm is easily solved in two steps. First, the managerial cost minimization sub-problem writes (for given levels of output y and production labor input P):

$$\min_{H,A} \{w_H.H + w_A.A\} \qquad \text{subject to } M \geq \left(\frac{y}{(rP)^\beta} \right)^{\frac{1}{\alpha}} \tag{9}$$

---

[68]Notice that this framework can be easily extended to allow for firm level heterogeneity in terms of productivity φ. Equation 6 would simply be replaced with $y = \varphi M^\alpha (r.P)^\beta$ and Y, rP and M in equations 13 to would be scaled by $\varphi^{\frac{1}{1-\alpha-\beta}}$.

where M corresponds to equation 7 above. Standard derivations imply:

$$\lambda . M = w_H . H + w_A . A \quad \text{where} \quad \lambda = \left(w_A^{1-\epsilon} + (w_H/\rho)^{1-\epsilon} \right)^{\frac{1}{1-\epsilon}} \tag{10}$$

$$\frac{\rho H}{A} \quad = \quad \left(\frac{w_A}{w_H/\rho} \right)^{\epsilon} \tag{11}$$

The conditional managerial input demand writes:

$$\frac{A}{M} = \left(\frac{\lambda}{w_A} \right)^{\epsilon} \text{ and } \frac{H}{M} = \frac{1}{\rho} . \left(\frac{\lambda}{w_H/\rho} \right)^{\epsilon} \tag{12}$$

Second, the global problem can be easily solved in P and M. Standard derivations imply:

$$y = \left(\frac{\beta}{w_P/r} \right)^{\frac{\beta}{1-\alpha-\beta}} . \left(\frac{\alpha}{\lambda} \right)^{\frac{\alpha}{1-\alpha-\beta}} \tag{13}$$

$$P = \frac{\beta}{w_P} . y \text{ and } M = \frac{\alpha}{\lambda} . y \tag{14}$$

A.3 Comparative Statics for a Multi-Location Firm

Assuming away general equilibrium effects on relative prices,[69] the theoretical predictions of section 2 correspond simply to the comparative statics exercise with respect to parameters r and ρ in the problem of the multi-site firm. Our main quantities of interest are y and P (equations 13 and 14), $\frac{H}{M}$ (equation 12) and the following additional indicator of affiliates' functional specialization:

$$\frac{P}{P+A} = \frac{\beta.\lambda}{\beta.\lambda + \alpha.w_P.\left(\frac{\lambda}{w_A} \right)^{\epsilon}} \tag{15}$$

First, the comparative statics with respect to r are relatively straightforward. A decrease in moral hazard problems leading to an increase in r leads to higher production worker demand and higher output:

$$\frac{\partial y/y}{\partial r/r} = \frac{\partial P/P}{\partial r/r} = \frac{\beta}{1-\alpha-\beta} > 0 \tag{16}$$

In contrast, the split of managerial workers between headquarters and affiliates is unaffected by a shock on r since the share of managers located at headquarters is independent of r (equations 12 and 15).

The comparative statics with respect to ρ show that both output and structural ratios are affected:

$$\frac{\partial y/y}{\partial \rho/\rho} = \frac{\partial P/P}{\partial \rho/\rho} = \frac{w_P}{\lambda} . \frac{\alpha}{1-\alpha-\beta} . \frac{H}{M} > 0 \tag{17}$$

$$\frac{\partial P/(P+A)}{\partial \rho} . \frac{\rho}{P/(P+A)} = (\epsilon - 1). \left(\frac{\lambda}{w_H/\rho} \right)^{\epsilon-1} \frac{\beta.\alpha.w_P.(\lambda/w_A)^{\epsilon}}{\beta.\lambda + \alpha.w_P.(\lambda/w_A)^{\epsilon}} \text{ positive iff } \epsilon > 1 \tag{18}$$

$$\frac{\partial H/M}{\partial \rho} \frac{\rho}{H/M} = \epsilon. \left(1 - \frac{w_H H}{\lambda M} \right) - 1 > 0 \text{ positive iff } \epsilon > 1 + \frac{w_H H}{w_A A} \tag{19}$$

---

[69]Report to Duranton and Puga (2005) for a thorough general equilibrium analysis.

B   Remote Corporate Control in France: the Disproportionate Weight of Paris

Figure 7 provides a break-down of the indicator of figure 2 (section 3.1) according to different "centres" of corporate control. It describes for each commuting zone, the share of employment under control of HQs located respectively in Paris, Lyon, Marseille or Strasbourg.

The main take-away from the comparison of the four panels of figure 7 is that the capital city, Paris, appears to have a disproportionate weight in terms of corporate control. Its sphere of "corporate influence" is particularly wide ranging, since most of the territory is reached by Parisian headquarters. Lyon and Marseille also reach very distant areas, but only occasionally and with a lower weight in terms of local employment.

Figure 7: Sphere of "Corporate Influence" of 4 of the Largest French Cities, in 2011

Share of private employment in each zone that is under control of HQs located in the respective cities in percentage of total private employment (except agriculture and personnel services)
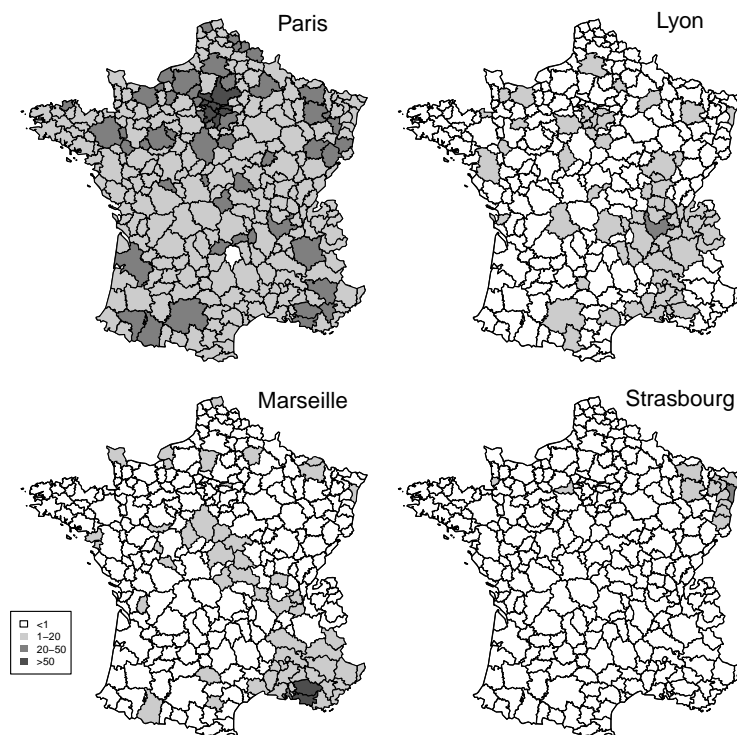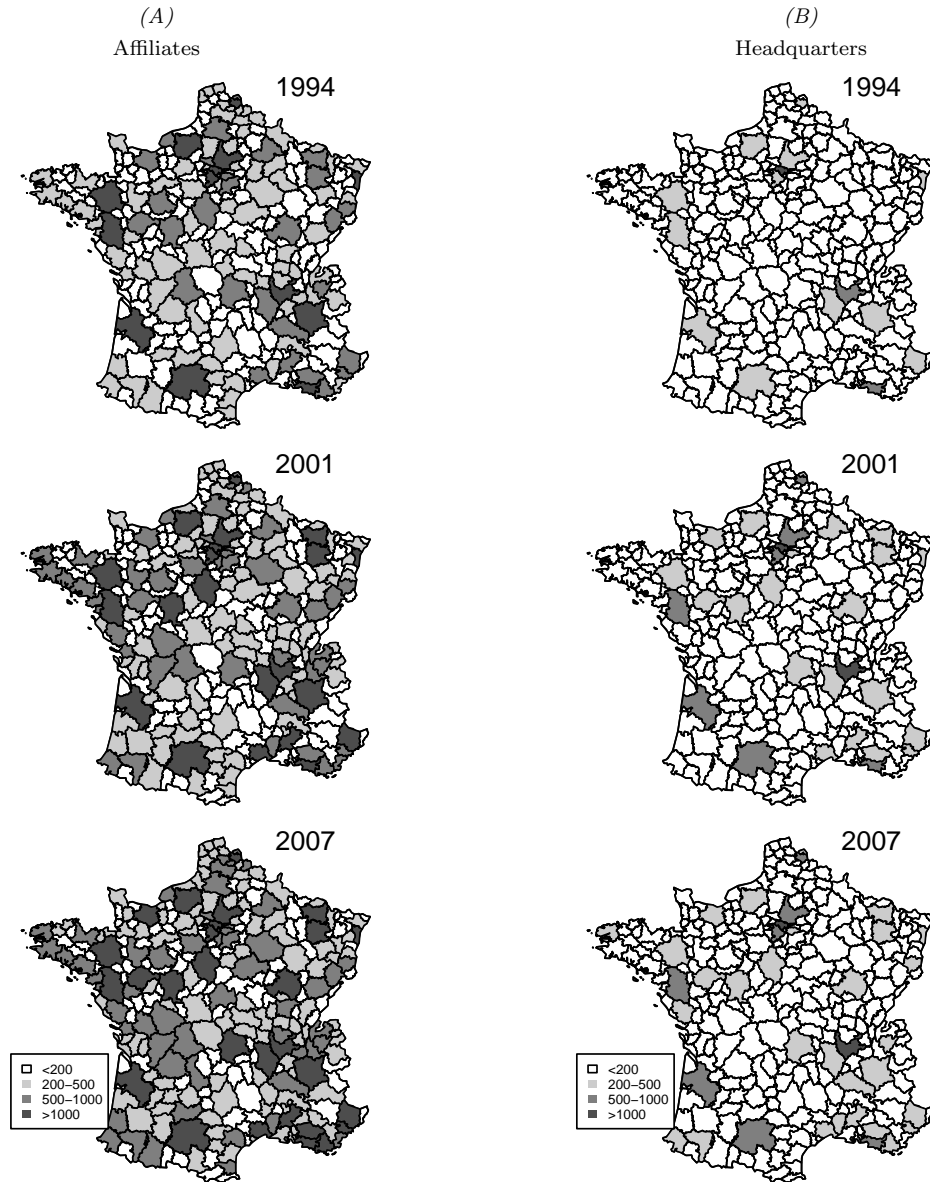

Sources: Matched DADS files and LIFI survey, covering the private sector (except agricultural activities and workers of the personnel service industries directly employed by households).
Notes: the four maps describe the share of private employment in each commuting zone that is under control of HQs located, respectively, in Paris, Lyon, Marseille or Strasbourg.

44

C   Geographical Repartition of the Estimation Sample

Figure 8: Localization of Affiliates and Headquarters of Multi-Site Business Organizations
Estimation Sample (Entire Population)

(A)
Affiliates

(B)
Headquarters

Sources: Matched DADS files and LIFI survey; business units (HQs or affiliates - see section 3.1 for definition) which are part of multi-location firms between 1993 and 2011.
Note: The different dates correspond to the opening dates of the Northern line (1993/1994), of the connection to Marseille (2000/2001) and to the opening date of the Eastern line (2006/2007).

Figure 9: Localization of Affiliates and Headquarters Benefiting from Rail Travel Reductions
At Dates of Main HSR Line Openings
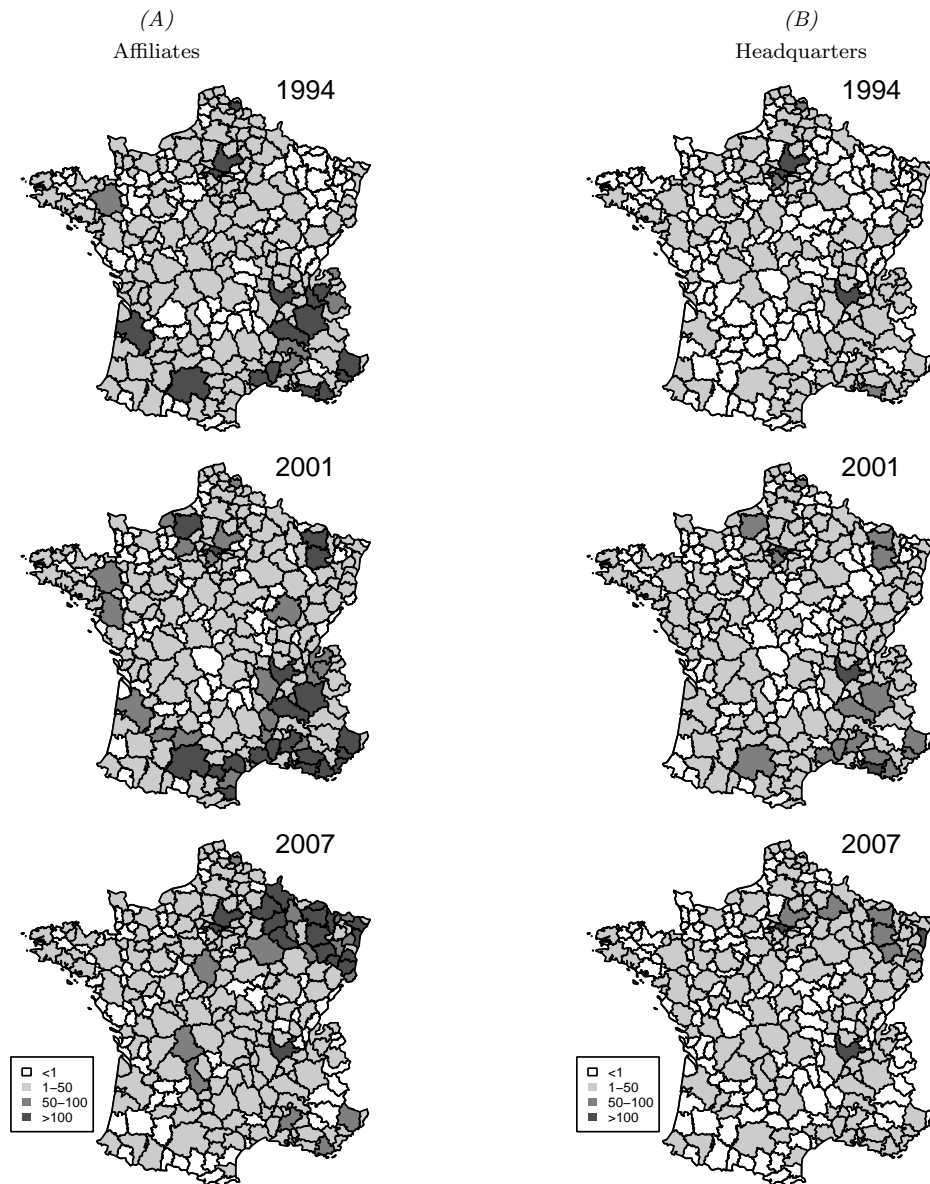
Sources: Matched DADS files and LIFI survey; business units (HQs or affiliates - see section 3.1 for definition) which are part of multi-location firms between 1993 and 2011.

D Computation of Rail Travel Times

This section describes the computation of rail travel times between headquarters and affiliates. We rely on a two-step procedure. First, we simplify the network of trains stations and select only one "main station" by commuting zone (where either affiliates or headquarters are located). Second, we collect (or construct) time tables for the resulting list of 316 such "main stations". Rail travel time between any affiliate and its headquarters is then measured by rail travel time between the respective "main stations" of the commuting zones where each are located.

Notice that travel times between headquarters or affiliates and the "main station" in their commuting zone are not taken into account for two reasons: first, they are typically short, and second, they do not vary over time (while our identification strategy typically relies on *variations* in travel times - see section 4). Furthermore, managers who are asked to travel for professional reasons might depart from home rather than from work (see the discussion in section 3.1): in absence of precise information about the location where managers live, and from which station they might depart, the main station in each commuting zones is by construction the best statistical guess we can get about it.[70]

D.1 Selection of the "Main Station" in Each Commuting Zone

We first select a set of 316 "main stations", among the set of train stations which were in operation in 1993, using the following criteria:

- In cases where only one station in the considered commuting zone is served by HSR, we select it as its "main station".

- In cases where several stations in the considered commuting zone are served by HSR, we select the station having the highest volume of long distance traffic using a score computed from the number of long distance services in each station.[71]

- In cases where none of the stations in the considered commuting zone is directly served by HSR, we select the station having the highest long distance traffic score.

Stations that were specifically built to accommodate HSR services during our period of estimation replace the previous "main station" in their respective commuting zones from the date they enter into service.

D.2 Time Tables Between "Main Stations"

We then collected historical as well as current timetables, in order to recover the fastest train service between any two directly connected "main stations":

- Travel times for train services which remained "local" over the entire period (i.e. those implying no HSR service) were simply approximated by the 2013 timetables, which are available on the open-data platform of the national rail company.

---

[70]In a small number of cases, the "main station" of the commuting zone is not the closest to the considered headquarters or affiliate. We checked in a previous version of our work (Charnoz et al., 2016) that this does not affect our empirical analysis.

[71]These scores are computed as a weighted number of long distance services available in each station, as of 2013 (for which precise information is exhaustively available). The weight of each service in this sum corresponds to the squared average distance between any two consecutive stops in this service. In the rare cases of score ties, we select the station located in the most populated municipality.

- For long distance services, we relied on rail fan web sites and the archives of the national rail company, as well as on the evaluation reports of the French Ministry for Transportation (LOTI reports). These sources enable us to assess rail travel times both *before* and *after* HSR line openings.

The obtained database contains an indicator of rail travel time between any two directly connected stations, at any date between 1980 and 2013. Rail travel times between any two *indirectly* connected stations are then computed as the shortest path between directly connected stations, under the assumption that any train change takes 15 minutes (which is a rather lower bound).

D.3   Descriptive Statistics

Figure 10 provides the full distibution of changes in our indicator of rail travel time in our estimation sample, at the dates of main HSR line openings. It complements figure 9 which described the geographical localization of the corresponding units. The opening of the Northern line in 1994 generated gains in travel times that were most often limited: the mass of the distribution is concentrated on values between 10 and 20 minutes. In contrast, the connection to Marseille in 2001 generated gains ranging between 50 and 65 minutes for two third of the benefiting affiliates. Likewise, the opening of the Eastern line in 2007 also generated large gains, staggered between 60 and 110 minutes, for two third of the beneficiaries.

Figure 10: Distribution of Travel Time Variations at Main Dates of HSR Line Openings
Northern line (1993/94), connection to Marseille (2000/01) and Eastern line (2006/07)
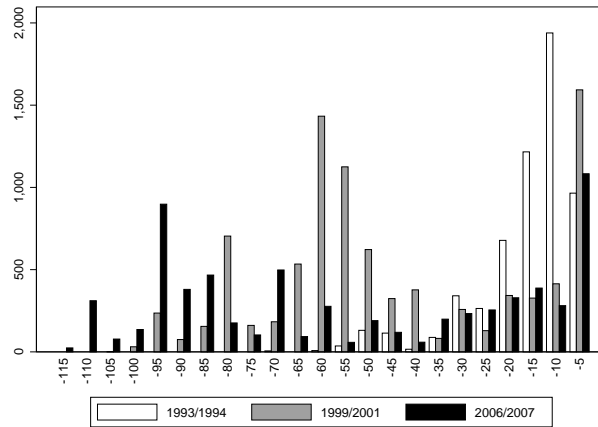
Sources: Matched DADS files and LIFI survey; affiliates (see definition in section 3.1) which are part of multi-site business organizations between 1993 and 2011.
Note: Values are expressed in minutes. Firms which did not experience any change in rail travel time at the respective dates were excluded.

E The French Market for Domestic Travels from 1993 to 2011

A limitation of our indicator of travel time is that it relies solely on rail travel times across commuting zones (see appendix D). We argue in section 4.3 that this should if anything produce an attenuation bias in our estimations. This Appendix provides complementary evidence for the different arguments put forward.

E.1 Market Shares for Business Travels: Short versus Long Distance

A first strategy to check whether rail travel times are a good approximation of actual travel times across commuting zones (both of their levels and variations) is to inspect the market shares of the differents modes of transportation, for various lengths of business trips. Unfortunately, we lack the direct information about which mean of communication, or which mode of transportation is used by managers to travel between affiliates and headquarters in our sample. However, we can rely on complementary, survey data to estimate aggregate usage. Table 12 is constructed from two different surveys, which were conducted by the French Statistical Institute and the French Ministry of Transport, respectively at the beginning (1994) and at the end (2008) of our period of analysis. This table first provides a description of the markets shares of road, air or rail for domestic business trips, with a break-down by categories of trip length. It shows that as of 2008, the overall market share of rail raises to 38% of business trips (46% when weighted by distance). This market share is much lower (23%) for short distance trips, which are overwhelmingly realized by road (77%), but it raises to 55% for medium and long distance trips. This suggests that our indicator of rail travel time *in levels* is probably fairly relevant for medium and long distance trips, but less relevant for short distance travels, for which information about road travel times would have been more appropriate. Fortunately, table 1 documents that given our definition of geographical dispersion for corporate groups, which requires affiliates and HQs to be located in different commuting zones (section 3.1), most of the trips between HQs and affiliates in geographically dispersed groups would be classified as "medium distance".[72]

Table 12: "Market Share" of Each Mode of Transportation by Market Segment
Domestic Business Travels Only

Distance:	Market shares in 1994					Market shares in 2008				
	Short (< 200km)	Medium (200 to 800km)	Long (> 800km)	All	All	Short (< 200km)	Medium (200 to 800km)	Long (> 800km)	All	All
Weight:	trips	trips	trips	trips	km	trips	trips	trips	trips	km
Road (car and bus)	78	64	24	70	58	77	39	15	57	44
Rail	21	30	21	24	26	23	55	52	38	46
of which HSR:	*0*	*19*	*12*	*8*	*14*	*2*	*24*	*42*	*13*	*26*
Air	0	6	43	5	13	0	6	33	4	10
Unanswered	1	1	11	2	3	0	2	0	1	1
Total	100	100	100	100	100	100	100	100	100	100

Sources: Insee, Transport and communication survey, 1994, and SOES, Transport and travel survey, 2008.

E.2 Expansions of the Networks of Alternative Transport Infrastructures

In our empirical setting, *variations* in travel are more relevant than their *levels* (section 4). Whether our indicator of rail travel time captures them accurately depends on two main (and correlated) factors: first,

---

[72] Average distance to HQ for affiliates never experiencing decreases in rail travel times is as high as 195km, and even 402km for those benefiting from the HSR network expansion.

whether travel times varied for other means of transport, and second, whether business travellers' choice of mode of transportation changed over time. The first aspect can be assessed by an analysis of the supply side for each transportation mode. The case of rail is described in the main text and in appendix D: travel times reductions result mostly from the expansion of the HSR network depicted on figure 4 and from the fact than when the dedicated infrastructure becomes available, speed increases from ca. 100-150 to ca. 300 km per hour.

The case of road can be analysed in the same way, thanks to figure 11. This map shows the expansion of the French highway network over our period of analysis. The network expansion was significant, especially in the western part of the country,[73] but not as large as the HSR network expansion. Moreover, speed increases brought about by the new availability of highways are more limited than in the case of HSR, since maximum speed on the standard road infrastructure is 90 km per hour (outside urban areas), while maximum speed on highways is 130 km per hour.

Figure 11: The French Highway Network between 1990 and 2010

(A) 1990/2000

(B) 2000/2010
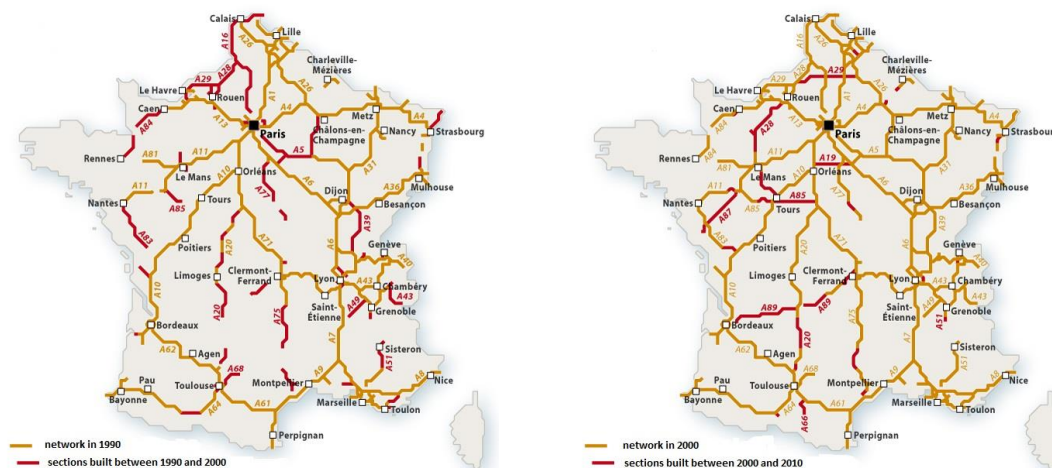
Sources: *La Documentation Française*, *Le Moniteur des Travaux Publics*, ADF, CNA, Highway Concession Companies.

The case of airlines is somewhat more difficult, because we were not able to gather precise historical information about line openings and closures as in Giroud (2013). Figure 12 however describes the overall evolution of the industry, as compared with rail. The liberalization episode of the late 1990's, which witnessed the entry and later exit of several airlines on the domestic market, is detectable in the aggregate data: the number of passengers-km increased between 1994 and 2001, and subsequently decreased between 2000 and 2004. However, the implied magnitudes appear to be really small in comparison with the increase in the market share of high speed rail. This is reassuring about the potential impact of these unmeasured adjustments of the airline market on our analysis.

---

[73]The main expansions were the connection between Bordeaux and Clermont-Ferrand, and the connections between Nantes, Le Mans, Tours and Poitiers. Interestingly, highway expansions appear to be geographically disconnected from HSR network expansions, which mitigates the risk that our indicator of rail travel times captures unintended confounding phenomena.

Figure 12: Market Shares of Rail and Air for the Passenger Market
All Travels (Business and Non-Business), 1993 - 2011
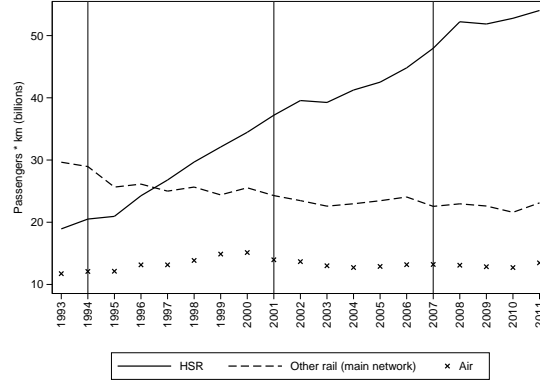
Sources: compiled by SOeS - Ministry in charge of Transports from rail operators and DGAC. The vertical bars correspond to the dates of the main HSR line openings.
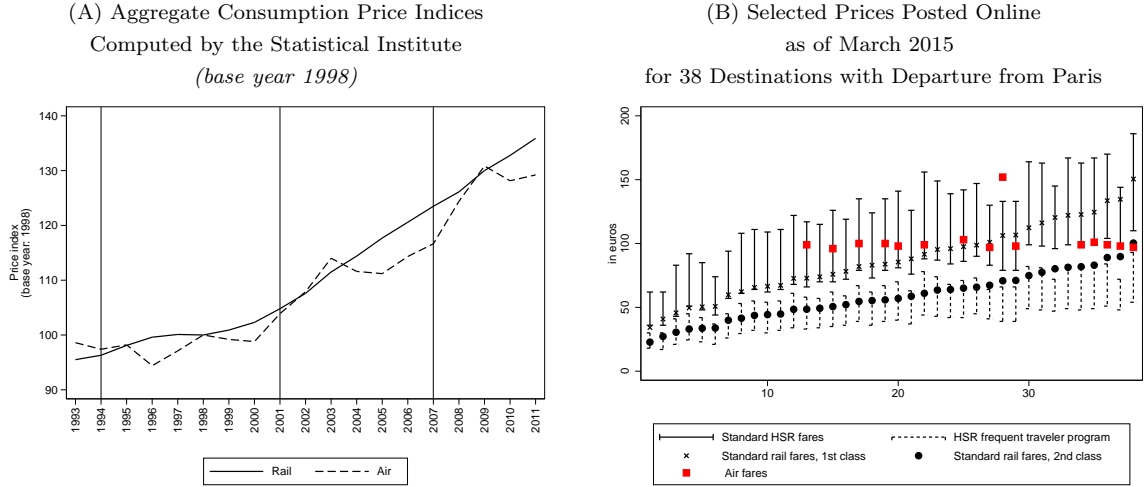
E.3   Discussion of Biases

Overall, the above discussion together with table 12 and figure 12 suggest that most of the variations in medium and long travel times were generated by the HSR network expansion, such that *variation* of our indicator of travel times is less likely to be affected by measurement error than its *level*. A second and related insight is that rail gained market shares over other means of transportation during our period of analysis. Table 12 shows that for medium distance business travels, the market share of rail increased by 25 percentage point, from 30% to 55%. For long distance business travels, the increase is even larger, from 21% to 52%. While this fact is reassuring about the relevance of our indicator of rail travel times as an indicator of overall travel time,[74] the gains in market shares also indicate that our indicator of rail travel times might in fact over-estimate travel time variations. To understand this, it is useful to go back to the discrete choice problem faced by business travellers when choosing their mode of transportation. Prices and travel times are likely to be the main drivers of this choice.[75] Figure 13 plots the evolution of the price index for both rail and air travels. These two indices exhibit a similar evolution, which suggests that pricing strategies have been very similar over the period. Therefore, variations in travel times are likely to have been the main determinants of the choice of switchers explaining the increase in the market share of rail (against road and air). These travellers are likely to have switched from a mean of transportation that was faster than standard rail (but slower than HSR), towards rail upon HSR line openings. As a consequence, for these travellers, variations in *rail* travel times will over-estimate their true variation in travel times.

If anything, over-estimating variations in travel times should bias our estimation results *downwards*, since

---

[74]The 2008 edition of the Transport and travel survey also enables to estimate that in 2008, 60% of HSR business travelers are "managers", as defined in section 3.2.

[75]The other dimensions of transportation services are likely to play in favor of HSR, since HSR line openings were almost always associated with improved rail service beyond travel time. For example, new and more comfortable coaches were most of the time introduced, with increased "workability" and a higher frequency of train services, at least for terminal cities. This last point is however more debated for the case of certain smaller or middle size cities, more precisely those that are bypassed by HSR services while they were previously served by traditional rail service (FNAUT, 2011; Emangard and Beaucire, 1985).

Figure 13: Prices of Transport by Rail and Air

(A) Aggregate Consumption Price Indices
Computed by the Statistical Institute
(base year 1998)

(B) Selected Prices Posted Online
as of March 2015
for 38 Destinations with Departure from Paris

Sources: Panel (A): French National Statistical Institute (Insee).
Panel (B): Online available prices as of March 2015, for 38 selected destinations (with departure from Paris) sorted by distance: Reims, Arras, Le Mans, Tours (St-Pierre-Des-Corps), Lille, Valenciennes, Dijon, Angers St-Laud, Metz, Poitiers, Nancy, Mâcon-Loché TGV, Rennes, Besanon-Viotte, Nantes, Niort, Lyon Part-Dieu, Angoulême, Lyon St-Exupéry TGV, Strasbourg, La Rochelle, Mulhouse, Valence TGV, Chambéry Challes-les-Eaux, Annecy, Grenoble, Bordeaux St-Jean, Lorient, Brest, Avignon TGV, Nîmes, Dax, Aix-en-Provence TGV, Montpellier, Marseille St-Charles, Toulon, Toulouse-Matabiau, Nice. Air fares are retrieved from the website of Air France (Hop!), the leading airline company.

we will relate a change in our various outcome variable to a change in our explanatory variable that is too large. Moreover, overlooking changes in travel times by air or road (which should occur mainly for the shortest trips, as argued above in section E.1) will also produce attenuation biases in our setting.[76]

---

[76]In theory, amplification bias could occur in cases where changes in travel times by air or road are temporally correlated with reductions in rail travel times and larger than the latter. This is however unlikely to happen frequently because section E.2 documented that the timing of the liberalization episode for airlines did not match the timing pattern of HSR line openings. Similarly, the expansion of the highways network appears to be geographically disconnected from the HSR network expansion.

F   Details of the Estimation Method

The principle of the estimation method proposed by Guimaraes and Portugal (2010) and Correia (2014) is to iterate on three sets of normal equations that are conveniently defined. Let us first rewrite equation 2 in matrix format as:

$$\underbrace{Y}_{y_{ijlt}} = \underbrace{X\ B}_{\beta.T_{ijt}+\gamma.X_{(ij)lt}} + \underbrace{D_I\ A^I}_{\alpha_{ij}} + \underbrace{D_{LT}\ A^{LT}}_{\alpha_{lt}} + \underbrace{\epsilon}_{\varepsilon_{ijlt}} \tag{20}$$

where X is a vector encompassing our *continuous* treatment variable and possibly additional explanatory variables (in our setting, group level controls such as its exposure to international demand conditions, i.e. export intensity), D_I is the vector of the affiliate level dummies (α_i), D_{LT} the vector of the commuting zone $\times$ year dummies (α_{lt}), B, A^I and A^{LT} are the corresponding parameters to be estimated. Explanatory variables of interest that are categorical variables might be more difficult to identifiy given the high number of "fixed effects" included in equation 20.[77]

The three sets of normal equations are defined as follows:

$$\begin{pmatrix} B = (X'X)^{-1}X'(Y - D_I\ A^I - D_{LT}\ A^{LT}) \\ A^I = (D'_I D_I)^{-1}D'_I(Y - X\ B - D_{LT}\ A^{LT}) \\ A^{LT} = (D'_{LT}D_{LT})^{-1}D'_{LT}(Y - X\ B - D_I\ A^I) \end{pmatrix} \tag{21}$$

The algorithm is initiated at $A^I_{(0)} = 0$ and $A^{LT}_{(0)} = 0$. The first of equation provides the first estimated value for $B_{(1)}$, which is plugged into the second set of equations to get $A^I_{(1)} = A^I_{(1)}\left(B_{(1)}, A^{LT}_{(0)}\right)$. Then $B_{(1)}$ and $A^I_{(1)}$ are plugged into the third set of equations to get $A^{LT}_{(1)}$. This procedure is iterated until the sum of squared residuals reached its minimum.

The advantage of this specific decomposition of normal equations into the three sets above is that it renders the estimation computationally tractable. The first set of equations only requires an actual matrix inversion, but it is of relatively low dimension. It can be performed by simple OLS on the following modified dependent variable: $Y - D_I\ A^I_{(n-1)} - D_{LT}\ A^{LT}_{(n-1)}$. The two last sets of equations simply correspond to the computations of means. First, the variable $(Y - X\ B_{(n)} - D_{LT}\ A^{LT}_{(n-1)})$ is averaged by classes of affiliates $\times$ years (classes generated by D_I). Second, the variable $(Y - X\ B_{(n)} - D_I\ A^I_{(n)})$ is averaged by classes of commuting zones $\times$ years (classes generated by D_{LT}). This algorithm, which consists in iterating sequentially across each set of equations, falls into the class of so-called "partitioned" algorithms ("zigzag" iterations) which has been analyzed in full length by Smyth (1996): while the iteration process is slow in general, the zigzag iteration is found to admit a global convergence result.

---

[77] Abowd et al. (2002) propose an algorithm that can be adapted to determine which are the estimable (identifiable) functions of categorical variables and "fixed effetcs".

G   Descriptive Analysis of the Extensive Margins

If proximity between affiliates and headquarters facilitates monitoring and information transmission, one might expect that it also matters on the "extensive margin", ie. for affiliate creations and destruction decisions. These events are not explicitly taken into account in our main regression framework of section 4, although in practice they are well captured by the affiliate level fixed effects (which "purge" regressions from most of the selection bias which could arise). A thoroughly rigorous econometric treatment of these events would require the specification of a discrete choice model of implantation across commuting zones, which is however beyond the scope of this paper. In particular, it would not be straightforward to adapt our main identification strategies to such a framework. We therefore leave this aspect for future research but propose as in Giroud (2013) a series of regressions describing the relation between affiliate creation or destruction and travel time.

G.1   Affiliate Creations

Table 13: Travel Time and the Reshuffling of Affiliates: Characteristics of Entries
1993-2011, affiliates of multi-location corporate groups only

Dependent variable:	Distance (ln) (1)	Travel time (2)	HSR (to HQ) (3)
Distance (ln)	-	-	0.155***
			(0.000)
Employment of affiliate	-0.058***	-0.033***	0.006***
(ln)	(0.002)	(0.002)	(0.000)
Gains in travel times	0.084***	-0.002	-0.002***
at other affiliates	(0.003)	(0.004)	(0.001)
Affiliate entry(ies)	-0.124***	-0.153***	-0.113***
in same group(dummy)	(0.009)	(0.012)	(0.002)
Affiliate exit(s)	-0.127***	-0.081***	-0.000
(dummy)	(0.009)	(0.012)	(0.002)
Other affiliates are HSR users	1.960***	1.693***	0.117***
	(0.008)	(0.011)	(0.002)
Total number of affiliates in group	0.046***	0.056***	0.011***
(ln)	(0.002)	(0.002)	(0.000)
Mean (un-logged) dependent variable	272	2.872	0.504
Observations	323,689	323,689	323,689

Sources: Matched DADS files and LIFI survey; affiliates (see definition in section 3.1) which are part of multi-location groups between 1993 and 2011.
Note: All regressions include industry (2 digits) × local (commuting zone) × time fixed effects to control for the local market conditions, as well as (4digit) industry dummies. Regressions also include the (ln) number of affiliates in the group as well as the group level exports in total sales to capture the cycle on international markets.

For the treatment of affiliate creations, we propose to follow Giroud (2013). More specifically, in table 13, we (simply) describe the characteristics Z_{ijlt}^{CREA} of the new affiliates[78] and investigate whether their geographic distance from headquarters, travel time or (conversely) the availability of HSR on the track to their headquarters are related to different characteristics, such as the size of the created affiliate, the total number of affiliates within group, *etc.* (X_{ijt}). We select these characteristics according to their likelihood to be associated with features of corporate groups which are likely to "tap" into the scarce HQ managerial resources

---

[78]Affiliate "creation" correspond to actual creations or to acquisitions in our setting. We only consider affiliate creation in pre-existing groups, which were furthermore already operating in the same industry (to abstract from global market entry decisions, which might obey different determinants).

that are to be shared across all affiliates - just as the management of affiliate creations.

The estimated equation therefore writes:

$$Z_{ijlt}^{CREA} = \alpha_{lt} + \gamma.X_{ijt} + \varepsilon_{ijlt} \tag{22}$$

where α_{lt} denote commuting zone $\times$ time fixed effects. Notice that such "fixed effects" are still identified in such a specification; they allow to control very precisely for (all types of) local shocks. However, it is not possible to insert affiliate level fixed effects to control for affiliate level unobserved heterogeneity. In practice, it is also (unfortunately) impossible to insert group level fixed effects, because there are too few affiliate creations per group and years, and the coefficients of interest are not sufficiently well identified in such a saturated specification.

Estimation results for this equation are reported in table 13. We obtain that larger affiliates, which are likely to be more difficult to manage remotely (or for which the "size of stake" is larger) tend to be created closer to their headquarters, whatever the indicator of distance. Similarly, when other affiliates experienced gains in travel time, thus freeing some HQ managerial resources, then affiliates tend to be created at a greater distance, and they are less likely to be created at proximity of the HSR network. These patterns suggest that when more managerial resources become available, then the trade-off underlying the location choice of affiliates (between higher gains from local market proximity and higher managerial costs from geographical dispersions) shifts in favour of higher distance and higher proximity to the final market. Last, overall, group level affiliate churning (creation *and* destruction), which might be demanding in terms of headquarters' managerial resources, are both negatively correlated with distant affiliate creations. Complementary (un-reported) regressions show that all of these results also hold industry by industry.

G.2 Affiliate Destruction

For affiliate destruction, we propose a more powerful empirical treatment, which is a direct extension of the regression framework in section 4. The estimated equation writes:

$$EXIT_{ijlt} = \alpha_j + \alpha_{lt} + \beta.T_{ijlt} + \gamma.X_{jt} + \varepsilon_{ijlt} \tag{23}$$

where $EXIT_{ijlt}$ is a dummy variable indicating that the considered affiliate is exiting from the group[79] and notations are otherwise similar to those in equation 2. The main difference with equation 2 is that the affiliate ($\times$ group) fixed effects are replaced with group level fixed effects. This is because affiliates are on average present for only 3.2 years in our panel, such that the inclusion of affiliate $\times$ group fixed effects would capture too much heterogeneity given this feature of the data.[80] Identification is otherwise very similar given the inclusion of the commuting zone $\times$ time fixed effects α_{lt}. In particular, the coefficients of interest in equation 23 are still identified from *variations* in travel times, and our estimates are safely purged from unobserved local shocks.

---

[79] Affiliate "destruction" corresponds to actual closures or to resale of affiliates, while the group itself still operates in the same industry, with other affiliates (to abstract from more global market exit decisions).

[80] The positive aspect of this feature of our data is that conversely, the affiliate level fixed effects in equation 2 are likely to purge efficiently for the selection biases arising from affiliate exit (and entry).

We estimate equation 23 as a linear probability model, using the same estimation procedure as in section 4.1, and obtained results are presented in table 14. The main result is that everything else equal, more distant affiliates (in terms of travel times) have a higher probability to be terminated. This pattern might be driven by a variety of reasons: either because these affiliates are more difficult to manage and/or because they are less profitable - or simply because they are also "politically" distant as in Bassanini et al. (2017). These correlations tend to be higher in business service industries, where their contribution to the probability of exit reaches 5 percentage points for the average travel time (2 hours). This however remains a small fraction of the entire share (47%) of exits to be explained in this industry. We also obtain that affiliates who benefited from gains in travel time thanks to the HSR network expansion were less likely to get terminated, while the probability of closure was also highly significantly reduced by gains at other affiliates. The latter findings broadly confirm the story of managerial resource constraints sketched in section 5.6.

Table 14: Travel Time and the Reshuffling of Affiliates: Exits
1993-2011, affiliates of multi-location corporate groups only

Dependent variable: probability of exit	Manufacturing Industries (1)	Personnel Services (2)	Retail and Trade (3)	Business Services (4)	Transport (5)
Travel time	0.017***	0.014***	0.014***	0.022***	0.017***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
Gains since entry	-0.014***	0.012**	-0.024***	-0.015***	-0.023***
	(0.004)	(0.005)	(0.002)	(0.004)	(0.006)
Gains at other affiliates	-0.011***	-0.018***	-0.010***	-0.011***	-0.016***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Share of exits	0.428	0.354	0.415	0.465	0.423
Observations	426,487	241,677	1,045,416	401,694	138,778

Sources: Matched DADS files and LIFI survey; affiliates (see definition in section 3.1) which are part of multi-location groups between 1993 and 2011.
Note: All regressions include headquarter (i.e., group) level fixed effects, as well as local (commuting zone) × time fixed effects to control for the local market conditions. Regressions also include group level exports in total sales to capture the cycle on international markets.