Evaluating the impacts of HSR stations on the creation of firms

This is a postprint version of the article: Matas, A., Raymond, J.L., & Roig, J.L. (2020), Evaluating the impacts of HSR stations on the creation of firms, Transport Policy, 99, 396-404. https://doi.org/10.1016/j.tranpol.2020.09.010

Keywords: high-speed rail, accessibility, economic impacts, territorial effects, sectoral effects, firm creation

1. Introduction

The broad consensus that investment in infrastructure has a positive impact on economic growth has led to continuous pressure on public policy to devote resources to new infrastructure. Such pressure is often increased by competition between regions to benefit from investments. Moreover, the lack of sound economic evaluation of the projects can lead to a waste of resources without obtaining the expected benefits on growth.

The characteristics of the investment policy in high speed rail (HSR) followed in Spain offer an appropriate context to evaluate the impact of this infrastructure on the economic development of the different regions in the country. The first HSR line was opened in 1992 between Madrid and Sevilla with 476 kilometres. The economic crisis at the beginning of the 1990s halted investment in HSR. However, from 1997 onwards the development of a high-speed network became a priority for the transport policy of the central government. Until 2018, investment in HSR amounted to 51,775 million euros, 21% of which have been funded by the European Union. Nowadays, Spain has the most extensive network of HSR[1]

[1] Length of lines on which trains can go faster than 250 km/h.

in Europe and the second most extensive in the world, with 2606 kilometres in operation and 904 under construction[2].

The decisions on HSR investment and the design of its routes have been based more on political grounds than on economic criteria (Albalate and Bel, 2011). Specifically, guided by equity concerns, policymakers have sought to provide similar HSR infrastructure endowments all over the country. In this sense, the commitment of successive governments has been that nine out of 10 citizens should have a HSR station no further than 30 kilometres away. It was expected that the high-speed network would contribute to reduce regional disparities by boosting economic growth in less developed regions.

As Figure 1 illustrates, the HSR system has been designed following a radial structure that connects Madrid with the provincial capitals. Initially, the network was conceived to provide the fastest route between the main capitals. However, in some cases, the pressures of regional and local governments have forced changes in the network layout in order to provide access to intermediate cities, thus increasing the number of stations.

Figure 1. High-speed network. 2018

[2] Mesa et al. (2017) detailed the evolution of the development of HSR in Spain.

Source: ADIF.

Following the network extension, demand has risen from 1.29 billion passenger-km in 1995 to 11.31 billion in 2018. In contrast with the network size, the rate of utilization is far below European levels. In Spain, HSR carries 4.9 million passenger-km per kilometre of network, whereas the equivalent figures for France and Germany are 23.1 and 17.1, respectively. The low levels of utilization can be explained by the fact that the network has been expanded without taking the potential demand into account.

Therefore, the criteria that guide the investment in HSR in Spain minimize the potential problems of reverse causality that often blur the estimation of the causal impact of transport infrastructure. In addition, the high investment and maintenance costs of such infrastructure justifies an ex-post evaluation of the degree of achievement of its expected benefits.

According to what has been said, this paper evaluates the impacts of the opening of a HSR station on the level of economic activity of the geographical area affected. Economic activity is measured as the number of new firms created in a given year[3]. Since HSR provides intercity passenger services, we expect a higher effect on those sectors where face-to-face connection is important (Lin, 2017). In this paper we contribute to the literature by estimating differential impacts on firm creation across four categories of activities: the service sector, tourism activities, knowledge-intensive activities, and the manufacturing sector. Our hypothesis is that since HSR facilitates human communication, the impact will be higher on the first three groups, whereas for the manufacturing industries, more dependent on freight traffic, the effect will be lower. Moreover, it can be expected that the impact will depend on the characteristics of the area that receives the investment. A second contribution of our research consists of estimating different coefficients for each province connected to the high-speed system. In this way, we provide some new insights on the territorial impact of HSR effects. Additionally, we test whether HSR leads to an increase or a reduction in the convergence of firm creation across provinces. Finally, we develop a methodological procedure to improve the accuracy of the estimation of small but positive effects.

The econometric strategy consists of estimating a fixed-effect model using panel data. Using panel data makes it possible to account for individual and temporal specific effects and, at the same time, testing for the existence of a dynamic structure in the data. We use panel data at the provincial level over the 1995–2017 period. Overall, the estimation results show that the effect of a new rail

[3] Additionally, we analysed the impacts on changes in residential location through two variables: changes in population and changes in house prices. However, we did not find any significant effect for either of these two variables.

station on firm creation is relatively small, if it exists at all, and it depends on the sector of activity and on the province characteristics.

We want to point out that the methodology and data used do not allow us to ascertain whether the estimated increase in the number of firms can be attributed to a net economic effect for the whole Spanish economy or to a relocation of firms from nearby provinces not connected to the HSR line. However, since one of the goals of HSR investment is to contribute to territorial redistribution, our analysis clearly addresses this objective.

The rest of the paper is organised as follows. In Section 2 we review those studies that have analysed the impact of HSR in Spain, and in Section 3 we show the data. Section 4 outlines the econometric strategy, and Section 5 offers the estimation results. In Section 6 we compute the distribution functions for the estimated coefficients. Finally, Section 7 provides the conclusions.

2. Related literature

The huge amount of resources devoted to HSR in Spain has prompted a relevant number of studies aimed at assessing its economic and territorial effects from different perspectives[4]. A first line of research has focused on ex-post cost-benefit analysis of investments in HSR corridors. The conclusion of these studies is that even in the most densely populated routes, demand is not high enough to

---

[4] In this paper we review the literature focused on the Spanish case. Blanquart and Koning (2017) provided a wide review of international evidence on the local economic impacts of high-speed rail; Lin (2017) reported evidence on the effects of China's HSR system on urban employment and specialization patterns; Li and Xu (2018) analysed the effects of two major HSRs in Japan on economic geography taking into account the sector and distance between cities; and Ahlfeldt and Feddersen (2018) estimated the impact of the Cologne-Frankfurt HSR in Germany in the counties of intermediate stops.

guarantee HSR profitability on socio-economic grounds (De Rus and Roman, 2006; De Rus and Nombela, 2007; Betancor and Llobet, 2015).

A second line of research has aimed to test a causal effect between the introduction of a new HSR line and some measure of economic activity using a difference-in-difference approach. Carbó et al. (2018), using panel data at the provincial level, found that the introduction of the HSR from Madrid to Barcelona had a positive effect on labour productivity, economic output, and the number of firms for the treated provinces. With regard to the number of firms, a variable similar to the one used in our study, these scholars estimated an overall growth due to HSR of 3.3% across all treated provinces and 2.8% in the intermediate provinces with HSR. Their results were confirmed for the provinces of Lleida and Tarragona by using the synthetic control method. Given the relevance of Spain as a touristic destination, some studies have looked at the role of HSR to promote tourism development. In this regard, Albalate and Fageda (2016) provided mixed evidence on the impact of HSR accessibility on tourist outcomes. On the one hand, the HSR negatively affects air traffic through a substitution effect. This may lead to an indirect negative effect on tourism as the air companies react to HSR competition by reducing the number of seats and frequency of flights. On the other hand, conditioned on the measure of accessibility and econometric technique used, HSR may have a positive (weak) effect on tourism. However, the net effect of HSR on tourism outcomes was found to not always be positive. The authors considered that the results could be attributed to a network design that does not respond to ridership needs, and which had a substitution effect on air transportation. Albalate, Campos, and Jimenez (2017) analysed the effect of new HSR corridors on the number of visitors and their overnight stays at several endline and intermediate stations at the municipality level. Using a panel database for the 2005–2012 period, they did not find any clear positive effect with respect to the tourism dependent variables. Only for the larger cities did there seem to be a positive, albeit very weak effect.

The impact of HSR on tourism has also been analysed by modelling destination choices of visitors. Pagliara et al. (2015), relying on data from a revealed preferences survey, concluded that the HSR system seemed to have a positive effect on tourists' decision to visit cities near Madrid, but the choice of Madrid as a tourist destination was not influenced by the HSR.

A third line of research has analysed how HSR changes the spatial distribution of accessibility. The underlying idea is that improving accessibility will enhance territorial cohesion and, in turn, contribute to the reduction of regional disparities. From a spatial perspective analysis, Gutiérrez Puebla (2004) pointed out that HSR might enhance competitiveness and territorial cohesion by increasing connectivity between cities. At the same time, however, he found that HSR promotes territorial polarization by benefiting the big connected cities at the expense of the cities that remain unconnected. Besides, the implementation of a HSR network might have a negative effect on the quality of service in the traditional rail network, further contributing to polarization. Monzon et al. (2018) assessed the impacts of new HSR links on "accessibility-based" territorial cohesion from 1990 to 2015. They found that new HSR lines have significantly increased the accessibility levels of the Spanish population to most destinations jointly with a reduction in the dispersion of accessibility values. Thus, better rail connections have achieved better levels of territorial cohesion in terms of accessibility. However, these authors did not assess the impact on the economic growth of the different regions.

From the reviewed literature it can be concluded that HSR has improved accessibility between connected cities by reducing rail travel time. Nonetheless, it is far from clear whether such improvements have translated into higher and more spatially equilibrated economic growth. Our research aims to provide new evidence on the economic impacts of HSR, differentiating across sectors and provinces.

3. Data

The sample comprises the Spanish peninsular provinces (NUTS3) observed between 1995 and 2017. We exclude those provinces with HSR connection before 1995, which correspond to the Madrid-Sevilla corridor[5]. The impact of HSR is measured with a dummy variable that takes a value of 1 if the province has a HSR station, and 0 otherwise. Table A.1 in Appendix 1 displays all the HSR corridors and stations included in the sample with the corresponding opening year (seven corridors and 26 stations). In all provinces the station is located in the capital city or in its surroundings. However, there are six provinces with more than one station. In general, the second station corresponds to a small city with a low traffic level.

The final sample includes 20 provinces with HSR stations and 23 without. Hence, we have 43 cross-section units and 23 temporal observations.

[5] The excluded provinces are Madrid, Ciudad Real, Cordoba, and Sevilla.

Regarding the dependent variable, we use data from the Central Business Register compiled by the National Institute of Statistics. Specifically, we evaluate the impact of HSR on the number of new registered companies grouped into four different divisions: manufacturing, services (excluding public administration), tourism-related activities[6], and knowledge-intensive activities[7]. The reason to distinguish among different groupings of sectors is that a connection to HSR will not affect all the economic activities with the same degree of intensity. First, we separate manufacturing from services. It can be expected that most of the activities related to the service sector require face-to-face contact and communication, while manufacturing is more dependent on freight services. Hence, we expect that HSR availability will have a higher capacity to increase the creation of new firms in the service sector than in the manufacturing one. Additionally, it can also be expected that HSR will positively affect the attraction of knowledge-intensive firms through a reduction in travel time that will facilitate personal interactions in a sector with a high level of human capital. Finally, improving rail accessibility can increase the tourist attractiveness of a city, with a consequent effect on related business.

The basic statistics for the number of new firms registered are shown in Table 1. The service sector presents the largest number of new firms, while this number is particularly low for the manufacturing industries. Moreover, the range between

---

[6] The definition of tourism activities corresponds to that used by the Instituto Nacional de Estadística in "La Cuenta Satélite del Turismo de España. Breve descripción metodológica", 2016.

[7] According to Eurostat, an economic activity is classified as knowledge intensive when the ratio of employees with tertiary education with respect to total employment is higher than 33%. We selected the narrower definition of knowledge-intensive activities that are those related to business industries.

the minimum and maximum values shows a high level of variability among provinces.

Table 1. Descriptive statistics for the dependent variable. Number of new firms

	Observations	Mean	Std. Deviation	Minimum	Maximum
Services	989	4489	6390	296	54200
Tourism	989	901	1041	74	8476
KIA*	989	1408	2312	79	19239
Manufacturing	989	293	481	9	5839

*KIA: knowledge-intensive activities.

Beyond cross-section fixed effects, we looked for additional control variables in order to capture the heterogeneity between provinces. In preliminary estimations, we included the economic cycle and two alternative definitions of the provincial level of education (the percentage of active population with tertiary studies and the average years of schooling). Due to the lack of sufficiently accurate data on provincial GDP, the economic cycle was computed according to total employment. The results showed that none of the educational variables were significant, probably due to an insufficient level of yearly variability. Thus, based on the principle of parsimony, only the employment cycle was maintained as a regressor.

4. Econometric strategy

To evaluate the impact of HSR, we exploit the specific features of the panel data methodology. First, we specify a dynamic equation to allow for temporal adjustments in the data. Specifically, we include the dependent variable lagged once and the current and the one-period lagged values for the explanatory variables. The number of lags is limited to one by the fact that the opening date of some stations takes place in the last years of the sample. Including more than one lag would imply losing relevant information on more recent HSR openings. Second, we control for unobserved province and temporal effects by including the corresponding dummy variables. Third, we include the employment cycle as a regressor to account for individual specific characteristics that vary over time. Fourth, to test whether connected and unconnected provinces have experienced different time trends in firm creation, we include a time-specific dummy variable for those provinces that have not benefited from high-speed investment[8]. Finally, to estimate the impact of HSR we create a dummy variable that takes a value of 1 if the province is connected to the high-speed network, and 0 otherwise. The dependent variable is log-transformed.

The general dynamic equation is specified as follows:

$$\ln(Y_{it}) = \gamma \cdot \ln(Y_{it-1}) + \sum_{k=0}^{1} \beta X_{it-k} + \sum_{k=0}^{1} \theta_k F_{it-k} + \alpha_i + \lambda_t + \delta_t D_{it} + \varepsilon_{it} \quad (1)$$

where Y_{it} is the number of firms created in province *i* and year *t*

F_{it} is a binary variable that takes a value of 1 if HSR is available in province *i* and year *t*, and 0 otherwise

X_{it} is the employment cycle

α_i are the province fixed effects

λ_t are the temporal fixed effects

[8] Ideally, we should include a specific time coefficient for each province. However, this would imply estimating 924 additional coefficients (42 provinces * 22 years), rendering the estimation unfeasible.

D_{it} are the temporal fixed effects referring to those provinces without HSR

The estimation results of the general model for each of the four sectors of activity are presented in Table A.2 in Appendix 1. To facilitate the interpretation of the estimates, and taking into account that our main interest lies in the long-run impact, we have calculated and reported the long-run coefficients for the impact of HSR for each province. All equations are estimated using generalised least squares and allowing for heteroscedasticity across panels and autocorrelation within the panel.

The estimations of the general equation showed that the coefficient of the lagged dependent variable took a low value for all sectors and was not statistically significant except for the manufacturing sector. Given these results, the next step was to exclude the lagged dependent variable and estimate the static equation. A potential drawback of this simplification strategy is that in the context of dynamic panel data the lagged dependent variable will be correlated with the error term. However, it can be shown that when the starting point is a general model that nests the unknown data generating process (DGP), estimating the general overparametrized model (dynamic specification when the true DGP is static) leads to estimated long-term effects that are very similar to those estimated when the correct model is specified. As shown in Appendix 2, the order of magnitude of the difference between the long-term coefficient estimated in the overparametrized model and the long-term coefficient estimated in the static model will be around the value of the coefficient divided by the number of time periods. With 23 temporal observations, such a difference will be less than 5%.

According to the previous reasoning, we proceeded to estimate the static equation, relying on the lagged values of the explanatory variables to capture the temporal adjustment. In this regard, the current values of the dummy variables accounting for the impact of HSR were not statistically significant for any province or sector; thus only the variables lagged once were maintained in the equation. In contrast, for the employment cycle only the current value showed significance. Hence, the equation finally estimated is:

$$\ln(Y_{it}) = \beta X_{it} + \theta_i F_{it-1} + \alpha_i + \lambda_t + \delta_t D_{it} + \varepsilon_{it} \quad (2)$$

Table A.3 compares the estimated long-term elasticities for the impact of HSR on each province, jointly with the corresponding standard errors, according to both the dynamic and the static specifications. As can be observed, there are no statistically significant differences among the estimated coefficients for any province in any sector. Hence, in our case, there is no evidence against the static or constrained model.

Finally, it has to be said that when estimating the impact of a new infrastructure we have to deal with a potential endogeneity problem of reverse causation. In our case, endogeneity may not be a severe problem since decisions on new railway lines are taken more on political than on economic grounds. Besides, working with panel data makes it possible to address the potential correlation between the regressors and the province fixed effects. The underlying idea is that any idiosyncratic changes in the number of new firms will be absorbed by the fixed effect.

5. Estimation results

Table 2 displays the estimated coefficients of Equation (2) for our variable of interest, while more comprehensive results are shown in Table A.4 in Appendix 1. We must keep in mind that we control for the employment cycle, cross-section, and time fixed effects and a time-specific fixed effect for those provinces without HSR. By similarity to the difference-in-difference approximation, we named the latter variable as “non-treated”. The employment cycle has a positive and significant effect for the services, tourism, and knowledge-intensive activities, while it has no effect on the manufacturing industries. Otherwise, the temporal effects specific for non-treated provinces are only statistically significant for the manufacturing sector in the first years of the sample. Thus, broadly, the results reveal that the temporal pattern in firm creation does not differ between connected and unconnected provinces.

Table 2. Impacts of HSR on firm creation by sector of activity

	ln(services)	ln(tourism)	ln(KIA)	ln(manufacturing)
Line Madrid–Barcelona–French border				
Girona (-1)	0.158	0.210*	0.205	0.217*
	(0.120)	(0.116)	(0.142)	(0.117)
Barcelona (-1)	0.162***	0.171***	0.142**	-0.016
	(0.049)	(0.054)	(0.057)	(0.047)
Tarragona (-1)	0.087	0.140**	0.081	0.154***
	(0.061)	(0.056)	(0.061)	(0.057)
Lleida (-1)	0.062	0.059*	-0.012	-0.004
	(0.040)	(0.032)	(0.063)	(0.045)
Huesca (-1)	0.028	0.056**	0.064	-0.064
	(0.029)	(0.027)	(0.055)	(0.063)
Zaragoza (-1)	0.052	0.080*	-0.010	-0.023
	(0.037)	(0.045)	(0.045)	(0.042)
Guadalajara (-1)	0.403***	0.318***	0.378***	0.395***
	(0.023)	(0.032)	(0.036)	(0.041)
Line Córdoba–Málaga				
Málaga (-1)	0.128***	0.087**	0.167***	0.134**
	(0.043)	(0.034)	(0.044)	(0.054)
Line Madrid–Valladolid–León				
Valladolid (-1)	-0.024	0.048	-0.055	0.020
	(0.035)	(0.044)	(0.040)	(0.053)
Segovia (-1)	-0.055*	-0.069**	-0.018	0.112*
	(0.031)	(0.034)	(0.042)	(0.068)
Palencia (-1)	0.073	0.139	0.048	-0.209
	(0.091)	(0.125)	(0.124)	(0.150)
León (-1)	0.021	-0.019	-0.038	-0.001
	(0.070)	(0.065)	(0.089)	(0.104)
Line Madrid–Toledo				
Toledo (-1)	0.130***	0.082**	0.152***	0.043
	(0.043)	(0.036)	(0.047)	(0.044)
Line Madrid–Valencia–Murcia				
Cuenca (-1)	0.029	-0.029	0.048	-0.202***
	(0.027)	(0.035)	(0.034)	(0.062)
Valencia (-1)	0.100***	0.065**	0.099***	0.018
	(0.032)	(0.031)	(0.034)	(0.046)
Albacete (-1)	0.076**	0.021	0.113***	0.064
	(0.035)	(0.039)	(0.036)	(0.068)
Alicante (-1)	0.124**	0.059	0.122***	0.130*
	(0.057)	(0.069)	(0.044)	(0.070)
Line Madrid–Ourense				
Ourense (-1)	0.030	0.049*	0.010	0.131**
	(0.036)	(0.029)	(0.037)	(0.055)
A Coruña (-1)	0.028	0.022	0.012	0.115***
	(0.027)	(0.034)	(0.042)	(0.043)
Line Madrid–Zamora				
Zamora (-1)	0.030	0.012	-0.068	-0.152
	(0.068)	(0.094)	(0.096)	(0.170)

Note: Standard errors in parentheses are corrected for heteroskedasticity across panels and for autocorrelation within panels.

With regard to our policy variable, the estimates clearly show that the impact of a HSR station differs across both provinces and sectors. However, the high standard errors of the estimated coefficients for most of the provinces and sectors introduce uncertainty about the actual impact of the HSR. To improve the accuracy, we develop a new methodological procedure based on a priori information. In the next section we present this procedure and comment upon the results.

Finally, we reestimated the equation by excluding the firms created in land transport activities from the dependent variable in order to check that the estimated coefficient for the service sector was not capturing new activities directly related to the construction of a HSR station. The results proved to be very similar. Since land transport includes activities other than those related to rail, we decided to use the service sector as a whole.

6. Distribution functions for the estimated parameters: the impacts of HSR

6.1. Methodology

Given that we analyse the effect of infrastructure on the creation of new firms in the treated provinces, it can be safely assumed that this effect will be either positive or null. Taking advantage of this a priori information, we propose to estimate the equation under the assumption that the estimated impact, θ, should be non-negative. A possible way to incorporate this constraint is to estimate a non-linear model (NLS) where the coefficient of the dummy variable that captures the HSR station is the square value of a certain γ coefficient. To present the methodology, the notation of Equation (2) is simplified to:

$$Y_{it} = Z'_{it}\delta + \theta_i F_{it} + \varepsilon_{it} \quad (3)$$

where Z_{it} includes all the explanatory variables except the dummy for the HSR.

Our proposal consists of estimating:

$$Y_{it} = Z'_{it}\delta + \gamma_i^2 F_{it} + \varepsilon_{it} \quad (4)$$

However, in our case, the estimation of a NLS model with such a high number of coefficients prevented the convergence of the estimator.

As an alternative, estimating the equation by NLS leads to exactly the same results as recovering the non-linear γ parameter and its standard error as follows:

$$\hat{\gamma} = \sqrt{\hat{\theta}} \quad (5)$$

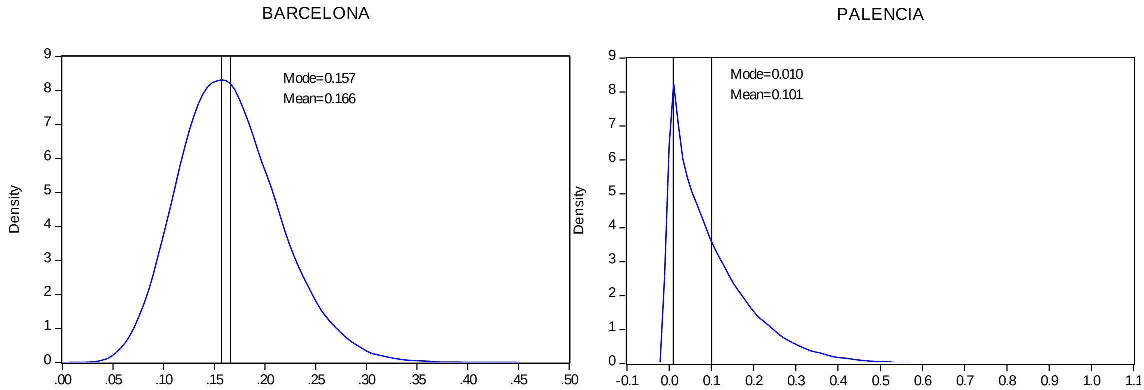
$$\sigma_{\hat{\gamma}} = 0.5 \cdot \frac{\sigma_{\hat{\theta}}}{\hat{\gamma}} \quad (6)$$

Thus, for those estimated coefficients greater than zero, we computed γ and its standard error. With this information, we constructed the one-tail distribution of $\hat{\theta}$ by simulation under the assumption that it takes only positive values. So, in addition to the initial point estimate of θ, the distribution of $\hat{\theta}$ provides more information on the actual impact of being connected to the HSR network.

For all the coefficients estimated with a positive sign, we have computed the corresponding one-tail distribution. In those cases where a negative coefficient was estimated, our assumption was that the impact of HSR was null. Given that the distribution is asymmetric, the mean is not the most probable value of the parameter, so we selected the mode as the point estimate of the impact. Additionally, we calculated the optimal confidence interval by simulation, taking the mode as the central point.

With the objective of illustrating how the degree of asymmetry of the distribution differs between two provinces, Figure 2 plots the corresponding distribution of the impact of HSR on the service sector for the provinces of Barcelona and Palencia.

Figure 2. One-tail distribution of HSR impact for Barcelona and Palencia (service sector)



In the case of Barcelona, the distribution is almost symmetric, and the mode and the mean are very similar, whereas the reverse is true for Palencia.

6.2. The impacts of the HSR

Table 3 displays the mode and the upper and lower limits of the 95% confidence interval for the one-tail distribution for the four groupings of sectors. Regarding the sectoral impacts, the results are in line with our expectations. The response to a HSR connection is stronger for service, tourism, and knowledge-intensive activities than in manufacturing industries. Therefore, the results confirm that the enhancement of intercity rail passenger transport provides greater benefits for those activities that rely on personal communication. For knowledge-intensive activities, it may well be that the HSR reinforces the role of science parks promoted by local governments.

Table 3. HSR impact on firm creation: mode and 95% confidence interval for one-tail distribution

	ln(services)				**ln(tourism)**				**ln(knowledge intensive)**				**ln(manufacturing)**		
	Lower limit	Mode	Upper limit		Lower limit	Mode	Upper limit		Lower limit	Mode	Upper limit		Lower limit	Mode	Upper limit
Guadalajara	0.360	0.403	0.449	Guadalajara	0.257	0.316	0.385	Guadalajara	0.311	0.375	0.452	Guadalajara	0.319	0.392	0.480
Barcelona	0.080	0.157	0.272	Girona	0.044	0.180	0.503	Malaga	0.092	0.162	0.265	Girona	0.048	0.180	0.506
Toledo	0.059	0.124	0.229	Barcelona	0.082	0.164	0.292	Girona	0.022	0.160	0.582	Tarragona	0.062	0.140	0.287
Malaga	0.057	0.119	0.227	Tarragona	0.052	0.126	0.270	Toledo	0.074	0.142	0.257	Malaga	0.050	0.120	0.260
Girona	0.010	0.110	0.483	Malaga	0.033	0.079	0.166	Barcelona	0.051	0.130	0.276	Ourense	0.044	0.120	0.261
Alicante	0.037	0.110	0.260	Toledo	0.027	0.075	0.168	Alicante	0.051	0.116	0.224	Alicante	0.029	0.100	0.303
Valencia	0.046	0.093	0.173	Palencia	0.003	0.070	0.490	Albacete	0.053	0.109	0.196	A Coruña	0.046	0.100	0.213
Tarragona	0.009	0.069	0.248	Zaragoza	0.016	0.068	0.192	Valencia	0.043	0.092	0.177	Segovia	0.018	0.090	0.285
Albacete	0.023	0.066	0.160	Valencia	0.018	0.056	0.140	Tarragona	0.006	0.050	0.247	Albacete	0.001	0.010	0.266
Lleida	0.009	0.048	0.165	Lleida	0.013	0.051	0.137	Huesca	0.002	0.040	0.214	Toledo	0.000	0.004	0.172
Zaragoza	0.005	0.041	0.151	Huesca	0.016	0.048	0.122	Cuenca	0.004	0.035	0.138	Valladolid	0.000	0.000	0.261
Cuenca	0.001	0.011	0.105	Ourense	0.009	0.042	0.122	Palencia	0.000	0.000	0.597	Valencia	0.000	0.000	0.220
Palencia	0.000	0.010	0.359	Segovia	0.001	0.024	0.175	Ourense	0.000	0.000	0.221	Barcelona		no effect	
Ourense	0.000	0.004	0.141	Albacete	0.000	0.010	0.167	A Coruña	0.000	0.000	0.236	Lleida		no effect	
Huesca	0.000	0.003	0.115	Alicante	0.000	0.010	0.274	Lleida		no effect		Huesca		no effect	
A Coruña	0.006	0.003	0.012	A Coruña	0.000	0.005	0.140	Zaragoza		no effect		Zaragoza		no effect	
Leon	0.000	0.000	0.395	Zamora	0.000	0.000	1.017	Valladolid		no effect		Palencia		no effect	
Zamora	0.000	0.000	0.312	Valladolid		no effect		Segovia		no effect		Leon		no effect	
Valladolid		no effect		Leon		no effect		Leon		no effect		Cuenca		no effect	
Segovia		no effect		Cuenca		no effect		Zamora		no effect		Zamora		no effect	

Note: “no effect” corresponds to the negative coefficients estimated in Equation 2

Beyond those provinces where the effect is null as a result of a negative coefficient in Equation 2, the wide range observed for the confidence interval reveals that the likely value of the impact will be zero for a certain additional number of provinces. To facilitate the interpretation of the results, we consider that the probability of observing a positive effect is not negligible when the mode is above 0.01 and the lower limit is above zero. Under this assumption, the HSR would have a positive impact on firm creation in 12 provinces for the service sector, 13 for tourism activities, 11 for knowledge-intensive activities, and eight for manufacturing industries[9]. Our findings agree with those of Lin (2017) for China. Her analysis of the differential impacts of HSR on employment across sectors indicates that the industries benefiting more from enhanced rail market access are either tourism-related or intensive in nonroutine cognitive skills.

Regarding the magnitude of the effect, the results have to be analysed leaving the province of Guadalajara aside[10]. Broadly, we estimate that being connected to a HSR can boost the creation of new firms between 1.1% and 18%. We have to bear in mind that Carbó et al. (2018), using Spanish data, estimated an average effect of HSR on the creation of new firms around 3.3%. Regarding international evidence, Lin (2017) reported that cities connected to HSR experienced a 7% higher growth in employment compared to those unconnected,

---

[9] We have to point out that the number of new firms created in the manufacturing sector is much lower than in the rest of the groupings. Consequently, the results can be more erratic.

[10] Special mention is needed for the case of Guadalajara, where the estimated impact is above 30% for all sectors of activities. Guadalajara is a province close to Madrid with a relatively low level of economic activity and population. Since the mid-1990s, its population has grown at an annual rate of 2.2%—well above the country average of 0.7%. A significant part of this growth is the result of a process of suburbanization from Madrid. Although transport improvements have facilitated this process, the growth in firms' creation cannot be entirely attributed to HSR connection.

whereas Ahlfeldt and Feddersen (2018) found an average causal effect of the HSR on GDP equal to 8.5%.

Finally, our work makes it possible to contribute to the debate on the territorial impacts of HSR effects. Looking at the territorial distribution of the effects, it is possible to distinguish differences between services and knowledge-intensive sectors, on the one hand, and tourism and manufacturing, on the other hand. If we consider the first two sectors, nine provinces (eight excluding Guadalajara) experience effects on firm creation above 5%. These provinces are the same in both sectors. The provinces are located along the three HSR lines that connect Madrid with the Mediterranean coast. Amongst these provinces, two different types appear. In the first place, we find the largest Mediterranean cities, from Barcelona to Malaga. The HSR seems to have induced a trend towards concentration of the sectors that are the most dependent on face-to-face interactions in these big cities. But additionally, in the case of the two biggest metropolitan areas, Madrid and Barcelona, medium-size cities within the range of 100 kilometres from the metropolitan centre have benefited from a certain decentralization and metropolitan integration as a result of HSR. That would be the case of Guadalajara and Toledo with respect to Madrid, and of Girona and, to a lesser extent, Tarragona with respect to Barcelona[11]. Reducing the time connection between these cities and the core of the metropolitan area makes them more attractive and fosters their metropolitan integration, as highlighted by Garmendia et al. (2012). Although the original goal of the HSR network was to provide interurban connections at distances around 500 kilometres, the pressure of local governments for connecting their regions (areas, territory) to the network resulted in the construction of intermediate rail stations. Besides, Renfe, the

[11] The capital provinces of Guadalajara and Toledo are located, respectively, 58 and 71 kilometres from Madrid, whereas those of Girona and Tarragona are located 100 kilometres away.

public rail operator, has contributed to the success of the short- and medium-distance connections by offering a high quality of service at relatively low prices for frequent users. Conversely, the territorial distribution of HSR effects on manufacturing new firms follows a decentralized pattern. The largest effects are found in small- and medium-size cities with the exception of Malaga.

These results follow those of Li and Xu (2018), who estimated the effects for the case of Japan. They found an effect of concentration around Tokyo in the case of services employment, whereas manufacturing is crowded out from the main centres due to the growing pressure of service employment on operational and living costs, which are not affordable for manufacturing activities.

Finally, the territorial distribution of the effects of HSR on new firms in the tourism sector follows an expected pattern. Firstly, the metropolitan clusters of Madrid and Barcelona are the main attractors. We should take into account that in these areas, hotel and catering activities provide services for both tourist and business trips. Secondly, we find a strong effect in the two main Mediterranean cities and areas outside Barcelona, namely Malaga and Valencia. Finally, a few interior areas beyond the two main metropolitan areas but within the range of short-stay leisure trips have seen their touristic activities reinforced by the time savings provided by the new connection.

Secondly, we tested whether a relationship between the impact of HSR on firm creation and the level of economic dynamics of the province exists. The dynamics of a province were approximated by the average number of firms created in each province before the opening of the first HSR station in 2003. To take the heterogeneity in province size into account, we transformed the variables in per capita terms (pc). By similarity to the “beta-convergence” approach, we estimated the following equation:

$$dln(Y_{pc}) = \alpha + \beta \ln(Y_{pc-1}) \quad (7)$$

where $dln(Y_{pc})$ is the estimated HSR impact and

Y_{pc-1} is the average number of firms per capita created between 1995 and 2002.

Taking into account that the previous equation can be expressed as:

$$\ln(Y_{pc}) = (1+\beta)\ln(Y_{pc})_{-1} + u \quad (8)$$

the following relationship between the level of dispersion in firms' creation before and after the introduction of the HSR can be derived:

$$\sigma^2_{\ln(Y_{pc})} = (1+\beta)^2 \sigma^2_{\ln(Y_{pc})_{-1}} + \sigma^2_u \quad (9)$$

Or:

$$\frac{\sigma^2_{\ln(Y_{pc})}}{\sigma^2_{\ln(Y_{pc})_{-1}}} = (1+\beta)^2 + \frac{\sigma^2_u}{\sigma^2_{\ln(Y_{pc})_{-1}}} \quad (10)$$

So, generally, $if\ \beta > 0$, the introduction of HSR would promote divergence in terms of dispersion of firm creation across Spanish provinces; if $\beta = 0$, the effect would be neutral, and if $\beta < 0$, the HSR would promote divergence.

Table 4 displays the results of the estimated equations for all the activity sectors using the seemingly unrelated regression method.

Table 4. Firm creation convergence

	Coef.	t-statistic	Observations	R^2
Service				
ln(firms created pc)	0.183	5.94	19	0.63
Constant	1.006	6.26		
Tourisms				
ln(firms created pc)	0.086	1.31	19	0.11
Constant	0.632	1.43		
Knowledge intensive				
ln(firms created pc)	0.106	2.64	19	0.23
Constant	0.742	2.84		
Manufacturing				
ln(firms created pc)	0.026	0.71	19	-0.006
Constant	0.239	0.88		

Note: The province of Guadalajara was excluded given its abnormally high coefficient.

The estimated coefficient takes a positive sign in all four equations, and it is statistically significant for the service and knowledge-intensive activities. Thus, our results provide evidence for the fact that the HSR network would have favoured firm creation in the more dynamic provinces, contributing to further increase the level of dispersion across provinces. These results represent bad news for the social and territorial cohesion criteria that have guided HSR investment in Spain.

7. Conclusions

This paper has analysed the effects of being connected to a HSR network on the creation of firms in Spain. Investment decisions have been guided more by territorial and social equity concerns than efficiency criteria. This has resulted in a rail network that tends to provide similar levels of infrastructure endowments independently of the potential demand.

The results point out that being connected to a HSR system may boost firm creation, but this is not always the case. Considerable heterogeneity exists across sectors and provinces. Connection to the HSR network leads to higher growth in firm creation in sectors where face-to-face relations are more important. This is the case for service, tourism, and knowledge-intensive activities.

Additionally, the development of a high-speed network may contribute to modify the territorial distribution of activities in different ways. In the cases of services and knowledge-intensive activities, HSR can reinforce the concentration of firm creation on the bigger cities, whereas in manufacturing no clear pattern emerges. Additionally, improving short-distance connections can reinforce the agglomeration benefits by enlarging the metropolitan areas. Finally, contrary to transport authorities' expectations, we find evidence that HSR would contribute to increasing the level of dispersion in terms of firm creation across provinces in the cases of service and knowledge-intensive activities. Therefore, our work calls for a rethinking of the strategy of basing investment decisions on territorial equity criteria.

References

Ahlfeldt, G. & Feddersen, A. (2018). From periphery to core: measuring agglomeration effects using high-speed rail. *Journal of Economic Geography*, 18, 355-390. DOI: 10.1093/jeg/lbx005.

Albalate, D. & Bel, G. (2011). Cuando la economía no importa: Auge y esplendor de la alta velocidad en España. *Revista de Economía Aplicada*, 55, 171-190.

Albalate, D. & Fageda, X. (2016). High speed rail and tourism: Empirical evidence from Spain. *Transportation Research Part A*, 85, 174-185. DOI: 10.1016/j.tra.2016.01.009

Albalate, D., Campos, J, & Jiménez, J.L. (2017). Tourism and high speed rail in Spain: Does the AVE increase local visitors? *Annals of Tourism Research*, 65, 71-82. DOI: 10.1016/j.annals.2017.05.004

Betancor, O. & Llobet, G. (2015). Contabilidad financiera y social de la alta velocidad en España. *Estudios sobre la Economía Española*, 2015/08. Fedea.

Blanquart, C & Koning, M. (2017). The local economic impacts of high-speed railways: theories and facts. *European Transportation Research Review*, 9:12, 1-14. DOI: 10.1007/s12544-017-0233-0.

Carbo, J.M., Graham, D.J., Anupriya, Casas, D. & Melo, P. (2018). Evaluating the causal economic impacts of transport investments: evidence from the Madrid-Barcelona high speed rail corridor. *Journal of Applied Statistics*, 46, 1714-1723, DOI: 10.1080/02664763.2018.1558188.

De Rus, G. & Nombela, G. (2007). Is investment in high speed rail socially profitable? *Journal of Transport Economics and Policy*, 41, 3-23.

De Rus, G. & Román, C. (2006). Análisis económico de la línea de alta velocidad Madrid-Barcelona. *Revista de Economía Aplicada*, 42, 35-79.

Garmendia, M., Ribalaygua, C. & Ureña, J.M. (2012). High speed rail: implication for cities. *Cities*, 29, S26-S31.DOI: 10.1016/j.cities.2012.06.005

Gutierrez Puebla, J. (2004). El tren de alta velocidad y sus efectos espaciales. *Investigaciones Regionales*, 5, 199-121.

Li, Z. & Xu, H. (2018). High-speed railroads and economic geography: Evidence from Japan. *Journal of Regional Science*, 58, 1-23. DOIU: 10.1111/jors.12384

Lin, Y. (2017). Travel costs and urban specialization patterns: Evidence from China's high speed rail system. *Journal of Urban Economics*, 98, 98-123. DOI: 10.1016/j.jue.2016.11.002

Mesa, E. et al. (2017), Atlas. High Speed Rail in Spain, Fundación de los Ferrocarriles Españoles, Madrid.

Monzón, A., Lopez, E. & Ortega, E. (2019). Has HSR improved territorial cohesion in Spain? An accessibility analysis of the first 25 years: 1990-2015. *European Planning Studies*, 27, 513-532. DOI: 10.1080/09654313.2018.1562656

Pagliara, F., La Pietra, A. & Vassallo, J.M. (2015). High speed rail and the tourism market: Evidence from the Madrid case study. *Transport Policy*, 37, 187-194. DOI: 10.1016/j.tranpol.2014.10.015

Appendix 1

Table A.1. High-speed stations included in the sample

Station	Province	Opening year
Line Madrid-Barcelona-French border		
Guadalajara	Guadalajara	2003
Calatayud	Zaragoza	2003
Zaragoza	Zaragoza	2003
Lleida	Lleida	2003
Huesca	Huesca	2005
Tarragona	Tarragona	2006
Barcelona	Barcelona	2008
Girona	Girona	2013
Figueres	Girona	2013
Line Córdoba-Málaga		
Antequera	Málaga	2006
Málaga	Málaga	2006
Line Madrid-Valladolid-León		
Segovia	Segovia	2007
Valladolid	Valladolid	2007
León	León	2015
Palencia	Palencia	2015
Line Madrid-Toledo		
Toledo	Toledo	2005
Line Madrid-Valencia-Murcia		
Cuenca	Cuenca	2010
Requena	Valencia	2010
Valencia	Valencia	2010
Albacete	Albacete	2010
Villena	Alicante	2013
Alicante	Alicante	2013
Line Madrid - Ourense		
Ourense	Ourense	2011
Santiago	A Coruña	2011
A Coruña	A Coruña	2011
Line Madrid-Zamora		
Zamora	Zamora	2015

Table A.2. Estimation results of the general model

	ln(services)		ln(tourism)		ln(KIA)		ln(manufacturing)	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Constant term	7.2327	0.2563	5.1671	0.2040	5.9015	0.2291	4.5305	0.1868
ln(depend.variable (-1))	-0.0312	0.0347	0.0668	0.0353	0.0205	0.0349	0.1010	0.0337
Long-term coefficients								
Employment cycle	0.0005	0.0002	0.0006	0.0002	0.0005	0.0002	-0.0001	0.0003
HSR impact								
Girona	0.1612	0.1167	0.2378	0.1176	0.2131	0.1383	0.2082	0.1229
Barcelona	0.1412	0.0408	0.1486	0.0503	0.1138	0.0520	-0.0492	0.0495
Tarragona	0.1084	0.0554	0.1539	0.0515	0.0720	0.0514	0.1578	0.0568
Lleida	0.0752	0.0385	0.0767	0.0320	-0.0052	0.0650	0.0013	0.0535
Huesca	0.0262	0.0297	0.0482	0.0249	0.0494	0.0569	-0.0759	0.0665
Zaragoza	0.0586	0.0391	0.0896	0.0498	-0.0103	0.0515	-0.0201	0.0504
Guadalajara	0.3842	0.0223	0.2882	0.0299	0.3587	0.0383	0.3850	0.0493
Malaga	0.1743	0.0462	0.1062	0.0337	0.1793	0.0458	0.1248	0.0509
Valladolid	-0.0230	0.0379	0.0595	0.0485	-0.0705	0.0417	-0.0139	0.0555
Segovia	-0.0418	0.0306	-0.0662	0.0340	-0.0134	0.0439	0.1253	0.0746
Palencia	-0.0323	0.0865	0.1249	0.1325	-0.0123	0.1008	-0.3142	0.1499
Leon	-0.0178	0.0753	-0.0240	0.0667	-0.0770	0.0845	-0.0268	0.1197
Toledo	0.1656	0.0455	0.0956	0.0387	0.1755	0.0521	0.0368	0.0484
Cuenca	0.0360	0.0278	-0.0293	0.0367	0.0365	0.0341	-0.1981	0.0686
Valencia	0.1020	0.0305	0.0728	0.0338	0.1049	0.0327	0.0178	0.0498
Albacete	0.0875	0.0350	0.0324	0.0430	0.1150	0.0359	0.0407	0.0483
Alicante	0.1140	0.0538	0.0541	0.0717	0.1186	0.0445	0.1265	0.0684
Ourense	0.0486	0.0343	0.0621	0.0252	0.0095	0.0358	0.1074	0.0622
A Coruña	0.0293	0.0283	0.0248	0.0346	0.0054	0.0431	0.1218	0.0406
Zamora	-0.0279	0.0752	-0.0595	0.1021	-0.0616	0.0934	-0.1848	0.1949
Province fixed-effects		yes		yes		yes		yes
Year fixed-effects		yes		yes		yes		yes
Year fixed-effects without HST		yes		yes		yes		yes
Number of observations		903		903		903		903

Note: Standard errors in parenthesis are corrected for heteroskedasticity across panels and for autocorrelation within panels

Table A.3.Comparison of the estimated long term elasticities for the impact of HSR

	General dynamic model			Static model			Difference of coefficients / Std. Error	
Service	**Coef.(C1)**	**Std. Err. (SE1)**	**t-statistic**	**Coef.(C2)**	**Std. Error (SE2)**	**t-statistic**	**(C(2)-C(1))/SE1**	**(C(2)-C(1))/SE2**
Girona	0.1612	0.1167	1.38	0.1580	0.1204	1.31	-0.0278	-0.0270
Barcelona	0.1412	0.0408	3.46	0.1619	0.0489	3.31	0.5086	0.4246
Tarragona	0.1084	0.0554	1.96	0.0870	0.0611	1.42	-0.3867	-0.3506
Lleida	0.0752	0.0385	1.95	0.0624	0.0401	1.56	-0.3330	-0.3199
Huesca	0.0262	0.0297	0.88	0.0284	0.0294	0.97	0.0773	0.0780
Zaragoza	0.0586	0.0391	1.5	0.0520	0.0372	1.4	-0.1691	-0.1775
Guadalajara	0.3842	0.0223	17.23	0.4033	0.0228	17.67	0.8564	0.8366
Valladolid	-0.0230	0.0379	-0.61	-0.0235	0.0354	-0.66	-0.0127	-0.0136
Cuenca	0.0360	0.0278	1.29	0.0287	0.0268	1.07	-0.2610	-0.2710
Valencia	0.1020	0.0305	3.34	0.0996	0.0322	3.09	-0.0802	-0.0759
Albacete	0.0875	0.0350	2.5	0.0756	0.0350	2.16	-0.3391	-0.3387
Alicante	0.1140	0.0538	2.12	0.1237	0.0570	2.17	0.1794	0.1693
Malaga	0.1743	0.0462	3.77	0.1278	0.0433	2.95	-1.0070	-1.0748
Ourense	0.0486	0.0343	1.42	0.0304	0.0360	0.85	-0.5293	-0.5046
A Coruña	0.0293	0.0283	1.03	0.0280	0.0274	1.02	-0.0452	-0.0468
Toledo	0.1656	0.0455	3.64	0.1302	0.0431	3.02	-0.7787	-0.8222
Segovia	-0.0418	0.0306	-1.36	-0.0550	0.0314	-1.75	-0.4334	-0.4223
Zamora	-0.0279	0.0752	-0.37	0.0299	0.0682	0.44	0.7679	0.8469
Palencia	-0.0323	0.0865	-0.37	0.0732	0.0914	0.8	1.2192	1.1543
Leon	-0.0178	0.0753	-0.24	0.0207	0.0701	0.29	0.5113	0.5489
Tourism	**Coef.(C1)**	**Std. Err. (SE1)**	**t-statistic**	**Coef.(C2)**	**Std. Error (SE2)**	**t-statistic**	**(C(2)-C(1))/SE1**	**(C(2)-C(1))/SE2**
Girona	0.2378	0.1176	2.02	0.2101	0.1162	1.81	-0.2390	-0.2362
Barcelona	0.1486	0.0503	2.95	0.1708	0.0537	3.18	0.4137	0.4415
Tarragona	0.1539	0.0515	2.99	0.1404	0.0557	2.52	-0.2423	-0.2619
Lleida	0.0767	0.0320	2.4	0.0589	0.0316	1.86	-0.5635	-0.5579
Huesca	0.0482	0.0249	1.94	0.0561	0.0270	2.08	0.2924	0.3174
Zaragoza	0.0896	0.0498	1.8	0.0800	0.0449	1.78	-0.2139	-0.1930

Guadalajara	0.2882	0.0299	9.63	0.3177	0.0325	9.78	0.9101	0.9885
Valladolid	0.0595	0.0485	1.23	0.0484	0.0443	1.09	-0.2510	-0.2295
Cuenca	-0.0293	0.0367	-0.8	-0.0289	0.0346	-0.84	0.0118	0.0111
Valencia	0.0728	0.0338	2.16	0.0646	0.0313	2.06	-0.2617	-0.2427
Albacete	0.0324	0.0430	0.75	0.0205	0.0390	0.53	-0.3049	-0.2764
Alicante	0.0541	0.0717	0.75	0.0588	0.0692	0.85	0.0673	0.0649
Malaga	0.1062	0.0337	3.15	0.0875	0.0340	2.58	-0.5520	-0.5562
Ourense	0.0621	0.0252	2.47	0.0493	0.0288	1.71	-0.4450	-0.5094
A Coruña	0.0248	0.0346	0.72	0.0215	0.0341	0.63	-0.0962	-0.0949
Toledo	0.0956	0.0387	2.47	0.0823	0.0361	2.28	-0.3687	-0.3438
Segovia	-0.0662	0.0340	-1.95	-0.0686	0.0341	-2.01	-0.0722	-0.0724
Zamora	-0.0595	0.1021	-0.58	0.0116	0.0938	0.12	0.7580	0.6967
Palencia	0.1249	0.1325	0.94	0.1386	0.1250	1.11	0.1093	0.1031
Leon	-0.0240	0.0667	-0.36	-0.0188	0.0647	-0.29	0.0806	0.0783
KIA	**Coef.(C1)**	**Std. Err. (SE1)**	**t-statistic**	**Coef.(C2)**	**Std. Error (SE2)**	**t-statistic**	**(C(2)-C(1))/SE1**	**(C(2)-C(1))/SE2**
Girona	0.2131	0.1383	1.54	0.2052	0.1425	1.44	-0.0552	-0.0569
Barcelona	0.1138	0.0520	2.19	0.1417	0.0573	2.47	0.4870	0.5367
Tarragona	0.0720	0.0514	1.4	0.0814	0.0613	1.33	0.1520	0.1812
Lleida	-0.0052	0.0650	-0.08	-0.0121	0.0634	-0.19	-0.1083	-0.1058
Huesca	0.0494	0.0569	0.87	0.0645	0.0545	1.18	0.2766	0.2648
Zaragoza	-0.0103	0.0515	-0.2	-0.0097	0.0454	-0.21	0.0140	0.0123
Guadalajara	0.3587	0.0383	9.36	0.3781	0.0356	10.63	0.5460	0.5072
Valladolid	-0.0705	0.0417	-1.69	-0.0545	0.0405	-1.35	0.3954	0.3839
Cuenca	0.0365	0.0341	1.07	0.0478	0.0341	1.4	0.3294	0.3298
Valencia	0.1049	0.0327	3.21	0.0989	0.0342	2.89	-0.1742	-0.1824
Albacete	0.1150	0.0359	3.2	0.1133	0.0364	3.11	-0.0454	-0.0460
Alicante	0.1186	0.0445	2.67	0.1221	0.0442	2.77	0.0794	0.0788
Malaga	0.1793	0.0458	3.91	0.1673	0.0440	3.8	-0.2725	-0.2616
Ourense	0.0095	0.0358	0.26	0.0097	0.0370	0.26	0.0049	0.0051
A Coruña	0.0054	0.0431	0.13	0.0118	0.0416	0.28	0.1532	0.1477

Toledo	0.1755	0.0521	3.37	0.1515	0.0466	3.25	-0.5148	-0.4605
Segovia	-0.0134	0.0439	-0.31	-0.0182	0.0424	-0.43	-0.1136	-0.1098
Zamora	-0.0616	0.0934	-0.66	-0.0680	0.0963	-0.71	-0.0673	-0.0694
Palencia	-0.0123	0.1008	-0.12	0.0484	0.1239	0.39	0.4907	0.6031
Leon	-0.0770	0.0845	-0.91	-0.0375	0.0888	-0.42	0.4442	0.4673
Manufacturing	**Coef.(C1)**	**Std. Err. (SE1)**	**t-statistic**	**Coef.(C2)**	**Std. Error (SE2)**	**t-statistic**	**(C(2)-C(1))/SE1**	**(C(2)-C(1))/SE2**
Girona	0.2082	0.1229	1.69	0.2167	0.1170	1.85	0.0727	0.0692
Barcelona	-0.0492	0.0495	-0.99	-0.0161	0.0472	-0.34	0.7023	0.6702
Tarragona	0.1578	0.0568	2.78	0.1535	0.0574	2.67	-0.0740	-0.0748
Lleida	0.0013	0.0535	0.02	-0.0035	0.0453	-0.08	-0.1062	-0.0898
Huesca	-0.0759	0.0665	-1.14	-0.0638	0.0633	-1.01	0.1914	0.1822
Zaragoza	-0.0201	0.0504	-0.4	-0.0226	0.0416	-0.54	-0.0585	-0.0484
Guadalajara	0.3850	0.0493	7.81	0.3952	0.0412	9.59	0.2485	0.2076
Valladolid	-0.0139	0.0555	-0.25	0.0196	0.0529	0.37	0.6318	0.6024
Cuenca	-0.1981	0.0686	-2.89	-0.2020	0.0622	-3.25	-0.0623	-0.0565
Valencia	0.0178	0.0498	0.36	0.0182	0.0459	0.4	0.0092	0.0085
Albacete	0.0407	0.0483	0.84	0.0640	0.0675	0.95	0.3449	0.4824
Alicante	0.1265	0.0684	1.85	0.1295	0.0699	1.85	0.0430	0.0439
Malaga	0.1248	0.0509	2.45	0.1342	0.0540	2.49	0.1732	0.1838
Ourense	0.1074	0.0622	1.73	0.1307	0.0553	2.36	0.4222	0.3753
A Coruña	0.1218	0.0406	3	0.1147	0.0427	2.69	-0.1664	-0.1748
Toledo	0.0368	0.0484	0.76	0.0434	0.0439	0.99	0.1484	0.1345
Segovia	0.1253	0.0746	1.68	0.1117	0.0677	1.65	-0.1995	-0.1810
Zamora	-0.1848	0.1949	-0.95	-0.1520	0.1705	-0.89	0.1922	0.1681
Palencia	-0.3142	0.1499	-2.1	-0.2093	0.1500	-1.4	0.6998	0.7002
Leon	-0.0268	0.1197	-0.22	-0.0006	0.1043	-0.01	0.2511	0.2187

Table A.4. Estimation results of the static model

	ln(services)		ln(tourism)		ln(KIA)		ln(manufacturing)	
	Coef.	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**
Constant term	7.0157	0.0735	5.5345	0.0627	5.9738	0.0915	5.0968	0.0788
Employment cycle	0.00058	0.0002	0.00062	0.0002	0.00041	0.0002	-0.00004	0.0002
HSR impact								
Girona (-1)	0.1580	0.1204	0.2101	0.1162	0.2052	0.1425	0.2167	0.1170
Barcelona (-1)	0.1619	0.0489	0.1708	0.0537	0.1417	0.0573	-0.0161	0.0472
Tarragona (-1)	0.0870	0.0611	0.1404	0.0557	0.0814	0.0613	0.1535	0.0574
Lleida (-1)	0.0624	0.0401	0.0589	0.0316	-0.0121	0.0634	-0.0035	0.0453
Huesca (-1)	0.0284	0.0294	0.0561	0.0270	0.0645	0.0545	-0.0638	0.0633
Zaragoza (-1)	0.0520	0.0372	0.0800	0.0449	-0.0097	0.0454	-0.0226	0.0416
Guadalajara (-1)	0.4033	0.0228	0.3177	0.0325	0.3781	0.0356	0.3952	0.0412
Málaga (-1)	0.1278	0.0433	0.0875	0.0340	0.1673	0.0440	0.1342	0.0540
Valladolid (-1)	-0.0235	0.0354	0.0484	0.0443	-0.0545	0.0405	0.0196	0.0529
Segovia (-1)	-0.0550	0.0314	-0.0686	0.0341	-0.0182	0.0424	0.1117	0.0677
Palencia (-1)	0.0732	0.0914	0.1386	0.1250	0.0484	0.1239	-0.2093	0.1500
León (-1)	0.0207	0.0701	-0.0188	0.0647	-0.0375	0.0888	-0.0006	0.1043
Toledo (-1)	0.1302	0.0431	0.0823	0.0361	0.1515	0.0466	0.0434	0.0439
Cuenca (-1)	0.0287	0.0268	-0.0289	0.0346	0.0478	0.0341	-0.2020	0.0622
Valencia (-1)	0.0996	0.0322	0.0646	0.0313	0.0989	0.0342	0.0182	0.0459
Albacete (-1)	0.0756	0.0350	0.0205	0.0390	0.1133	0.0364	0.0640	0.0675
Alicante (-1)	0.1237	0.0570	0.0588	0.0692	0.1221	0.0442	0.1295	0.0699
Ourense (-1)	0.0304	0.0360	0.0493	0.0288	0.0097	0.0370	0.1307	0.0553
A Coruña (-1)	0.0280	0.0274	0.0215	0.0341	0.0118	0.0416	0.1147	0.0427
Zamora (-1)	0.0299	0.0682	0.0116	0.0938	-0.0680	0.0963	-0.1520	0.1705
Non-treated*1996	-0.028	0.0447	-0.0866	0.0438	-0.069	0.0523	-0.157	0.0561
Non-treated*1997	-0.079*	0.0443	-0.121***	0.0433	-0.138***	0.0518	-0.200***	0.0556
Non-treated*1998	-0.039	0.0443	-0.077*	0.0433	-0.060	0.0518	-0.076	0.0556
9Non-treated*1996	-0.020	0.0443	-0.051	0.0433	-0.034	0.0518	-0.115	0.0556
Non-treated*2000	-0.066	0.0443	-0.076*	0.0432	-0.095*	0.0517	-0.171*	0.0555
Non-treated*2001	0.017	0.0443	-0.016	0.0433	0.021	0.0518	-0.020	0.0556
Non-treated*2002	-0.024	0.0443	-0.039	0.0433	-0.045	0.0517	-0.081	0.0556
Non-treated*2003	0.014	0.0443	0.020	0.0433	-0.000	0.0517	-0.031	0.0555
Non-treated*2004	0.0003	0.0442	0.023	0.0432	-0.007	0.0517	-0.050	0.0552
Non-treated*2005	-0.007	0.0437	-0.035	0.0428	-0.014	0.0513	-0.052	0.0544
Non-treated*2006	-0.023	0.0437	-0.034	0.0427	-0.018	0.0515	-0.065	0.0542
Non-treated*2007	-0.006	0.0434	-0.013	0.0424	-0.039	0.0511	-0.055	0.0538
Non-treated*2008	0.006	0.0434	-0.001	0.0424	-0.025	0.0511	-0.105	0.0535
Non-treated*2009	0.017	0.0430	0.021	0.0419	0.021	0.0505	-0.034	0.0530
Non-treated*2010	-0.024	0.0430	-0.027	0.0419	-0.029	0.0505	-0.020	0.0530
Non-treated*2011	-0.002	0.0429	-0.006	0.0420	0.018	0.0502	-0.086	0.0531
Non-treated*2012	-0.010	0.0424	-0.007	0.0413	-0.036	0.0496	-0.041	0.0525
Non-treated*2013	-0.030	0.0421	-0.063	0.0411	-0.052	0.0491	-0.068	0.0521
Non-treated*2014	0.011	0.0420	-0.041	0.0409	0.012	0.0489	-0.020	0.0519

Non-treated*2015	-0.016	0.0419	-0.031	0.0405	-0.028	0.0487	-0.072	0.0515
Non-treated*2016	0.014	0.0409	-0.027	0.0377	-0.002	0.0475	0.021	0.0492
Province fixed-effects	**yes**		**yes**		**yes**		**yes**	
Year fixed-effects	**yes**		**yes**		**yes**		**yes**	
Number of observations	**946**		**946**		**946**		**946**	

Appendix 2. Nested models and long-term effect with panel data

Let's assume that the starting point is a general model that nests the data generating process (DGP). It can be shown that the long-term coefficients estimated in the general overparametrized model (dynamic specification when the true DGP is static) are very similar to those estimated if the static correct model is specified.

We assume that the true unknown DGP is given by:

$$Y_{it} = \beta_1 X_{it-1} + \mu_i + \varepsilon_{it} \tag{1}$$

where "i" is the individual, "t" is time, and μ_i is the individual fixed effect. As usual, ε_{it} is the random term.

Since no a priori information on the dynamic structure of the model is available, the starting general model estimated by OLS is:

$$Y_{it} = \hat{\alpha} Y_{it-1} + \hat{\beta}_0 X_{it} + \hat{\beta}_1 X_{it-1} + \hat{\mu}_i + \hat{\varepsilon}_{it} \tag{2}$$

When N (the number of cross-section units) and T (the number of time periods) tend to infinity, we have:

$$plim(\hat{\alpha}) = 0 \tag{3}$$

$$plim(\hat{\beta}_0) = 0 \tag{4}$$

$$plim(\hat{\beta}_1) = \beta_1 \tag{5}$$

$$plim(\widehat{lte}) = plim\left(\frac{\hat{\beta}_0 + \hat{\beta}_1}{1 - \hat{\alpha}}\right) = \beta_1 \tag{6}$$

where lte is the long-term effect.

Hence, the estimated long-term coefficient in the overparametrized model tends to the true value of the coefficient.

When N tends to infinity but T is finite, Nickell (1981) showed that:

$$plim(\hat{\alpha}) \cong -\frac{1}{(T-1)} \tag{7}$$

From (6), it can be derived that under the orthogonality between Y_{it-1} and X_{it} :

$$plim\left(\widehat{lte}\right) = \frac{plim(\widehat{\beta}_1)}{1-plim(\widehat{\alpha})} \cong \frac{\beta_1}{1+\frac{1}{T-1}} = \frac{T-1}{T}\beta_1 \tag{8}$$

When the static model is estimated, the results are:

$$Y_{it} = \tilde{\beta}_1 X_{it-1} + \tilde{\mu}_i + \tilde{\varepsilon}_{it} \tag{9}$$

$$plim\left(\widetilde{lte}\right) = plim\left(\tilde{\beta}_1\right) = \beta_1 \tag{10}$$

In this case, the bias of the estimated long-term effect in the general model is:

$$plim\left(\widetilde{lte}\right) - plim(\tilde{\beta}_1) \cong \frac{\beta_1}{T} \tag{11}$$

When the dynamic model is estimated, the bias will be lower the larger the number of temporal observations. The implication of this result is that if the static model is valid and if the dynamic model nests the static one, both models will offer a similar estimate of the long-term effects.