Evaluating the Causal Economic Impacts of Transport Investments: Evidence from the Madrid-Barcelona High Speed Rail Corridor

Jose M. Carbo\textsuperscript{a}, Daniel. J. Graham\textsuperscript{a}, Anupriya\textsuperscript{a}, Daniel Casas\textsuperscript{a} and Patricia C. Melo\textsuperscript{b}

\textsuperscript{a}Department of Civil Engineering, Imperial College London, London, SW7 2AZ, UK; \textsuperscript{b}ISEG, University of Lisbon, Portugal

\section*{ARTICLE HISTORY}
Compiled December 3, 2018

\section*{ABSTRACT}
This paper evaluates economic impacts arising from the introduction of high-speed rail (HSR) between Madrid and Barcelona. Using difference-in-differences estimation we estimate an average treatment effect for provinces with stops on the HSR line of 2.4\% for economic output, 3.3\% for numbers of firms, and 1.1\% for labour productivity. We complement our DID results with a synthetic control analysis for Lleida and Tarragona, two provinces that we argue were assigned HSR stations largely due to their incidental location. We find that both the number of firms and labour productivity are substantially higher in these provinces than in their synthetic counterparts.

\section*{KEYWORDS}
High-speed rail; Economic Impacts; Causal inference

\section*{1. Introduction}
Investments in major transport projects are frequently justified on the grounds that they will induce positive transformational effects on the economy. This is evident, for example, in the cases currently being made for proposed high speed rail (HSR) investments in the UK and California. Critics of HSR often question whether such economic benefits will actually materialise and whether they will be of sufficient magnitude to warrant investment. Such uncertainty is understandable, because it is very hard to forecast a-priori how a ‘post-treatment’ equilibrium will pan out under substantially different transportation capacity. For this reason, it is valuable to conduct ex-post statistical analysis of the impacts of past relevant transport interventions to help inform future scenarios.

To this end, the substantive objective of this paper is to obtain empirical evidence of the economic impact of a major HSR investment. In so doing, we are contributing to a literature on the economic impacts of transport with numerous relevant theoretical and empirical studies. For reviews see Redding and Turner [10] and Melo et al. [8], as well as the brief summary provided in on-line supplementary materials to this paper. The key statistical challenge we face is to ensure that the evidence is causal in nature,
in the sense that it measures the actual impact of an HSR intervention net of other effects. This is challenging because transport interventions tend to be characterised by non-random spatial assignment, but the data we have available are non-experimental (or observational). In the case of HSR, the implication is that investments may tend to be endogenous to economic activity, and thus the key issue to be addressed in obtaining a causal understanding of impacts in the observational setting involves identifying and adjusting for confounding and selection bias. Aside from recent high quality work by Ahlfeldt and Feddersen [4] on the economic impacts of HSR in Germany, most studies of HSR have been insufficiently concerned with issues of causality [for a review see 5].

The subject of our analysis is the Madrid-Barcelona line of the Spanish HSR network. Spain has developed the second-largest HSR network in the world; the result of a large investment plan of the Spanish government and the EU over a 25 year period with huge cost. The Madrid-Barcelona connection alone, which covers 621 km and links the two largest cities in Spain, cost EUR 8.97 billion in total. A detailed description of the evolution of the Spanish HSR network is given in on-line supplementary materials. Our study is similar in objectives and approach to that of Ahlfeldt and Feddersen [4], but we consider impacts on a wider range of economic outcomes for a different case study.

The data we have collated for estimation are for 50 Spanish provinces (47 peninsular and 3 island) for the period 1995 to 2014. The three main categories of variable of interest are economic outcomes, treatment status, and provincial baseline characteristics (or covariates). The data used to represent variables in each category are as follows.

1. **Economic outcomes** - the outcome variables of interest are: economic output, measured as provincial Gross Value Added (GVA); labour productivity, measured as average GVA per employee; total employment; and the number of registered companies. These economic data are obtained from Instituto Nacional de Estadística (INE).

2. **Treatment status** - we classify Spanish provinces into control and treatment groups based on whether the province received an HSR service or not.

3. **Covariates** - to adjust for non-random assignment we require data on provincial economic characteristics that can predict future economic growth. Following the seminal work of Abadie and Gardeazabal [3], we have gathered data on education level, investment and sectoral decomposition of GDP. As a proxy for education we use average years of education. As an indicator of investment, we use gross capital formation over GDP. These two variables have been collected from the human capital, investment and capital stock series of Instituto Valenciano de Investigaciones Económicas (IVIE). We decompose GVA into agriculture, construction, manufacturing and service sectors; and include other relevant factors for economic growth such as GVA per employee, number of firms, employment, and average compensation per employee. Again, these economic data were obtained from INE.

The paper is structured as follows. Section 2 describes the methods used for analysis. Results form our application are presented in section 3. Conclusions are then drawn in the final section.
2. Methodology

We use difference-in-differences (DID) and Synthetic Control Method (SCM) approaches to quantify the causal impact of HSR on spatial economic performance.

2.0.1. Difference-in-Differences Method

We treat the data on Spanish provinces indexed by \( i \) \((i = 1, ..., N)\) over years \( t \), \( (t = 1, ..., T) \), as realizations of a random vector, \( Z_{it} = (Y_{it}, D_{it}, X_{it}) \) where \( Y_{it} \) represents our response variable, \( D_{it} \) the treatment status, and \( X_{it} \) is a set of covariates. Our estimand of interest is the Average Treatment Effect (ATE) defined as the difference in mean response under treated \((D=1)\) and control status \((D=0)\) respectively

\[
\tau = E[Y_{it}(1)] - E[Y_{it}(0)]. \tag{1}
\]

We divide the data on Spanish provinces into control and treatment groups, \( D_{it} \in \{0, 1\} \), and one of two time periods, \( T_i \in \{0, 1\} \), depending on the year in which the province received the HSR service. In time period \( T_i = 0 \) there are no treated units. DID estimation provides us with a valid framework to control for temporal and confounding biases, by making comparisons between control and treated groups in the pre-treatment and post-treatment period. We use the following DID specification to estimate the ATE from HSR

\[
\ln(Y_{it}) = \beta_0 + \tau(DID) + \sum_k \beta_k X_{it} + \sum_t \beta_t \gamma_t + \sum_j \beta_j \alpha_j + u_{it} \tag{2}
\]

where the \( \beta \)'s are parameters to be estimated, \( X_{it} \) is the set of \( K \) covariates added to adjust for systematic differences in control and treated groups in the two time periods, \( \alpha_j \) and \( \gamma_t \) are province and year specific dummy variables that capture provincial heterogeneity and yearly shocks respectively, the term \( DID \in \{0, 1\} \) equals 1 for treated provinces in post-treatment stage and 0 otherwise, and \( u_{it} \sim N(0, \sigma_u) \) is a random error. The covariates are gross capital formation over GVA (named as investment), average years of education, average compensation, and the weight in GDP of the agricultural, industrial, construction and services sectors. Our estimate of the ATE is given by \( \hat{\tau} \).

A critical assumption required for valid identification in DID estimation is that a ‘parallel trend’ exists, which requires that the treatment and control groups must follow the same trend over time in absence of the treatment \([\text{e.g. 6}]\), i.e.

\[
E[Y_{i,1}(0) - Y_{i,0}(0)|D_i(1)] = E[Y_{i,1}(0) - Y_{i,0}(0)|D_i(0)]. \tag{3}
\]

Some papers conduct a visual inspection for this assumption. Here we use the following standard test developed in the literature \([\text{e.g. 7, 9}]\) to select the corresponding control group

\[
\ln(Y_{it}) = \delta_0 + \sum_j \delta_j \alpha_j + \sum_i \delta_i \alpha_i \ast year_t + \sum_k \delta_k X_{it} + \sum_t \delta_t \gamma_t + v_{it} \tag{4}
\]

where, \( \delta_j \) are parameters to be estimated, \( X_{it} \) is the set of covariates mentioned previ-
ously, \( \alpha_j \) and \( \gamma_t \) are province and year specific dummy variables, and \( \nu_{it} \sim N(0, \sigma) \) is a random error. This equation compares the average response of the potential control provinces relative to the treated provinces in the pre-treatment years, that is, prior to 2003. In other words, this equation tests the null hypothesis that there is no significant difference in trend between the province specific interaction variables \( (\alpha_i \times \text{year}) \) and the base case (i.e. the trend of the treated provinces where all these interaction terms are zero). Those provinces which have statistically significant parameters associated with these interaction terms are not deemed suitable controls.

### 2.0.2. Synthetic control method

Under the SCM approach, our objective is to construct a synthetic control region (SCR) by weighting control provinces that resemble the relevant economic characteristics of the treated provinces, in this case Lleida and Tarragona, before receiving treatment. We choose Lleida and Tarragona because we argue that they received HSR stations due to their incidental location on the Madrid-Barcelona route, rendering their assignment to treatment likely exogenous to their economic performance. The control provinces did not have access to HSR stations during the period 1992-2015. After building our synthetic control province, we compare its economic evolution with that of the treated province.

Let \( J \) be the number of available control provinces and let \( K \) be the number of economic characteristics. We assign a weight to each control province. These weights are chosen so that the SCR better approximates the economic characteristics of the treated province. The optimal \( J \times 1 \) vector of weights \( W = (w_1, \ldots, w_J) \) is obtained from the following minimization problem:

\[
\min_W (X_1 - X_0 W)' V (X_1 - X_0 W) \text{ such that } \sum_{i=1}^{J} w_i = 1, \tag{5}
\]

where \( X_1 \) is a \( K \times 1 \) matrix containing all the economic characteristics of the treated province and \( X_0 \) is the \( K \times J \) vector containing the economic characteristics of the control provinces. The solution to this problem is a vector of optimal weights \( W^*(V) \), which should add up to one.

The optimal vector depends on \( V \), a \( K \times K \) diagonal matrix that assigns weights to each economic characteristic. Following standard practice [e.g. 1–3], we choose the matrix \( V \) so that the variable of interest in the treated province is best replicated by its synthetic counterpart before treatment. Let \( T_p \) be the number of years observed previous to the treatment. The problem of choosing \( V \) can be specified as

\[
\min_V (Z_1 - Z_0 W^*(V))' (Z_1 - Z_0 W^*(V)) \tag{6}
\]

where \( Z_1 \) is a \( T_p \times 1 \) vector with a time series of the variable of interest in the treated province, and \( Z_0 \) is a \( T_p \times J \) vector with a time series of the variable of interest in the control provinces.

Once we have built our SCR, we compare its evolution with respect to the treated province during the years following the treatment. Let \( T \) be the number of years observed after the treatment. We are interested in the gap between the treated province and its synthetic analogue, i.e. \( Y_1 - Y_0 W \), where \( Y_1 \) is a \( T \times 1 \) vector with a time series of the variable of interest in the treated province for the post treatment years, and \( Y_0 \)
3. Application

3.1. Results from difference-in-differences models

We perform two DID analyses, one using all provinces in the corridor as treatment units and one using only provinces with intermediate stops (i.e. excluding Madrid and Barcelona). Table 1 summarizes results for all treated provinces and Table 2 for intermediate provinces.

The results indicate that there have been positive and significant impacts from HSR on labour productivity, on economic output (i.e GVA), and on numbers of firms. We find that GVA per employee grew by approximately 1.1% more, on average, in the treated provinces than the controls. The effect is slightly larger, 2.1%, when we exclude Madrid and Barcelona. For employment, we find no statistically significant effects in either of our DID analyses. This suggests that the increases in labour productivity that we observe have come about via growth in GVA or improvements in labour efficiency, rather than through numbers employed. In fact, we estimate a statistically significant increase of 2.3% in GVA in all HSR treated provinces relative to controls, and a comparable estimate for the intermediate provinces of 2.4%. In terms of numbers of companies, we estimate an overall growth due to HSR of 3.3% across all treated provinces and 2.8% in the intermediate provinces, with both these estimates being statistically significant at the 99% confidence level. Growth in numbers of firms could have arisen either due to a net positive impact from HSR in the Spanish economy as a whole, or due to negative spillovers as firms relocate from adjacent control provinces, but we are not able to distinguish these mechanisms in our data.

3.2. Results from synthetic control models

3.2.1. Economic characteristics comparison

We use our SCM models to look in more detail at the evolution of two outcomes of interest, labour productivity and numbers of firms, for the provinces of Lleida and Tarragona. Table 3 compares the economic characteristics of these treated provinces to their synthetic counterparts and to the other available controls. Both Lleida and Tarragona have higher GVA per capita, higher average education, and a higher share of construction sector GDP than the average of the controls. On the other hand, Lleida has a lower number of firms and higher share of agriculture, while Tarragona has a higher number of firms and a lower share of agriculture than the remaining controls. Overall, however, the two SCRs provide a good match to the characteristics of the treated provinces.

3.2.2. Evolution of the number of companies and GVA per employee

Figure 1a shows the evolution of the number of firms in Lleida and in its SCR. The two red vertical dotted lines mark the years 2003 and 2007. The HSR line between Madrid and Barcelona was built in three stages. In 2003 the line connected Madrid and Lleida, in 2006 it connected Lleida and Tarragona, and in 2008 it connected Tarragona and Barcelona. Consequently, we focus on 2003 and 2007. The SCR provides an excellent
### Table 1.: Summary of DID regression results for all treated provinces in the Madrid-Barcelona corridor.

<table>
<thead>
<tr>
<th>Variable</th>
<th>log (avg. GVA per employee)</th>
<th>log (no. of employees)</th>
<th>log (GVA)</th>
<th>log (no. of companies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DID</td>
<td>0.0108**</td>
<td>0.0098</td>
<td>0.0227***</td>
<td>0.0329***</td>
</tr>
<tr>
<td></td>
<td>(0.0059)</td>
<td>(0.0081)</td>
<td>(0.0061)</td>
<td>(0.0082)</td>
</tr>
<tr>
<td>GDP share: service sector</td>
<td>-0.0408</td>
<td>-0.1688***</td>
<td>-0.3106***</td>
<td>-0.3459**</td>
</tr>
<tr>
<td></td>
<td>(0.0665)</td>
<td>(0.0855)</td>
<td>(0.0709)</td>
<td>(0.1728)</td>
</tr>
<tr>
<td>GDP share: industry sector</td>
<td>0.5056***</td>
<td>-0.8389***</td>
<td>0.1504*</td>
<td>-0.2084</td>
</tr>
<tr>
<td></td>
<td>(0.0632)</td>
<td>(0.0862)</td>
<td>(0.0855)</td>
<td>(0.1590)</td>
</tr>
<tr>
<td>GDP share: construction sector</td>
<td>-0.9062***</td>
<td>0.6167***</td>
<td>-0.1786</td>
<td>-0.2340</td>
</tr>
<tr>
<td></td>
<td>(0.1763)</td>
<td>(0.2169)</td>
<td>(0.2061)</td>
<td>(0.2428)</td>
</tr>
<tr>
<td>log(active population)</td>
<td>-0.0463*</td>
<td>0.4407***</td>
<td>0.3198***</td>
<td>0.4210***</td>
</tr>
<tr>
<td></td>
<td>(0.0238)</td>
<td>(0.0314)</td>
<td>(0.0253)</td>
<td>(0.0386)</td>
</tr>
<tr>
<td>log(avg. years of education)</td>
<td>-0.0306</td>
<td>0.2906***</td>
<td>-0.0235</td>
<td>0.1087</td>
</tr>
<tr>
<td></td>
<td>(0.0775)</td>
<td>(0.0835)</td>
<td>(0.0675)</td>
<td>(0.0863)</td>
</tr>
<tr>
<td>log(gross fixed capital formation)</td>
<td>-0.0265**</td>
<td>0.0415**</td>
<td>0.0716***</td>
<td>0.0846***</td>
</tr>
<tr>
<td></td>
<td>(0.0129)</td>
<td>(0.0163)</td>
<td>(0.0159)</td>
<td>(0.0183)</td>
</tr>
<tr>
<td>log(avg. compensation per employee)</td>
<td>0.6631***</td>
<td>-0.6556***</td>
<td>0.0975</td>
<td>0.0601</td>
</tr>
<tr>
<td>Province dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>N</td>
<td>399</td>
<td>475</td>
<td>342</td>
<td>368</td>
</tr>
<tr>
<td>R square</td>
<td>0.9827</td>
<td>0.9233</td>
<td>0.9928</td>
<td>0.8653</td>
</tr>
</tbody>
</table>

*Figures in brackets denote the standard errors associated with the estimates.
**Significance: (***) 99 percent, (**) 95 percent, (*) 90 percent.

### Table 2.: Summary of DID regression results for treated provinces in the intermediate of Madrid-Barcelona corridor.

<table>
<thead>
<tr>
<th>Variable</th>
<th>log (avg. GVA per employee)</th>
<th>log (no. of employees)</th>
<th>log (GVA)</th>
<th>log (no. of companies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DID</td>
<td>0.0206***</td>
<td>0.0048</td>
<td>0.0237***</td>
<td>0.0280***</td>
</tr>
<tr>
<td></td>
<td>(0.0067)</td>
<td>(0.0095)</td>
<td>(0.0071)</td>
<td>(0.0098)</td>
</tr>
<tr>
<td>GDP share: service sector</td>
<td>-0.0256</td>
<td>-0.1462</td>
<td>-0.2933***</td>
<td>-0.3393*</td>
</tr>
<tr>
<td></td>
<td>(0.0677)</td>
<td>(0.0903)</td>
<td>(0.0728)</td>
<td>(0.1727)</td>
</tr>
<tr>
<td>GDP share: industry sector</td>
<td>0.4237***</td>
<td>-0.7836***</td>
<td>0.1856*</td>
<td>-0.1107</td>
</tr>
<tr>
<td></td>
<td>(0.0700)</td>
<td>(0.0981)</td>
<td>(0.0998)</td>
<td>(0.1650)</td>
</tr>
<tr>
<td>GDP share: construction sector</td>
<td>-1.0463***</td>
<td>0.7640***</td>
<td>-0.1053</td>
<td>-0.0285</td>
</tr>
<tr>
<td></td>
<td>(0.1801)</td>
<td>(0.2298)</td>
<td>(0.2147)</td>
<td>(0.2515)</td>
</tr>
<tr>
<td>log(active population)</td>
<td>-0.0566**</td>
<td>0.4498***</td>
<td>0.3130</td>
<td>0.4128***</td>
</tr>
<tr>
<td></td>
<td>(0.0250)</td>
<td>(0.0349)</td>
<td>(0.0267)</td>
<td>(0.0400)</td>
</tr>
<tr>
<td>log(avg. years of education)</td>
<td>-0.0333</td>
<td>0.2710***</td>
<td>-0.0343</td>
<td>0.0924</td>
</tr>
<tr>
<td></td>
<td>(0.0807)</td>
<td>(0.0962)</td>
<td>(0.0704)</td>
<td>(0.0884)</td>
</tr>
<tr>
<td>log(gross fixed capital formation)</td>
<td>-0.0190</td>
<td>0.0317**</td>
<td>0.0656***</td>
<td>0.0762***</td>
</tr>
<tr>
<td></td>
<td>(0.0136)</td>
<td>(0.0180)</td>
<td>(0.0172)</td>
<td>(0.0194)</td>
</tr>
<tr>
<td>log(avg. compensation per employee)</td>
<td>0.6011***</td>
<td>-0.6412***</td>
<td>0.1482</td>
<td>0.2693</td>
</tr>
<tr>
<td>Province dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year dummies</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>N</td>
<td>361</td>
<td>418</td>
<td>304</td>
<td>340</td>
</tr>
<tr>
<td>R square</td>
<td>0.9826</td>
<td>0.9136</td>
<td>0.9924</td>
<td>0.857</td>
</tr>
</tbody>
</table>

*Figures in brackets denote the standard errors associated with the estimates.
**Significance: (***) 99 percent, (**) 95 percent, (*) 90 percent.
fit for the number of firms in Lleida prior to 2003. Immediately after 2003, the number of firms in Lleida grew faster than in the SCR. This divergence increased after 2007, when the HSR line was enhanced. We find that, by 2009, the number of firms in Lleida is 7% higher than in the SCR. This difference is represented in figure 1b, which displays the percentage difference between Lleida and synthetic Lleida.

The impact of HSR on the number of firms in Tarragona is similar, as shown in figure 2a. Synthetic Tarragona also provides a good fit for the number of firms in Tarragona before 2003. After 2003, the number of firms in Tarragona starts to diverge from its SCR, although the HSR service was not available until the end of 2006. This might indicate that there was an anticipation effect, with firms moving or being created in the province before the station was opened. After the station opened, the number of firms in Tarragona was 5% higher than in the SCR by 2010. For both Lleida and Tarragona, the gap with respect their synthetic analogues was reduced from 2010 onwards.

The labour productivity effects of HSR are also found to be positive and significant (see figures 3 and 4). The SCRs for both Lleida and Tarragona deliver a good fit before the treatment years, 2003 and 2007 respectively. Immediately after the introduction of HSR, there is a positive difference in GVA per employee relative to the synthetic control provinces. In Lleida, this difference is around 12% by 2010, and in Tarragona 4%. In both cases, the gap is reduced from 2010 onwards. Supplementary analyses to check the robustness of these results are reported in on-line supplementary materials.

### Table 3.: Companies

<table>
<thead>
<tr>
<th></th>
<th>Lleida</th>
<th>Synth for Lleida</th>
<th>Tarragona</th>
<th>Synth for Tarragona</th>
<th>Rest of controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Companies</td>
<td>28472</td>
<td>28560</td>
<td>47548</td>
<td>46331</td>
<td>33457</td>
</tr>
<tr>
<td>GVA per capita</td>
<td>18947.21</td>
<td>16661.82</td>
<td>21105.79</td>
<td>17400.22</td>
<td>13996.8</td>
</tr>
<tr>
<td>Investment</td>
<td>27.68%</td>
<td>27.81%</td>
<td>33.03%</td>
<td>32.81%</td>
<td>28.80%</td>
</tr>
<tr>
<td>Education</td>
<td>09.46%</td>
<td>09.63%</td>
<td>09.53</td>
<td>09.59</td>
<td>09.18</td>
</tr>
<tr>
<td>Agriculture</td>
<td>09.82%</td>
<td>07.19%</td>
<td>02.61%</td>
<td>05.60%</td>
<td>07.30%</td>
</tr>
<tr>
<td>Construction</td>
<td>10.82%</td>
<td>08.97%</td>
<td>10.62%</td>
<td>10.14%</td>
<td>08.81%</td>
</tr>
<tr>
<td>Industry</td>
<td>13.78%</td>
<td>19.19%</td>
<td>22.2%</td>
<td>21.44%</td>
<td>18.74%</td>
</tr>
<tr>
<td>Services</td>
<td>55.72%</td>
<td>57.93%</td>
<td>53.26%</td>
<td>53.02%</td>
<td>55.42%</td>
</tr>
<tr>
<td>Avg compensation</td>
<td>24581.4</td>
<td>24033.8</td>
<td>25897.14</td>
<td>23963</td>
<td>22036.04</td>
</tr>
</tbody>
</table>

(a) Synthetic Lleida and Lleida

(b) Difference with synth Lleida

Figure 1.: Number of companies, Lleida

To summarise our DID and SCM results, we find convincing evidence that the introduction of HSR between Madrid and Barcelona generated substantial positive effects.
Figure 2.: Number of companies, Tarragona

Figure 3.: GVA per employee, Lleida

Figure 4.: GVA per employee, Tarragona
on labour productivity and on economic output. We do not find significant effects on employment, and we are therefore led to believe that improvements in labour productivity have been driven by expansions in economic output, or in efficiency gains, rather than through employment change. Finally, we find evidence of increased economic activity in provinces that received HSR stops, as measured by numbers of firms, but we are not able to distinguish whether these gains represent net increase for the Spanish economy as a whole or displacement from control provinces. On-line supplementary materials provide additional comparison of our DID and SCM results as well as more on potential spillover effects.

4. Conclusions

This paper has quantified spatial economic impacts arising from the introduction of high-speed rail (HSR) between Madrid and Barcelona. Spain has the largest HSR network in Europe and the second largest in the world, but the effect of HSR remains understudied. Since HSR investments are endogenous to economic activity, estimating causal effects on the economy is a complex task. Most existing studies in the field have done little to address issues of causality. In this paper we follow two methodologies to quantify causal economic impacts. First, we use a difference-in-differences specification with province and year fixed effects. This provides estimates of average treatment effects for provinces with stops on the HSR line of 2.4% for economic output, 3.3% for numbers of firms, and 1.1% for labour productivity. We do not find significant effects on employment and therefore attribute labour productivity growth to expansions in GVA and in numbers of firms. Second, we assess the economic impact of HSR on the Lleida and Tarragona provinces through synthetic control methods. We argue that these two provinces received HSR stations due to their geographical location on route between Madrid and Barcelona, and thus their assignment to the HSR network can be considered largely exogenous to their baseline economic performance. We find that both labour productivity and numbers of firms are substantially higher in Lleida and Tarragona than in their synthetic counterparts due to the introduction of HSR.

References

