A NEW APPROACH TO MEASURE THE IMPACT OF HIGHWAYS ON BUSINESS LOCATION WITH AN APPLICATION TO SPAIN*

Teresa Garcia-Milà
Universitat Pompeu Fabra and Barcelona GSE

José G. Montalvo
Universitat Pompeu Fabra, ICREA-Academia and IVIE

June 2012

In this paper we present a new approach to measure the effect of infrastructures on firms' location. We use fixed size rectangles around the roads as the unit of analysis instead of the traditional approach based on regions or municipalities. We apply this approach to Spain, which is an interesting case since during the period 1984- 2000 there was a very active process of increasing the capacity of roads by transforming national roads into highways/dual carriageways. The results show that the transformation of national roads into high capacity roads did not increase the number of firms in the catchment area relative to the non-treated group.

JEL classification numbers: R4; O4; R58
Keywords: firms' location, infrastructure, geographic information system.

---

* Corresponding author: José G. Montalvo, email: jose.garcia-montalvo@upf.edu. The paper has benefited from the comments of participants on many conferences and seminars. The authors acknowledge the financial support of the Ministerio de Ciencia e Innovación (SEJ2007-64340). This research is not the result of a for-pay consulting relationship. We thank Xavier Vinyals, Sandro Shelegia and Cristina Vila for their excellent research assistance. We are grateful to the Institute of Geographic Studies and, in particular, Jordi Martín, for his collaboration in the construction and transformation of the data in several stages of the paper.

1. Introduction

Credible and sensitive methods to evaluate the effect of public infrastructures on economic development are critical to understand the economic relevance of these programs. This is especially important since very often governments use the construction of public infrastructures as a generic device to spur growth. There is a renewed emphasis on the beneficial economic effects of public infrastructures by international development agencies like the World Bank, and the continuing efforts of the EU to finance these kinds of projects. In 2010 the Obama Administration announced a "bold plan"[1] to renew and expand the infrastructures in the US which includes a 50 billion dollars up-front investment in six-years. This renewed interest for the investment in infrastructures should be matched with the search for credible analysis on its economic impact.

The results of the economic literature on this issue are controversial. The emergence of the literature on the effects of public infrastructure on productivity and growth took place around the end of the 80's and the beginning of the 90's. Most of the initial papers, focused on estimating aggregate production functions, found (incredibly) large elasticities of production to public capital. However, the model specifications and econometric techniques used by these initial papers had many pitfalls. Varying the level of aggregation of the data (national, regional, metropolitan), and using more sophisticated techniques to estimate production functions, the effect of public infrastructure on growth reduces drastically, and even tends to disappear. Even if one could rely on the results obtained in the estimation of aggregate production functions, it is under question whether that is the appropriate way to evaluate the impact of a new, or improved, infrastructure on the area that has received the investment.

In this paper we propose a microeconometric methodology to measure the impact of new highways on local economies, diverging from the mainstream studies in two dimensions. On the one hand, we focus on the exact location of new firms and the employment generated by them, considering as the unit of analysis a fixed size rectangles around roads/highways. In the past, most of the studies on the impact of infrastructures have taken the country, region, county or municipality as the

[1] These are the exact words included in the document of the Department of Treasury (2010).

geographical unit of reference. Instead, we divide the highways in rectangles of 20km long by 10km wide, and we allocate new firms/employment to each of these rectangles. As far as we know, this is the first time that the geographical area of reference is not a politically determined entity as a municipality but an exogenously constructed area[2]. This approach has many advantages and some disadvantages. In particular, production aggregates like GDP are calculated using the reference of political boundaries and, therefore, we will not be able to rely on those measures. Instead, we use the number of new firms and the new employment generated by them in each rectangle. On the other hand, we use matching techniques to control for the possibility of endogeneity in the decision of the investment and the location of new highways.

The outline of the paper is the following. In section 2 we review alternative procedures to evaluate the impact of public infrastructures and, in particular, highways. Section 3 presents the basic approach, the data and the estimation procedures proposed to analyze the effect of the transformation of Spanish national roads in highways. Section 4 presents the results obtained by applying our methodology to the Spanish case. Section 5 contains the conclusions.

2. The impact of public infrastructure

In this section we cover a brief summary of the literature on the effect of public infrastructures. We divide the section in two parts: methodological issues and the impact of public infrastructures in Spain. The discussion of the previous estimations for the Spanish case will lead naturally to the next section which discusses our estimation procedure.

2.1. The impact of public infrastructures: methodological issues

The traditional literature dealing with the economic effect of public infrastructures has relied on the estimation of aggregated production functions with public capital as an additional input of production. Although we can find early estimates in Eberts

[2] Except for the initial location, the beginning of all the segments is determined by the ad-hoc choice of the length of the segment.

(1986), it was not until Ashauer (1989)[3] obtained a very large output elasticity of public capital (larger even than the one for private capital) that studies on this topic suddenly spurred. The initial studies estimated Cobb-Douglas production functions using aggregate post-war time series data for the United States. The results obtained by Ashauer, supported by Munnell (1990a) and Lynde and Richmond (1993), were questioned by Aaron (1990) and Tatom (1991) arguing that the strong result was due to the spurious correlation that arises from the common trend of output and public capital. When these authors corrected for nonstationarity of the national time series, the estimates were not significant, and therefore the relationship washed out.

To partly overcome the time series problems, state level data were incorporated into the analysis. Munnell (1990b) and Garcia-Milà and McGuire (1992) obtain much lower values for the output elasticity of public capital when estimating state level production functions. Their analysis, though, has a potential endogeneity problem related to an omitted variable bias, state specific productivity shocks, not observable but correlated to the observable inputs. In Holtz-Eakin (1994), Evans and Karras (1994), Holtz-Eakin and Schwartz (1995) and Garcia-Milà, McGuire and Porter (1996), state specific productivity differences are taken into account through a panel data estimation that includes state-specific effects. The estimates for the output elasticity of public capital turn out to be very low or even zero. Furthermore, Garcia-Milà, McGuire and Porter (1996) find evidence that private capital could be measured with error, which in turn would put under question the panel data estimates. More recent work by Fernald (1999) finds that, as a consequence of the construction of the interstate highway system in the 1960's, there was a large increase of productivity in industries that are vehicle intensive (for instance gas utilities or, obviously, transportation) relative to industries that depend less on vehicles.

Since the middle of the 90's there has been a growing interest for the application of econometric methods that achieve cleaner identification than the traditional production function approach. These methods provide a credible approach to deal with sample selection and endogeneity problems. The number of applications of these methods to the evaluation of economic policies is increasing exponentially.

[3] It is interesting to notice that the document to justify the new US plan to invest in infrastructures (An economic analysis of infrastructure investment) still refers to this paper despite the long period of time passed since its publication.

By analogy with natural and medical sciences, the highest ranked methodology among the techniques that can prevent endogeneity problems is the experimental approach, which is based on generating a treated and a control group selected randomly and tightly controlled. Several well know experiments like Mexico's PROGESA program, the Proempleo program of Argentine or the vouchers program of Colombia have attracted a lot of attention among the economist. However, the purely randomized experimental approach cannot work easily in the context of the evaluation of the economic impact of public infrastructures. Only very recently Quintana and Gonzalez (2010) have used an experimental design of randomly assigned treatment (pavement covering) to analyze the effect of paving roads in Acayucan (Mexico). They show that newly paved roads raise the price of housing, which they interpret as an increase in living standards, which exceeded the cost of paving.

Despite this example, the strictly experimental approach is not easy to implement in the case of infrastructure and it is not feasible in the context of large public infrastructures. For this reason the literature on the effect of infrastructures has turned into natural or quasi-experiments[4]. In the context of the impact of public infrastructure on economic performance there are an increasing number of papers that deal with the endogeneity and sample selection issue by using either instrumental variables, matching methods, or assuming some kind of exogenous variation with respect to the construction of the infrastructure.

In the first line of research, some authors have searched for instruments to study the effect of the construction of large infrastructures. Hooks et al. (2004) consider the impact of the construction of prisons on total employment growth of the counties. They analyze data on new prisons from 1960 until 1994 and evaluate the impact of these infrastructures on the pace of counties' employment growth. They compare metropolitan and nonmetropolitan counties using a simple differences estimator. The methodology is not quasi-experimental in strict sense, since they do not deal with the endogeneity in the location of the prisons. Nevertheless, Hooks et al. (2004) use instruments (unemployment rate in 1970 and total housing units in 1950) to deal with

[4] Donaldson (2010) departs from these two methods by using a structural model to analyze the impact of the construction of the railroad in India. However, in the estimation, he uses the argument that the network of railroads was constructed for military reasons and not economic arguments. To address the possibility of endogeneity bias Donaldson (2010) also estimate the contrafactual effect of the railroads that were planed but never constructed.

that problem, although it is not clear how exogeneous is the variation of those instruments. Duflo and Pande (2007) analyze the effects on productivity of the construction of large irrigation dams in India. They also use instrumental variables. In particular, they argue that there is a non-monotonic relationship between the river gradient and its suitability for a dam construction. They claim that the river gradient can be used as a suitable instrument since it affects the possibility of having a dam but it does not affect directly productivity. They find a large increase in irrigated area and agricultural production in the downstream from the dams.

However, in the context of the analysis of the impact of the construction of highways in the US the ultimate instruments are the National Interstate Highway Plans prepared in the 40's. Many authors have used these plans as an instrument to analyze several types of infrastructures. The basic argument is that those Plans were basically designed to facilitate national defense and not to improve economic development in those areas, at least conditional on observable variables[5]. Baum-Snow (2007) shows that highways caused suburbanization in the US. To deal with the potential endogeneity of highway assignment he uses the number of highways planned in the 1947 National Interstate Plan as an instrument for the number of highways built. There are other recent studies that consider the US Interstate Highway System as a viable source for a policy experiment. Michaels (2008) uses the construction of highways to analyze the effect of reducing trade barriers on the demand for skills. Michaels (2008) considers the possibility that political or economic conditions may have affected the specific placement of highways in contrast with the original design. He proposes to use two instruments to deal with this problem: an indicator of having the highway planned in 1944 and the orientation of the nearest large city with respect to each county's geographic centroid. Michaels (2008) finds that highways facilitated the use of large trucks reducing barriers to trade across counties. However, highways had no effect on the demand for high-skilled workers relative to low-skilled in manufacturing.

A second alternative for identifying the effect of the construction of infrastructures is the use of quasi-experiment and matching techniques[6]. This avenue has been less frequent in the literature. Rephann and Isserman (1994) consider a

[5] The instrument is based on the intention to treat argument.

[6] Dehejia and Wahba (1999) show that, for instance, matching on the propensity score produces very similar results to a randomized experiment.

quasi-experiment to examine the impact of highways construction on counties, which obtained a link or are close to one. Their approach does not use matching on the propensity score but a different technique, which implies three steps. First, they apply a sequential caliper, then they calculate a similarity measure (based on the Mahalanobis distance) and, finally, they use the so-called "optimal matching" to set the twins (treated observation-control). Rephann and Isserman (1994) conclude that the benefits of the interstate system are concentrated on the areas close to large cities or with a high degree of urbanization in the pre-treatment period. Isolated rural areas and areas close to the interstate network do not receive any benefit.

A third alternative is to justify the existence of some exogenous variation in the data. This is the approach adopted by Chandra and Thompson (2000). They analyze the relationship between interstate highway construction and the level of economic activity. They argue that the construction of a new highway is endogenous in the case of metropolitan areas but it is unrelated with past economic performance in the case of non-metropolitan areas. As we have already indicated, the US Interstate Highway System was designed to connect major metropolitan areas in the US and to serve to national defense[7]. In principle the fact that a highway goes through a particular non-metropolitan area is a consequence of the need to link two metropolitan areas with the lowest cost. Therefore, for non-metropolitan areas the highway could be considered an exogeneous event.

Chandra and Thompson (2000) use this exogeneity assumption[8] to justify a quasi-experiment in which the treatment group is the non-metropolitan counties in the US that received an interstate highway and the control group is the non-metropolitan counties that never had an interstate. Chandra and Thompson (2000) also study the effect of a new highway on areas that are close to the counties that received the highway but the infrastructure does not cross their territory. Their findings show that non-metropolitan counties that received a highway experienced an increase in earnings compared to counties where the highway did not cross through. However, counties that where adjacent to highway counties suffered a reduction in retail trade and government earning. Chandra and Thompson (2000) claim that this finding explains why some authors find no statewide impact of public infrastructure on output.

[7] Federal Aid Highway Act of 1944.
[8] Chandra and Thompson (2000) present some evidence that support this assumption (page 482).

2.2. The impact of public capital: the Spanish case

The literature on the impact of public capital on the economic activity is very prolific for the case of Spain. This is probably in part the outcome of availability of high quality regional data on capital, both private and public, elaborated by IVIE[9]-Fundación BBVA, which distinguishes Spain from other European countries. The Spanish studies also started using aggregate time series, as is the case in Argimón et. al., (1993), and Bajo y Sosvilla (1994). Later work moves to regional production function estimates that use panel data techniques and in some cases control for spatial productivity spillovers (Mas et. al. (1996). Further work has been done estimating cost functions (Avilés et. al. 2001, Boscá et. al. 2002, Moreno et. al. 2002), or evaluating the effect of public capital on TFP (Cantos et.al. 2005).

When human capital is included, together with public capital, in the estimation of a production function (as in de la Fuente and Vives 1995), the impact of public capital is reduced and part of the productivity effect is shared with human capital.

Although there is also a large spread of results for the Spanish case, all available estimates, to our knowledge, find a positive (even if it is small) output elasticity of public capital that ranges between 0.07 to 0.02. Nevertheless, the drawbacks pointed out in the previous section on the estimations of production functions as a way to identify the impact of public capital on the economy, apply in general to all these studies.

Only recently Holl (2004) estimates the impact of road transport infrastructure on manufacturing location using a microeconometric approach. Holl (2004) considers a fixed effect Poisson specification to analyze if the municipalities that are closer to a highway attract more new business than the ones further away. Holl (2004) argues that highway construction can be assumed to be exogenous to changes at the municipality level because the decision about the route of the highway is taken at a higher governmental level. Using this exogeneity assumption, she regresses the number of new manufacturing establishments on proxies for intra and inter-regional demand accessibility, supplier accessibility and distance to the closest highway. Holl (2004) finds that highways affect the spatial distribution of new manufacturing

[9] IVIE stands for *Insituto Valenciano de Investigaciones Económicas*

establishments increasing their number in municipalities close to highways. She also finds sectoral differences in the attractiveness of municipalities close to highways.

3. A pseudo-experimental approach to the impact of highways.

Improvements in transport infrastructure have an ambiguous impact on local economies. They present the typical trade-off involved in market integration: new transport infrastructures can change the relative importance of concentrating factors (for instance market size and agglomeration economies) versus dispersing forces (factor costs and competition). In particular, better transport connections can make areas of lower economic activity more attractive for firm location (to gain access to market in core areas) but, at the same time, competition from firms in economic agglomerations may increase as they can now more easily supply locations at a distance, and benefit from cost and demand linkages. Due to its network character, transport infrastructure may lead to positive or negative spillover effect among closeby areas. In this paper we propose to measure the local impact of upgrading roads to highways considering the likely endogeneity problem in the upgrading decision. The basic units of analysis are segments of roads of 20 km-long and 10 km-wide. Our objective is to estimate the difference in the number of new establishments, or employment, in the catchment area of segments transformed into highways/dual carriageways versus the roads not upgraded. We consider that the transformation of roads into highways may not be independent of the characteristics of the areas transformed. For this reason we propose to treat this problem as a pseudo-experiment, and match the treated and control segments using their observable characteristics. In this respect our exercise is related with the approach in Rephann and Isserman (1994). However, there are significant differences. First, we consider a segment of a road as the basic unit of analysis while Rephann and Isserman (1994) work with counties[10]. In fact they restrict the treated counties to contain at least nine miles of interstate highway trying to capture counties with a relevant presence of a highway. Our approach guarantees that the unit of analysis corresponds to the area of influence of the highway since our segments are constructed on top of the roads and, therefore, the boundaries correspond exactly to the same distance from the main road in all the

[10] The construction of the segments, described in the next section, guarantees their exogeneity with respect to any "a priori" selection.

cases. Second, we consider the location of new establishments in each segment, and the employment generated by them, as the variable of analysis while Rephann and Isserman (1994) work with county income growth rates. Third, we compare several alternative matching estimators, including propensity score methods. We also perform an intense scrutiny of the segments, in form of formal tests, to make sure they are comparables and that the overlap condition is satisfied. Finally, we have larger samples than Rephann and Isserman (1994).

Our focus on the impact of transport infrastructure on firm location at a low level of geographic disaggregation is related with the approach in Holl (2004). However, Holl descends to the level of municipalities while we use segments of highways. Holl (2004) considers the potential simultaneity problem between firm location and the placement of new highways. To deal with this problem of potential endogeneity she runs a Poisson fixed effects estimation. We approach the endogeneity issue differently by using a matching estimator.

3.1. The pseudo-experimental approach.

Following the standard notation of the Rubin's causal model for potential outputs let's consider the existence of N units, i=1,…N, which in our case are segments of roads. Each unit has two potential outputs, $Y_i(0)$ for the outcome under the control and $Y_i(1)$ for the outcome under the treatment. Additionally there may be a set of pre-treatment variables, X. Each unit is either exposed to the treatment (W_i=1) or not exposed (W_i=0). Each observation is completely characterized by a triple (W_i, Y_i, X_i), where Y is the realized outcome.

$$Y_i = \begin{matrix} Y_i(0) & \text{if } W_i = 0 \\ Y_i(1) & \text{if } W_i = 1 \end{matrix}$$

Therefore, the problem is transformed into a missing variables situation in which one of the outcomes is always missing for each observation. We are interested in estimating the average treatment effect (ATE) of the transformation of a road into a highway[11] and, therefore, we want to evaluate the expression

$$ATE = E(Y(1) - Y(0))$$

[11] There may be other objects of interest like the average treatment effect on the treated, the population average treatment effect, marginal treatment effects, etc.

In the context of this framework there are two basic conditions for identification. First of all we need unconfoundness, which implies that the treatment and the potential outcomes are independent conditional on a set of X variables.

$$(Y(1)-Y(0)) \perp W \mid X$$

The term unconfoundeness is used mainly on statistics (Rosembaum and Rubin 1983). It corresponds to selection on observables (Fitzgerald, Gottschalk and Moffit 1998) or conditional mean independence. Secondly, we need the overlap condition. The overlap condition implies that the probability of receiving the treatment must be positive.

$$1 > \Pr(W = 1 \mid X) > 0$$

This assumption implies that when, for instance, we match observations in the treated and the control group we can always find observations in both groups with similar probabilities of having received the treatment.

To deal with the likely selection bias in the location of highways one can use several matching procedures. For instance, the propensity score technique summarizes the influence of the observables on the treatment in a scalar indicator. The propensity score, *e(X),* is the conditional likelihood of receiving treatment

$$e(X) = \Pr(W = 1 \mid X = x) = E(W \mid X = x)$$

This method balances the observed covariates between the treatment group and the control group. Therefore, after matching on the propensity score the control and the treatment groups are identical in terms of the observed characteristics. Obviously, if the set of X variables does not contain the most important determinants of the placement of the highways then it would not be possible to reproduce the results of a randomized experiment. Rosenbaum and Rubin (1983) show that under unconfoundeness, the matching of control and treatment groups generate an observational analog to a randomized experiment. In fact Dehejia and Wahba (1999) show, using the data of Lalonde (1986) from the US National Supported Work Demonstration (NSWD) program, that matching on the propensity score provides very similar answers to the original randomized experiment[12]. Once the matching has been conducted the evaluation uses differences, differences-in-differences or panel data techniques to estimate the average treatment effect.

The use of matching techniques has increased at a fast rate during the last years. Most of the applications are centered on the evaluation of active labor market programs (training, search help, subsidized employment, etc.). However, in the context of development economics there are several evaluations of the impact of the construction of roads in underdeveloped countries. For example social funds usually target poor communities. Some areas apply, some areas do not, and some others are rejected. Therefore, some areas receive a road while others in the same targeted area with poor infrastructure will receive nothing. Van de Walle (2002) reports the evaluation of a rural road in which using simplistic regressions of incomes of villages that get the program and those that do not, seems to indicate a large income gain due to road construction when, in fact, there was none. Van de Walle and Cratty (2005) compare the kilometers of road rehabilitated in communities that participated in an aid-financed project with a control group of communities that did not participated, finding no sign of any impact.

Another issue is the distinction between the analyses of the construction of new roads versus the improvement of already built roads. In the case of the impact of new roads the problem of finding suitable matches for treated units is quite difficult. Much easier is the issue of the improvement of old roads since it is eventually possible to find a matching with a non-improved road that can be used as a control. Any matching technique should include the socio-economic conditions of the areas covered by the improved road and the control road (population density, mix of agricultural/non-agricultural production, education level, sectoral composition of employment, etc.).

3.2. Empirical strategy and data

In this section we consider the application of the methodology proposed in the previous subsection to the evaluation of the impact of the transformation of national

[12] Lalonde (1986) had already shown that observational methods (like regression) provide the wrong answer due to a large selectivity effect. The non parametric nature of the matching method avoids the strong parametric assumptions implied by other methods of correction for selectivity like Heckman's two stages. In addition it is well known that the parametric methods are very sensitive to specification changes.

roads into highways in Spain during the period 1984-2000. Figure 1 presents the situation of the Spanish national roads/highways system in 2000. While in 1984 Spain had 2,286 Km. of highways and dual carriageways, in 2000 that number had increased to 10,443 Km., the major part of the change being an upgrade of national roads to high capacity roads. Since we want to analyze the transformation of a national road into a highway we are dealing with a simple situation in terms of constructing a matching for the treated areas. The basic idea is to analyze the number of new establishments and the creation of net employment around the new highways in comparison with the number of new establishments and net job creation around the non-transformed national roads.

[Figure 1 around here]

First of all we divide national roads and highways in segments of 20 kilometers. The Spanish system is basically radial, which means that most of the national roads and highways begin in Madrid and end in the coast. Therefore, we start measuring the segments from Madrid towards the endpoint in the coast. Since it is possible that the last part of the segment has less than 20 km we construct also a variable that measures the length of the segment. For roads and highways that do not begin in Madrid the rule to start the segments is from South to North. In order to check the robustness of the results we have constructed rectangles of two different sizes for the catchment area[13]: 10 km wide and 20 km wide[14].

For the matching of transformed and untransformed roads we are going to use several procedures, including the propensity score matching. As covariates we consider all the pre-treatment information available in the Census about demographics, labor market characteristics, sectoral mix and education in each municipality.

The proposed methodological approach requires of a database with the exact location of establishments to be able to determine how many firms are in the catchment area of the different segments of roads and highways analyzed. Probably

[13] We use the term catchment area in a loose sense. In general, a catchment area is not a fixed segment but depends on the capacity of the infrastructure to attract activities that are developed around it. We are going to use "catchment area" as a fixed area.

[14] We present the results for the 10km wide areas. The results are basically the same for the 20km wide rectangles. Results are available upon request.

the only database that contains such detailed information about location is the data of Duns & Brandstreet for Spain[15]. This dataset contains the *id* of more than 1 million active establishments, their exact address, their sector of activity (SIC, four digits), the year of birth of the firm, the number of current employees, an indicator for headquarter, the region, province and municipality as well as the zip code. The addresses were transformed into Universal Transverse Mercator (UTM) ED50 zone 30 coordinates by Geomarketing, Arvato Services using GSI software. In some cases (around 10%) the exact address was not available in the dataset, or the GIS system could not find it. In those cases, and since the postal code was available for 99,9% of the establishments, the UTM coordinates were calculated for the center of the centroid corresponding to the postal code.

Once the exact location of each firm has been determined, we assign them to their corresponding road segment. Some firms are not included in any of the segments and, therefore, they are not considered. In other cases one firm could be potentially included in two segments. This happens more often when the catchment area is 20 km and it is less frequent if the wideness of the segment is reduced to 10 km. If that is the case we allocate each firm to multiple segments of different road categories[16].

[Figure 2 around here]

Figure 2 presents the basic structure of our data in a graphical example. It shows the location of the establishments in one particular four digits sector, which has been chosen to have a low number of firms in the north-centre quadrant of Spain to simplify the graphical interpretation. The difference between the two figures is the size of the catchment area (10 km or 20 km). Catchment areas are classified in two groups: the ones that have been transformed into highways and the ones that have not been transformed.

3.3. Estimation procedure

[15] We use the 2003 edition of the Duns & Brandstreet data set for Spain.

[16] We have also explored alternative allocation criteria (the widest road, the smallest road, etc.). The results presented in the following pages are robust to those alternative criteria.

The estimation of the average treatment effect of the transformation of a national road into a highway/dual carriageway could proceed using several methods. We could estimate regressions for the two conditional functions derived from the transformed and untransformed samples or we could use a non-parametric estimator. We are going to adopt this second strategy using a matching estimator. Matching leads to consistent estimators under weaker assumptions than least squares estimation of a parametric regression.

The basic problem for the estimation of the effect of the transformation of a national road in a highway is the existence of missing potential outcomes corresponding to the situation that would have been observed had the road not been transformed (treated units had they not been treated) and the outcome of roads that were not transformed had they been transformed (not treated units had they been treated). Matching estimators solve the missing potential outcomes problems by imputing the missing value using the average outcomes of the nearest neighbors of the opposite treatment group. Basically, matching estimators impute the missing outcomes for transformed roads had they not been transformed by finding other segment in the data whose covariates are similar but that were not exposed to the treatment. This is similar to a non-parametric kernel regression with the number of neighbors defining the bandwidth of the kernel. There are two basic elements in the matching estimation: the definition of "distance" of a neighbor and the number of neighbors included in the imputation of the missing outcomes. The adjustment based on the distance can take into account all the covariates directly or use the propensity score. If we consider all the covariates then the imputation is performed in the following way. Let's consider the sample (W_i,Y_i,X_i) for i=1 to N. Define d_m as the distance from the covariates of unit i, X_i, to the m^{th} nearest match with the opposite treatment. Assuming that there are no ties then we include the observations such that m units of the opposite treatment are the nearest to unit i. Define $l_m(i)$ as the index l that satisfies that

$$\sum_{j:W_j \neq W_i} 1\left\{\| X_j - X_i \| < \| X_l - X_i \|\right\} = m$$

Where 1(.) is an indicator function. The $l_1(i)$ would be the nearest match to i. The set of the closest m matches is defined as $L_m(i)=[l_1(i), l_2(i), \dots l_m(i)]$. The outcomes are imputed using the following criteria

$$\hat{Y}_i(0) = \begin{cases} Y_i & if \quad W_i = 0 \\ 1/m \sum_{j \in L_m(i)} Y_j & if \quad W_i = 1 \end{cases}$$

$$\hat{Y}_i(1) = \begin{cases} 1/m \sum_{j \in L_m(i)} Y_j & if \quad W_i = 0 \\ Y_i & if \quad W_i = 1 \end{cases}$$

The simplest matching estimator can be calculated as

$$\hat{\delta}_m^{SM} = 1/N \sum_{i=1}^{N} [\hat{Y}_i(1) - \hat{Y}_i(0)]$$

Where m is the number of neighbors used to calculate the imputed values of the potential outcomes. This estimator is the one used in most of the literature on program and project evaluation, even though it is theoretically biased. However, if the number of controls is sufficiently large with respect to the treatment units this problem is not important. In addition, usually the actual bias is small. In the following section we compare the results of this type of estimators with the bias-corrected estimators proposed by Abadie and Imbens (2006).

The estimator depends on the definition of the distance used to characterize the "closeness" of the covariates. Three are the basic alternatives[17]:

a. the standard Euclidean metric: $\| x_i - x_j \| = (x_i - x_j)'(x_i - x_j)$

b. the generalized Euclidean metric: $\| x_i - y_j \| = (x_i - x_j)' diag(\Sigma^{-1})(x_i - x_j)$

c. the Mahalanobis metric: $\| x_i - x_j \| = (x_i - x_j)' \Sigma^{-1} (x_i - x_j)$

where x_i is a vector of characteristics of segment i, and Σ is the covariance matrix of the covariates.

An alternative way to measure closeness is based on a single indicator, the propensity score, instead of using all the covariates. The propensity score is the conditional probability of receiving a treatment

$$e(x) = \Pr(W_i = 1 \mid X = x) = \frac{\exp(X\beta)}{1 + \exp(X\beta)}$$

The propensity score can be used to weight the observations, block on it or use the score as a variable in a regression. We are going to estimate the treatment effect using

[17] The results of the following section are obtained using a generalized Euclidean metric but they are qualitatively unaffected by the use of any of the other metrics. Results are available upon request.

$$\hat{\delta}^{PS} = \frac{\sum_{i=1}^{N} \frac{W_i Y_i}{\hat{e}(X_i)}}{\sum_{i=1}^{N} \frac{W_i}{\hat{e}(X_i)}} - \frac{\sum_{i=1}^{N} \frac{(1-W_i) Y_i}{1-\hat{e}(X_i)}}{\sum_{i=1}^{N} \frac{1-W_i}{1-\hat{e}(X_i)}}$$

Following Hirano, Imbens and Rider (2003) this estimator will be efficient if the propensity score is estimated with a nonparametric estimator.

4. Results

Given the nature of our data we consider the comparison of several alternative estimations. First of all, we have constructed rectangles of two different sizes: 10km wide for 20 km long (10KM sample) and 20km wide for 20 km long (20KM sample) [18]. Notice that some of these rectangles may be shorter than 20 km long depending on their location and the limit of the coast. For this reason the variable that we analyze is the number of firms divided by the length of the segment measured in kilometers [19]. We have calculated the number of firms in each rectangle. However, when firms belong to the catchment area of several roads, we assign firms to multiple roads/highways. When we use the propensity score as the matching mechanism we estimate the following logit model[20]

$$\begin{aligned} e(x) = \Pr(T_i = 1 / X = x) = \Lambda(&\beta_0 + \beta_1 EDU1_i + \beta_2 EDU2_i + \beta_3 PART_i + \beta_4 UNEMP_i \\ &+ \beta_5 RAGRI_i + \beta_6 RCONS_i + \beta_7 RSERV_i + \beta_8 RWEMP_i + \beta_9 RNEMP_i + \beta_{10} DEN_i) \end{aligned}$$

where T=1 if the segment has been transformed in a highway and T=0 otherwise. The pre-intervention variables are the proportion of workers with secondary education (EDU1), proportion of university graduates among the working population (EDU2), participation rate (PART), unemployment rate (UMEMP), proportion of working population in the agricultural sector (RAGRI), proportion of construction workers in the working population (RCONS), proportion of workers in the service sector over total working population (RSERV), proportion of entrepreneurs with

[18] The results obtained with the 20km wide segments support the findings using the 10KM wide segments.

[19] During the remaining of this section we will refer to the treatment effect in terms of “firms per 10 square kilometers”, which has to be interpreted as the number of new firms per rectangle of 1km by 10km.

[20] The same explanatory variables are included when using alternative definitions of distance.

workers in the total working population (RWEMP), proportion of entrepreneurs without workers (RNEMP)[21] and population density (DEN). Initially we consider the characteristics of the municipality where the segment starts[22].

[Table 1 around here]

In order to check the sensitivity of the results to the characteristics of the municipalities in the different segments we consider two alternative choices: the municipality where the segment begins and the municipality where it ends. Table 1 presents the basic descriptive statistics for both choices. The first fact reflected by table 1 is the small difference in the average characteristics of the municipalities where segments begin and end. The average level of education of the municipalities at the beginning of the sample period is low, with a proportion of primary education (complete or partial) reaching 40% and a proportion of tertiary education around 5.9%. The participation rate of the municipalities included in the sample is 48% with an average unemployment rate reaching 18.5%. The productive structure shows a large agricultural sector, at least compared with the current situation. The proportion of entrepreneurs with employees is low (around 5%). Finally, density is high compared with the average density of Spain. Notice that national roads and highways do not run through high mountains and places with difficult access. In addition, notice that we are assigning the characteristics of the municipality to the whole segment.

Table 2 shows that the test of the mean propensity score by blocks (five) cannot reject the null hypothesis that the characteristics of the segments are appropriately balanced for treatments and controls. The tests of differences in means for the balancing property by blocks are satisfied (reject any significant difference) in all the blocs[23]. Figures 3 and 4 show a high degree of overlap in terms of pre-intervention

[21] Notice that the last two variables add up to the proportion of entrepreneurs in the working population and not to 1

[22] We have also used other variables in the logit model such as the number of exits in each segment of road/highway or the distance to the closest city larger than 250.000 inhabitants. The results are basically the same. Since there are very many possible exercises depending on the combinations of elements for the estimation (rectangle size, metric for matching, method of estimation, explanatory variables in the propensity logit, characteristics of the initial or final municipality, etc.) we have decided to show a limited number of exercises. The rest are available upon request.

[23] Results available upon request.

characteristics[24]. All these indicators point in the direction of having a well-balanced matching of treatments and controls using the propensity score.

Tables 3 to 5 present the estimation of the treatment effect of the transformation of national roads into highways. We analyze two types of interventions. The first type of intervention is summarized in T, which takes a value equal to 1 if a national road has been transformed into a dual carriageway or a highway[25] and 0 otherwise. The second intervention, T1, is equal to 1 if a national road was transformed into a highway/dual carriageway or a regional road into a national road, and 0 otherwise. We present estimations for segments of 10 Km wide, although the results for segments of 20 Km. wide yield similar qualitative results. We calculate three different matching estimators using the socio-economic characteristics of the municipality at the initial, and at the end part of the segment.

We use two different measures to capture the effect of the treatment, the net creation of firms (Table 3) and the net creation of employment[26] (Table 4). For both measures we estimate the treatment effect restricting the sample to only those segments with no missing values, that is with at least one firm in the segment (NM), and also we estimate the effect with the full sample including those segments that have no firms and therefore no employment (All), assigning to them a value of zero firms/employment. The last three columns in each table present the matching estimators. In all cases we have matched each treated segment with four controls. Abadie and Imbens (2006) show that four matches perform extremely well in terms of mean squared error. The first of these columns, (m(4)) uses a generalized Euclidean distance as the metric for proximity. The second column (BC) contains a bias-corrected estimator following the proposal in Abadie and Imbens (2006). The correction uses the same regressors included in the logit model. Finally, the third column presents the results of the estimation using propensity score matching. In this case the estimation of the standard deviation of the average treatment effect is obtained by bootstrapping.

[24] These results are common to all the estimations based on the propensity score presented in the following paragraphs.
[25] In previous versions of the paper we considered in the control group all the segments that were untransformed.
[26] The employment is also divided by the actual length of the segment to control for the fact that not all the segments are 20 km. long.

[Table 3 around here]

In Table 3 we are interested in the treatment effect of the transformation of national roads into highways/dual carriageways, measured as the number of new firms located in those transformed segment versus the control group (untransformed segments) over the size of each segment[27]. The first row of table 3 presents the estimated average treatment effect considering transformations over the whole period analyzed but taking into account only the segments with at least one firm within the segment. The point estimator is always positive. When we use four matching controls for each treatment and the generalized Euclidean distance the effect of the intervention is almost 26 firms per 10 square kilometers, which is statistically insignificant[28]. The bias-corrected estimator leads to a very similar result, with the effect of the intervention being 28 firms per 10 km^2 and a coefficient that is statistically not significant. Finally, the last column contains the result of the estimation matching on the propensity score. The average treatment effect is higher than the other two estimates, a little above 35 firms per 10 km^2, but still not statistically significant.[29] Results using the socio-economic characteristics of the municipality at the end of the segment yield similar results, if anything slightly smaller in the estimated coefficients, which are statistically insignificant. The next two rows consider all segments, including those that have no firms, with estimated values lower and still insignificant. The second panel of Table 3 presents the results for an intervention (T1) that adds to the treatment the transformation of regional roads into national roads. The estimations obtained are qualitatively similar, with estimated values that tend to be smaller than their counterparts in the first panel, with the exception of the estimate for non-missing, initial municipality, propensity score matching, that yields a point estimate value of 50 firms, higher than the others but still not statistically significant.

[Table 3 around here]

[27] Notice that, since the length of the segment is not identical in all the cases we need to divide by this length in order to get a comparable measure across segments.
[28] The results are also not significant if we use only one match for treated unit.
[29] Additional estimations, using the Mahalanobis criteria for distance, and the bias-corrected estimator with robust standard error estimation, yield similar results for the analysis presented in Tables 2 to 4.

Table 4 measures the effect of the intervention by the net employment created by the firms located in each segment, instead of simply counting the number of firms. The idea of considering this alternative measure is to correct for the potential bias that could arise if the firms located in certain areas were systematically smaller or larger than the firms in other areas. The table is divided in two panels; the top panel analyzes the effect of transforming a national road into a highway/dual carriageway (T), while the second panel takes also into account the transformation of regional roads into national roads (T1). The first two rows of the first panel present the sample estimated average treatment effect considering transformations over the whole period analyzed but taking into account only the segments with at least one firm within the segment, and therefore positive employment. When the covariates refer to the initial municipality the estimated effect of the intervention is between 350 and 480 jobs per 10 square kilometers, depending on the matching method used in the estimation. The results are in all cases statistically insignificant. When we consider the socio economic characteristics of the municipality at the end of the segment, we obtain slightly smaller estimates that are also statistically insignificant. If all segments are considered, including the ones that have no firms and therefore no employment, all estimates, when compared with their counterparts in terms of matching method or municipality criteria, are smaller and still insignificant.

In order to control for the possibility that it takes time for a road transformation to have an impact on the economy, we present estimates that restrict the treated segments to those that were transformed before 1994 (Table 4).

[Table 5 around here]

Table 5 present estimates that consider only those transformations that were constructed before 1994. This time break up coincides with the end of the Plan General de Carreteras 1984-1991, which in the data is clearly identified because they were many transformations completed in 1992 and 1993. All the transformations that occurred before 1994 were from national roads into highways/dual carriageway, with no cases of transformation of regional roads into national roads. Therefore the analysis for the period before 1994 presents results only for treatment T, given that T and T1 are identical. As could be expected, the estimated values presented in Table 4 are in all cases higher than their counterparts in the first panels of Tables 2 and 3. For the intervention effect measured in net firm creation (first half of Table 4), the estimated effect ranges from 24 to 54 firms, although in all cases the coefficients are statistically not significant. When the intervention effect is measured in terms of net job creation (second half of Table 4), the estimated impact ranges from 300 to 760 net jobs created, but again in all cases the results are statistically insignificant.

[Figure 3 around here]

The results of all the exercises show that the transformation of national roads into highways did not have a significant effect on the economic activity located around them. There are different explanations for this result. Many of the roads transformed were financed by the European Structural Funds (ESF) allocated to Spain. It is well-known that given the size of the funds corresponding to Spain, and that in large they had to be spent in Objective 1 regions, it was not always easy to find economically sensible infrastructures to justify the funds. Upgrading national roads was a simple way to spend the resources allocated to infrastructures in Spain. Therefore, the decisions on the roads to be upgraded were not based only on economic considerations. Besides the availability of ESF resources, mainly for Objective 1 regions, and some controversial criteria of geographical “fairness” in the distribution of highways were important in the decision of upgrading, as well as the consistence with the traditional Spanish road system of radial distribution. These criteria were quite important since the upgraded roads did not always coincide with the ones with the highest traffic index among all the potential roads.

We do not interpret the previous finding as implying that the construction of highways is irrelevant for business creation. The results indicate that the distribution of the location of business is not affected by the transformation of roads into highways. Since this is the implicit objective of many regional development programs, the finding of this paper imply that the transformation of roads into highways do not lead to the location of more firms around the new infrastructures would have they kept the simple road status. However, this result does not imply that the construction of highways is irrelevant for economic activity: the number of firms in all type of segments may be increasing as a consequence of the network externalities produced by the increase of the capacity of some of the roads.

4. Conclusions.

The evaluation of the impact of public infrastructures is a very important exercise, given that the size of the budget for public works is quite large in all levels of government. During a long time the estimation of aggregate production and cost functions has dominated the evaluation of the productive effect of infrastructures, despite the problematic nature of such estimations. Recently, economists have moved into the application of more credible methods of evaluation using pseudo-experiments and matching estimators.

In this paper we present a new approach to measure the impact of the transformation of national roads into highways with an application to Spain. During the period 1984 to 2000 there was a very active process of transformation of national roads into highways/dual carriageways. But, did the transformed roads attracted more firms than the untransformed segments? To answer this question we divide the Spanish national roads/highways system into 20-km long segments. Then, we use the GIS location of each new firm to assign it to the catchment area of one of these segments. Once we obtain the number of new firms in each segment we use several matching estimators to compare the number of new firms per squared kilometer in the transformed and untransformed segments. The treatment effect is statistically insignificant. Therefore, we can conclude that the transformation of national roads into high capacity roads (highways or dual carriageways) did not have an additional attraction effect of firms with respect to the segments that were not transformed.

These results imply that the transformation of roads into highways is not an effective way to stimulate a differential regional development in economically lagging areas. Our finding cannot be interpreted to imply that new infrastructures do not generate economic activity since both, transformed and untransformed areas, are attracting firms.

FIGURES AND TABLES

Figure 1. The Spanish system of national roads and highways (2000).

Red: National roads
Blue: Dual carriageways
Dark blue: highways.

Figure 2. Firm location and catchment areas

Catchment area: 10 km | Catchment area: 20 km

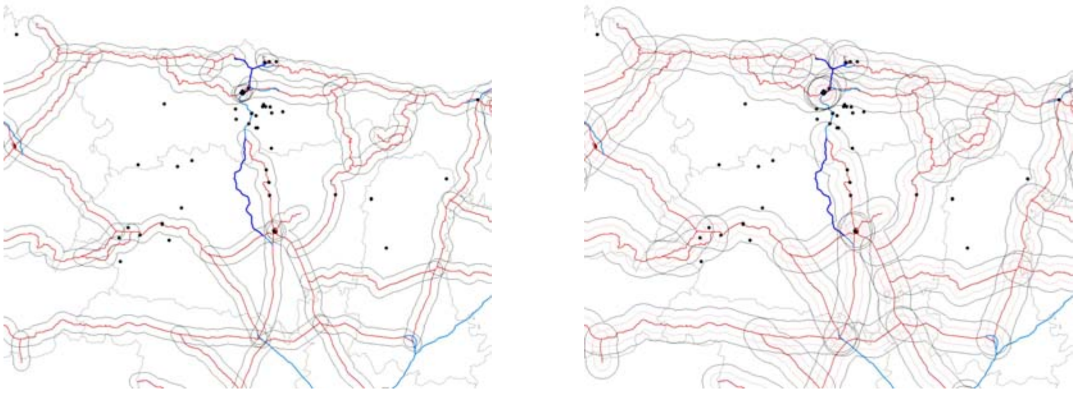


Figure 3. Support overlap in propensity matching estimation: 10 KM for non-missing segments

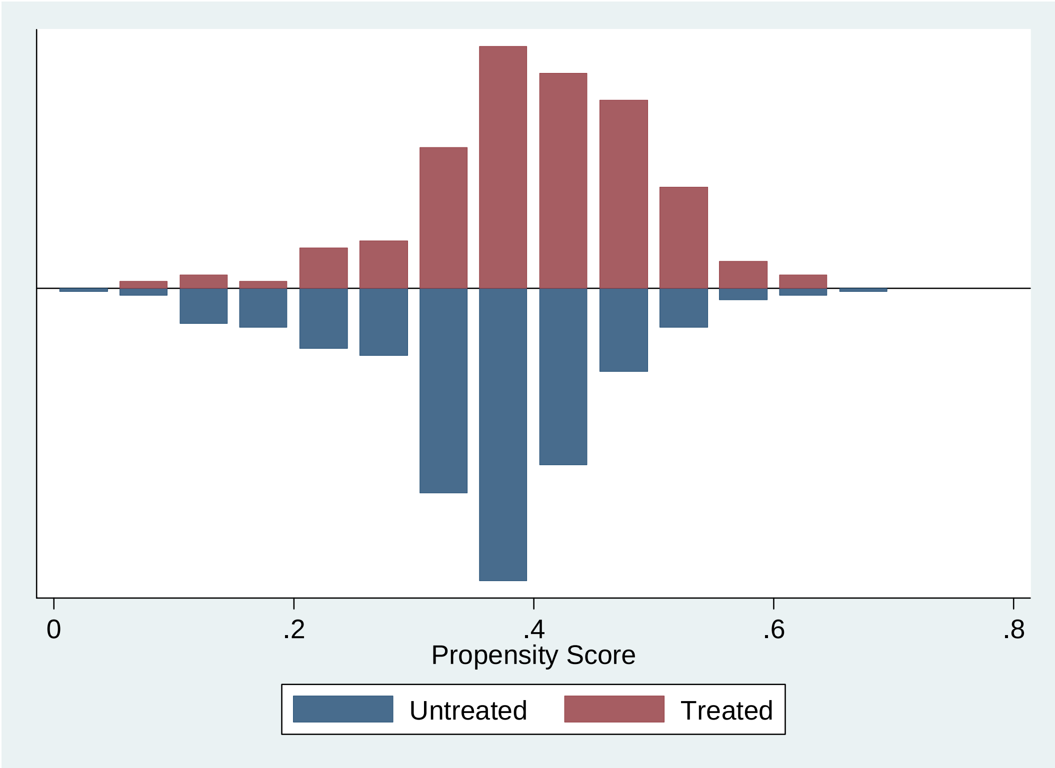
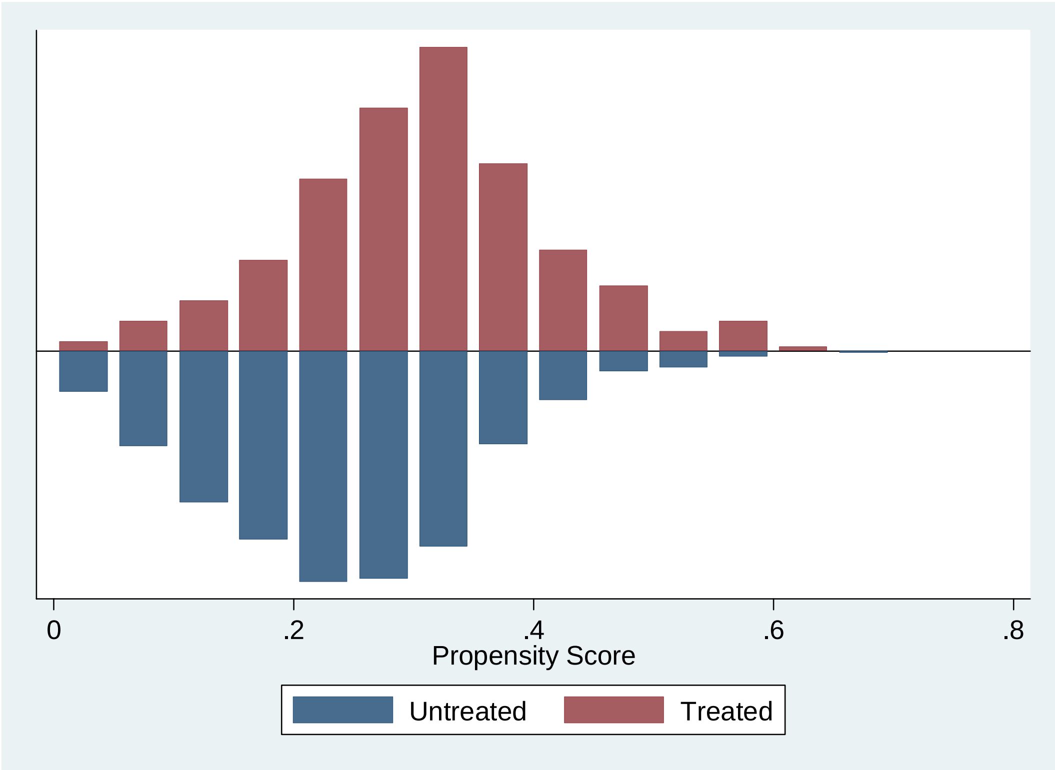


Figure 4. Support overlap in propensity matching estimation: 10 KM for all segments



REFERENCES

Aaron, H., (1990) "Discussion of Why is Infrastructure Important", in Alicia Munell (ed,) *Is there a shortfall in public capital investment?*, Boston: Federal Reserve Bank of Boston, 51-63.

Abadie, A., Diamond, A. and J. Hainmueller (2009), "Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program," *Journal of the American Statistical Association*.

Abadie, A., and G. Imbens (2006), "Large sample properties of matching estimators for average treatment effects", *Econometrica*, 74(1), 235-267.

Argimón, I., J. M. González-Páramo, M. J. Martín, and J. M. Roldán (1994). "Productividad e infraestructuras en la economía española". *Moneda y Crédito* 198, 207-245.

Ashauer, D., (1989), "Is public expenditure productive?", *Journal of Monetary Economics,* 23, 177-200.

Avilés, A., Gómez, R. and Sánchez-Maldonado, J. (2001): "The effects of public infrastructures on the cost structure of Spanish industries", *Spanish Economic Review* 3(2), 131-150.

Bajo, O. and Sosvilla, S. (1994): "Does public capital affect private sector performance. An analysis of the Spanish case, 1964-88", *Economic Modelling* 10(3), 179-186.

Baum-Snow, N. (2007), "Did highways cause suburbanization?," *Quarterly Journal of Economics*, 775-805.

Boscá, J.E., Escribá, J. and Murgui, M.J. (2002): "The effect of public infrastructures on the private productive sector of Spanish regions", *Journal of Regional Science*, Vol. 42, No. 2, 301-326.

Cantos, P., Gumbau, M., and J. Maudos (2005): "Transport infrastructures, spillover effects and regional growth: evidence of the Spanish case", *Transport Reviews*, 25(1), 25-50.

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

Dehejia, R. and S. Wahba (1999), “Causal effects in non-experimental studies: re-evaluating the evaluation of training programs,” *Journal of the American Statistical Association*, 94, 1053-1062.

De la Fuente, A. and X. Vives (1995). “Infrastructure and education as instruments of regional policy: Evidence from Spain”. *Economic Policy* 20, 11-54.

Department of Treasury and Council of Economic Advisors (2010), *An economic analysis of infrastructure investment.*

Donaldson, D. (2010), “Railroads of the Raj: estimating the impact of transportation infrastructure,” NBER working paper 16487.

Duflo, E. and R. Pande (2007), “Dams,” *Quarterly Journal of Economics*, 122 (2), 601-646.

Eberts, R.W. (1986), “Estimating the contribution of urban public infrastructure to regional growth”, Working Paper 8610, Federal Reserve Bank of Cleveland.

Evans, P and G. Karras (1994), “Are government activities productive? Evidence from a panel of US States,” *Review of Economics and Statistics*, 76(1), 1-11.

Fernald, J.G (1999), “Roads to Prosperity? Assessing the Link Between Public Capital and Productivity, *American Economic Review*, 89-3, 619-638.

Fitzgerald, J., Gottschalk, P., and R. Moffit (1998), “An analysis of the impact of sample attrition on the second generation of respondents in the Michigan Panel Study of Income Dynamics”, *Journal of Human Resources*, 33(2), 300-344.

Garcia-Mila, T., and T. J. McGuire (1992), “The contribution of publicly provided inputs to states’ economies”, *Regional Science and Urban Economics*, 22, 229-241.

Garcia-Milà, T., T.J. McGuire and R. Porter (1996), “The effect of public capital in state-level production functions reconsidered”, *The Review of Economics and Statistics*, 177-180.

Hirano, K., Imbens, G., and G. Ridder (2003), “Efficient estimation of average treatment effects using the estimated propensity score”, *Econometrica*, 71(4), 1161-1189.

Holl, A. (2004), “Manufacturing location and the impacts of road transport infrastructure: empirical evidence from Spain,” *Regional Science and Urban Economics*, 34, 341-363.

Holtz-Eakin, D. and A. Schwartz (1995), “Spatial productivity spillovers from public infrastructure: evidence from state highways,” *International Tax and Public Finance*, 2, 459-468.

Holtz-Eakin, D. (1994), "Public sector capital and the productivity puzzle," *Review of Economic and Statistics*, 12-21.

Hooks, G., Mosher, C., Rotolo, T. and L. Lobao (2004), "The prison industry: carceral expansion and employment in US counties," *Social Science Quarterly*, 58 (1), 37-57.

Lalonde, R. (1986), "Evaluating the econometric evaluation of training programs," *American Economic Review*, 76, 604-620.

Lynde C., and J, Richmond, (1993) "Public capital and total factor productivity", *International Economic Review*. 34-2, 401-14.

Mas, M., J. Maudos, F. Perez and E. Uriel (1996), "Infrastructures and Productivity in the Spanish Regions", *Regional Studies*, vol. 30-7, 641-649.

Michaels, G. (2008), "The effect of trade on the demand of skills- evidence from the Interstate Highway System," *Review of Economics and Statistics*, 90 (4), 683-701.

Moreno, R., Artís, M., López-Bazo, E. and Suriñach, J. (2002): "Public infrastructures and the performance of manufacturing industries: short-and long-run effects", *Regional Science and Urban Economics* 32, 97-121.

Munell, A.H., (1990b), "How does public infrastructure affect regional economic performance", in Alicia Munell (ed,) *Is there a shortfall in public capital investment?*, Boston: Federal Reserve Bank of Boston., 69-63.

Quitana-Domeque, C. and M. Gonzalez-Navarro (2010), "Street pavement: results from an infrastructure experiment in Mexico," Industrial Relations Section, Princeton University, Working Paper 556.

Rephann, T. and A. Isserman (1994), "New highways as economic development tools: an evaluation using a quasi-experimental matching method," *Regional Science and Urban Economics*, 24, 723-751.

Rosenbaum, P and D. Rubin (1983), "The central role of the propensity score in observational studies for causal effects," *Biometrika*, 70, 41-55.

Tatom, J., "Public capital and private sector performance" (1991), *St. Louis Federal Reserve Bank Review*, 73, 3-15

Van de Walle, D. (2002), "Choosing rural road investment to help reduce poverty," *World Development*, 30.

Van de Walle, D. And D. Cratty (2005), "Do aid donors get what they want? Microevidence of fingibility," Policy Research Working Paper, World Bank.

Table 1. Basic statistics for the characteristics of municipalities in the Census of 1991 at the beginning (end) of each road segments.

	Beginning	End
EDU1	39.37%	39.65%
EDU3	5.91%	5.95%
PART	48.36%	48.51%
UNEMP	18.49%	18.39%
RAGRI	18.82%	18.66%
RCONS	12.83%	12.77%
RSERV	17.62%	17.53%
RWEMP	4.77%	4.74%
RNEMP	17.29%	17.40%
DEN	684	746

Note: Proportion of workers with secondary education (EDU1), proportion of university graduates among the working population (EDU3), participation rate (PART), unemployment rate (UNEMP), proportion of working population in the agricultural sector (RAGRI), proportion of construction workers in the working population (RCONS), proportion of workers in the service sector over total working population (RSERV), proportion of entrepreneurs with employees in the total working population (RWEMP), proportion of entrepreneurs without employees (RNEMP) and population density (DEN).

Table 2. Test for equality of the propensity score within blocks.

	Not missing		All	
Block	Mean Difference	t-stat	Mean Difference	t-stat
1	0.008	0.36	-0.07	-0.8
2	-0.004	-0.59	-0.06	-1.63
3	-0.014	-1.91	-0.04	-1.19
4	0.005	0.25	-0.01	-1.35

Table 3. Estimation: segments of 10KM.Firm creation (net)

Intervention	Sample	Municipality	N	m(4)	BC	PS
T	NM	Initial	414	25.58	28.25	35.35
				(1.04)	(1.08)	(1.16)
		End	410	22.18	24.93	29.91
				(1.03)	(1.15)	(1.25)
	All	Initial	1005	16.35	12.98	14.38
				(1.53)	(1.21)	(0.97)
		End	1004	20.78	22.59	22.19
				(1.75)	(1.90)	(1.60)
T1	NM	Initial	433	23.58	25.33	50.37
				(0.94)	(1.02)	(1.40)
		End	429	15.13	17.97	20.01
				(0.74)	(0.88)	(0.69)
	All	Initial	1082	12.2	10.01	15.04
				(1.36)	(1.11)	(0.90)
		End	1082	14.83	16.96	20.46
				(1.48)	(1.69)	(1.53)

Note for tables 2 to 5: t-ratio between squared brackets. m represents the number of matches. BC: Bias-corrected estimator (Abadie and Imbens 2006). PS: estimation using propensity score matching. T: dummy variable that takes value 1 if a national road has been transformed in a highway/dual carriageway and 0 if it has not been transformed. T1: dummy variable takes value 1 if T is equal to 1 or a regional road has been transformed in a national road. NM considers only segments that have no missing values, that is at least one firm in each segment; All includes also the segments with no firms, to which assigns a zero. Initial considers socio economic characteristics of the municipality at the beginning of the segment; End takes the municipality at the end of the segment. N is the number of segments included in the estimation.

Table 4. Estimation: segments of 10KM. Employment creation (net in thousands jobs)

Intervention	Sample	Municipality	N	m(4)	BC	PS
T	NM	Initial	414	0.35	0.37	0.48
				(0.95)	(1.01)	(0.94)
		End	410	0.29	0.33	0.38
				(0.96)	(1.08)	(1.02)
	All	Initial	1005	0.2	0.16	0.18
				(1.40)	(1.12)	(0.94)
		End	1004	0.28	0.3	0.29
				(1.70)	(1.84)	(1.71)
T1	NM	Initial	433	0.29	0.33	0.31
				(0.84)	(0.95)	(0.78)
		End	429	0.2	0.24	0.26
				(0.72)	(0.86)	(0.72)
	All	Initial	1082	0.15	0.12	0.19
				(1.22)	(1.03)	(0.95)
		End	1082	0.19	0.22	0.28
				(1.42)	(1.63)	(1.64)

Table 5. Estimation: segments of 10KM. Segments transformed before 1994.

Dependent	Sample	Municipality	N	m(4)	BC	PS
Fims	NM	Initial	349	35.9	53.65	43.61
				(1.06)	(1.50)	(0.58)
		End	345	35.56	32.97	31.15
				(1.08)	(0.97)	(0.75)
	All	Initial	889	23.85	24.34	26.04
				(1.36)	(1.39)	(1.24)
		End	888	28.52	29.9	23.77
				(1.42)	(1.51)	(1.36)
Employment	NM	Initial	349	0.47	0.76	0.6
				(0.98)	(1.49)	(0.70)
		End	345	0.49	0.47	0.41
				(1.05)	(0.99)	(0.76)
	All	Initial	889	0.3	0.31	0.34
				(1.21)	(1.28)	(1.03)
		End	888	0.39	0.42	0.34
				(1.42)	(1.51)	(1.27)