Roads to innovation: evidence from Italy

Bottasso, Anna and Conti, Maurizio and Robbiano, Simone and Santagata, Marta

University of Genoa, Department of Economics

15 December 2020

Roads to Innovation: Evidence from Italy*

Bottasso A. Conti M.† Robbiano S. Santagata M.

University of Genoa

December 2020

Abstract

In this study we leverage on the ancient Roman roads network as a source of exogenous variation in order to identify the causal effect of the modern highways network on innovation using Italian NUTS-3 regional data. Our results suggest that a 10 percent increase in the highways stock in a region causes an increase in the number of patents of about 2-3 percent over a five years period. We document that this positive effect on innovation might in part be explained by a reduction in travel costs that foster collaborations between inventors living in different regions. We also find that the innovation enhancing effect of highways declines over time, possibly because of the introduction of ICT, or the increasing congestion over the Italian network. Finally, we find also evidence of important heterogeneous treatment effects associated to region population density and we cannot rule out the existence of negative spillovers across regions, suggesting possible reorganization of innovative activity across space.

1 Introduction

The role of transport infrastructure investments in fostering growth has been extensively studied in the economics and regional science literature. As documented by recent surveys, like Ferrari et al. (2019), and meta-analysis, like Melo et al. (2013) and Bom and Ligthart (2014), transport infrastructures have been found to display significant impacts on different economic outcomes. A reduction in transport costs associated to transport infrastructure investments can generate higher productivity of other inputs and lower production costs, can increase trade and competition by

*A previous version of this paper was included in the PhD dissertation of Marta Santagata and Simone Robbiano. The authors thank seminar participants at the 61 Conference of the Italian Economic Association. The usual disclaimer applies.

†Corresponding author. *E-mail address:* mconti@economia.unige.it

enlarging relevant markets and can favor the exploitation of scale economies. Furthermore, greater accessibility contributes to raise the market potential of different locations, thus affecting the spatial allocation of human capital and economic activities (agglomeration economies). Another channel through which transportation infrastructures can have important economic effects is the fostering of knowledge creation and diffusion process.[1] By lowering transportation costs, transport infrastructure may indeed make the interactions among inventors easier. This in turn tends to increase the spread of (localised) knowledge across space. Hence, it becomes crucial to understand if transport infrastructure investments, among other possible policy tools, are able to stimulate the innovation and knowledge spillovers that are often constrained by geography (Jaffe et al., 1993). However, with the notable exception of Agrawal et al. (2017), this issue has been substantially neglected by previous literature.

This study investigates the impact of road infrastructures (motorways) on the innovative capacity of Italian (NUTS-3) regions.[2] In particular, we exploit cross sectional variation at regional level in order to evaluate whether larger highway networks tend to make the spatial diffusion of knowledge easier, which in turn tends to foster innovative activity. In order to give a causal interpretation to our results we need to address a difficult identification issue linked to possible simultaneity between regional technological evolution and transport infrastructure investments. Indeed, such investments are typically not randomly allocated whenever governments tend to build infrastructures in low-income and low-innovation regions, or when high growth driven by local innovation fosters the demand for mobility and therefore the construction of highways. Moreover, there might be omitted factors that drive both infrastructure and innovation. In order to tackle this issue, we follow the historical route instrumental variable approach suggested by the urban and regional economics literature (Redding and Turner, 2015) and pioneered by Duranton and Turner (2012). This approach grounds on the idea that the presence of a transport network built in the past can be a good predictor for successive infrastructure investments.[3]

In this study we consider the ancient Roman roads dating back to 117 A.D. as an instrument for the modern motorways endowment of the Italian regions.[4] Following the literature, we argue that the Roman road network is reasonably exogenous, given that Roman roads were built mainly for military purposes; therefore, conditionally

[1]More recent endogenous growth theory models rest on knowledge spillovers as one of the most important engine of growth (Romer, 1990; Aghion and Howitt, 1990; Acemoglu and Akcigit, 2012).

[2]In this work the terms "region" and "NUTS-3 region" will be used interchangeably to indicate the Italian NUTS-3 statistical territorial unit.

[3]Studies that have followed this approach include Duranton et al. (2014), Duranton (2015), Agrawal et al. (2017), Baum-Snow et al. (2017) and Martincus et al. (2017).

[4]Such instruments has been firstly adopted by Garcia-López et al. (2015) when studying the impact of highways on the sub-urbanization of Spanish cities and have been successively employed in different studies like Percoco (2015), Holl (2016), Roca and Puga (2017), De Benedictis et al. (2018), Garcia-López (2019).

on a set of geographic controls, we assume that there are not important local unobservables, that explain both the construction of Roman Roads in certain areas and their (very) long run patterns of growth.

Following Agrawal et al. (2017) we estimate our model for the year 1988, which precedes the large diffusion of the Information and Communication Technology (ICT) and we let regional innovation performance depend on the motorways stock lagged five years. Indeed, the ICT revolution might have boosted or reduced the impact of road infrastructures on innovative activity, depending on the substitutability or the complementarity between personal interactions and ICT in the knowledge production process.

Our main result is that the stock of highways has a positive and significant impact on regional innovative activity. Conditioning on a set of geographic and inventor control variables, estimates suggest that an increase of 10% in the length of the motorways network leads to an increase of about 2-3% in regional innovation as measured by forward citation weighted sum of patents.[5] These findings are confirmed when we use unweighted patent counts and when we consider patents by region and technological fields as observation units; moreover, they are robust to the inclusion of a set of geographic and inventor control variables and to the use of an alternative measure of innovative capacity based on the address of the applicants instead of the inventors' place of residence.

Overall results can be associated to different transmission mechanisms, such as those related to agglomeration economies, working through the attraction of human capital, the enlargement of the relevant market and the diffusion of knowledge across space. Indeed, Agrawal et al. (2017) find that, in regions with denser highways networks, patent documents tend to cite inventors that are located in more distant areas within the same region. In this study we offer novel evidence supporting the importance of improved knowledge flows, by estimating a gravity model of innovation and we show that, across region-pairs, the number of collaborations tend to be higher in those region-pairs that have denser highway networks.

We also find that the impact of highways network tends to decline over time, either because of the diffusion of information and communication technologies or in the light of the large increase in traffic that occurred over the sample period. Moreover, we uncover evidence of important heterogeneous treatment effects, as we find that roads favor innovation particularly in those regions where inventors are more scattered over the territory: this is exactly what we should find if we believe that roads foster innovation by making communication easier within a region. However, in contrast to Agrawal et al. (2017), there is no evidence in favour of significant heterogeneous effects in sectors characterised by a different level of technological turnover. This result might be related to the low presence of firms characterized by high tech-

[5]Similar results are found by Agrawal et al. (2017) for US metropolitan statistical areas.

nological turnover in our sample. Finally, in some empirical specifications we find mild evidence of displacement effects. Such result is in line with the hypothesis of a spatial reorganization process of economic activity taking place, so that the increase of innovative activity in one region takes place, at least in part, at the expense of nearby ones (Redding and Turner, 2015). However, this issue would deserve further attention, possibly with a General Equilibrium analysis approach.

This study is organised as follow: Section 2 describes related literature, Section 3 presents our database and in Section 4 we describe the identification strategy. Empirical results are discussed in Section 5 which is followed by the conclusions.

2 Related literature

This study is related to different strands of literature. First, it fits to the wide and rapidly expanding literature on the effects of roads infrastructure on regional growth and productivity.[6] Within this context, Fernald (1999) is perhaps one of the first studies offering a convincing identification strategy that exploits industry differentials of sensitivity to transport costs: industries relying relatively more on road services should be particularly affected by improvements in the road network. By applying a variant of the Difference-in-Difference (DiD) identification strategy to a set of US industries over time, he finds that regional TFP growth was positively affected by the construction of the US highways system. Another important study which in turn pioneered the above mentioned historical route identification approach is Duranton and Turner (2012), who find that an increase in the stock of highways within US MSAs leads to an increase in employment of about 1.5% after 20 years. Moreover, authors suggest that their result is unlikely to simply reflect the spatial reorganization of economic activity. A recent paper by Ghani et al. (2016) evaluates the impact on productivity, employment, output and number of establishments of the so called Golden Quadrilateral (GQ) project, a recent major investment program which involved a massive upgrade of the GQ highways network in India. Authors finds that such investments significantly affected the growth of manufacturing activity, the number of firms and labour productivity.[7]

A more specialized literature this paper contributes to has in turn focused on the impact of highways and railways on innovation. In the seminal work by Agrawal et al. (2017), authors analyze the impact of interstate US highways on regional innovation. They apply an Instrumental Variable (IV) approach to deal with possible

[6] For a recent survey that focuses on studies employing counterfactual impact evaluation methods, see Redding and Turner (2015). Ferrari et al. (2019) provides an up-to-date review of the economic effects of transport infrastructure for each transport mode.

[7] Another interesting study on the impact of the massive investments in highways in China is the one by Xu and Nakajima (2017) who find a positive effect of highways construction on investment and output, with notable differences across types of regions and industries.

endogeneity of highway endowment and find that a 10% increase in interstate highways leads to a 1.7% increase in regional patenting activity over a five years period. In particular, they suggest that roads facilitate knowledge creation and diffusion also by favoring knowledge flows within metropolitan statistical areas; in particular, they show that the average distance between the location of a given patent and the patents it cites from the same region, is larger in MSAs with denser highways networks.

Following the Agrawal et al. (2017) approach, Wang et al. (2018) use an IV approach to examine the impact of road development on local innovation in China. In particular, to overcome endogeneity issues associated to roads endowment, they use city mean slope as an estimation instrument, arguing that the mean slope can be considered a proxy for the relative cost of road construction. Authors suggest that a 10% improvement in road density increases the average number of approved patents per company of 0.71%. Turning to railways, Yamasaki (2017) analyses the effect of rail access on the adoption of steam energy by analyzing the expansion of the Japanese rail network between the late 1800s and early 1900s. Using a Difference-in-Difference identification design together with an IV approach, the author suggests that the growth of rail access from 1888 to 1892 accounts for 67% of the growth of steam energy observed over the period 1888-1902.[8] Lin (2017) estimates a *DiD* model on a panel of Chinese cities observed over the period 2003-2013 in order to assess the impact of high-speed rail (HSR) on a number of economic outcomes, including patent applications, as a proxy for innovation activities within a city. Among other results, authors find that high-speed rails stimulate innovation by favoring greater scientific collaboration between cities and diffusion of knowledge. Finally, in Dong et al. (2020), the relation between knowledge diffusion and the construction of China's high speed rail is analyzed for the period 2006 - 2015. By instrumenting the construction of HSR with the Chinese railroad network in 1962 and the geographic slope for cities, authors show that, in Chinese cities connected to the HSR network, researchers experienced a significant increase in productivity, both in terms of quantity and quality of scientific publications.[9]

Finally, our study is related to the issue of promoting innovation at regional level. Starting with the seminal work of Jaffe et al. (1993), knowledge accumulation tends to be considered geographically localised. Indeed, innovation is fostered by several common features of the local "milieu", like clusters of high-tech firms, presence of research centres and by any other characteristic that may favour knowledge spillovers. Moreover, innovation benefits from local inter-firm alliances, mutual information and interactions between firms, scientists and specialised suppliers. Such relations promote knowledge flows and learning processes, thus allowing knowledge exchanges

[8] Authors construct their instrument by calculating the "cost-minimizing route" between destinations using slope information to account for costs of construction.

[9] See also Inoue et al. (2017), who use a Diff-in-Diff design to identify the effect of the opening the Nagano-Hokuriku Shinkansen high speed rail on firm level innovation in Japan.

of both formal and informal nature. Therefore, agglomeration processes favour the transmission of tacit knowledge, that can support the emergence of more stable and longer research joint projects (Baptista, 1998; Bennett et al., 2000; Love and Roper, 2001; Guillain and Huriot, 2001; Hervas-Oliver and Albors-Garrigos, 2008). In particular, agglomeration is likely to reduce search costs, uncertainty (Feldman, 1999) and transaction costs associated to joint projects, so that firms can exploit the benefits of increasing returns from collaboration (Izushi, 2003; Abramovsky and Simpson, 2011).[10]

Our study contributes to the above-mentioned strands of literature in a number of ways. First, it is the first paper that analyzes the impact of road infrastructure on regional innovation performance using data for an EU country; in fact, to the best of our knowledge, empirical evidence on this issue has never been provided other than for US or China. Second, although we closely follow the seminal contribution of Agrawal et al. (2017) in terms of the identification strategy (historical route instrumental variable approach), our study is based on units of observations (NUTS-3 data) that significantly differ from US metropolitan statistical areas along various dimensions: Italian NUTS-3 regions, unlike US MSAs, always share borders with each other and are very heterogeneous in term of population density and economic development. Such characteristics allows us to better analyze the issue of spillover effects of transport investments on nearby regions at a quite narrow level of spatial aggregation. Moreover, while Agrawal et al. (2017) report evidence of knowledge flows within regions, our results suggest that more developed highways networks are associated to higher collaborations across regions.

3 Data

Our study analyzes the relation between roads and innovation, as measured by per capita weighted patent fractional counts, for 89 Italian NUTS-3 regions as defined in 1974.[11] In particular, in our analysis we rely primarily on fractional patent counts

[10]It is worth noting that New Economic Geography (NEG) literature proposes some theoretical models where location choices and growth are jointly determined. In their seminal work, Black and Henderson (1999) develop a model where urbanization generates geographically localised knowledge spillovers. Fujita and Thisse (1996, 2002, 2003) demonstrate the existence of a core-periphery equilibrium where firms and R&D facilities are geographically clustered. Baldwin and Martin (2004) show how growth affects geography by means of a cumulative causation processes of human, physical and knowledge accumulation; in the same spirit, Minerva and Ottaviano (2009) highlight the effects of public investments on growth and agglomeration patterns.

[11]Despite the number of Italian NUTS-3 regions has recently increased, we consider the 1974 local Administrative setting that counts 95 NUTS-3 regions: this is because information on motorways kms in 1983 is available only for the 1974 setting. Moreover, we drop from the sample six regions that do not have highways (Brindisi, Matera, Agrigento, Ragusa, Siracusa and Grosseto) so that our sample includes 89 regions. Indeed, dropped regions are equipped with road infrastructure similar to motorways, which however are not easily measurable for our sample period.

weighted for forward citations as a measure for innovation output.[12] The innovation literature (Pakes and Griliches, 1980; Agrawal et al., 2017) has identified forward citations as an indirect measure of the invention value, as the number of citations received by a patent reflects its importance in the development of subsequent technologies (Trajtenberg, 1990; Harhoff et al., 2003; Hall et al., 2005). The economic literature recognizes patents as fundamental instruments of appropriation of the innovative activity; moreover, technologies with greater impact on welfare and economic development are more likely to be patented (Pakes and Griliches, 1980). However, patents measure inventions but do not measure all innovative activity (Smith, 2005) and not all inventions are patented. Nevertheless, as argued by the innovation literature, patents are an effective measure of local technological capacity.

We recover annual data on patents from the European patent Office (EPO) repository (EPO - Patstat) that includes bibliographical and legal status patent data on several countries at NUTS-3 regions level. Patent data refer to patent applications filed directly under the European Patent Convention or to patent applications filed under the Patent Co-operation Treaty and designating the EPO (Euro-PCT). A detailed set of information on applications, like the number of forward citations, applicants and inventors and their characteristics, the relative technological IPC class of the patent and NACE-2 statistical classification of economic activity are included.[13] We recover patent data for the period 1978 - 2015 and we "regionalise" raw patent information by means of inventors address (NUTS-3 codes). Data are finally classified according to different technological fields, following the WIPO systematic technology classification, based on the codes of the International Patent Classification (IPC). In particular, we identify five patent classes, namely Electrical Engineering, Instruments, Chemistry, Mechanical Engineering and a fifth class including residual ones. Data are collected until 2015, since the two last years of available data underestimate application counts because of the delays in the publication of patent data (eighteen/twenty-four months since application).[14]

Turning to roads infrastructure regional endowment, we consider the total number of kilometres of motorways in each NUTS-3 region as provided by the Italian Central Institute of Statistics and the Automobile Club of Italy.[15] As far as data on the length of Roman roads is concerned, in particular for those defined as major/consular roads, we rely on the Digital Atlas of Roman and Medieval Civilization (DARMC), which provides georeferenced data at regional (NUT-3) level on the road network of

[12] The geographical distribution of patent applications is assigned according to the "inventor criterion", i.e. according the inventor place of residence. If a patent has more than one inventor, the patent application is distributed equally between all of them and consequently between their NUTS-3 regions of residence.

[13] WIPO IPC-based technology field classification. Source: WIPO IPC Technology Concordance Table.

[14] See Bronzini and Piselli (2016) for more details.

[15] See https://ebiblio.istat.it/SebinaOpac/resource/statistica-degli-incidenti-stradali/IST0010868.

the Roman Empire in 117 AD and we calculate the length of the major Roman roads in each Italian NUTS-3 region.[16] Figure 1 shows the resulting map; in particular, in the Italian peninsula the total length of major roads is almost 10,000 kilometres.[17]

Figure 1: Roman Road Network in Italy: Major Roads.

Source: Authors' elaboration from McCormick et al. (2013).

Following Duranton and Turner (2012) and Agrawal et al. (2017), we include in the analysis a complete set of control variables, like (NUTS-3) regions surface, the difference between maximum and minimum altitude and an index of terrain ruggedness.[18] Table 1 reports basic descriptive statistics for the main variables used in the study and Figure 2 shows the territorial distribution of weighted patents in 1988.

[16]The main predecessor of this database is Talbert (2000) which provides maps of the entire Greek and Roman empires, covering the territory of over 75 modern countries.

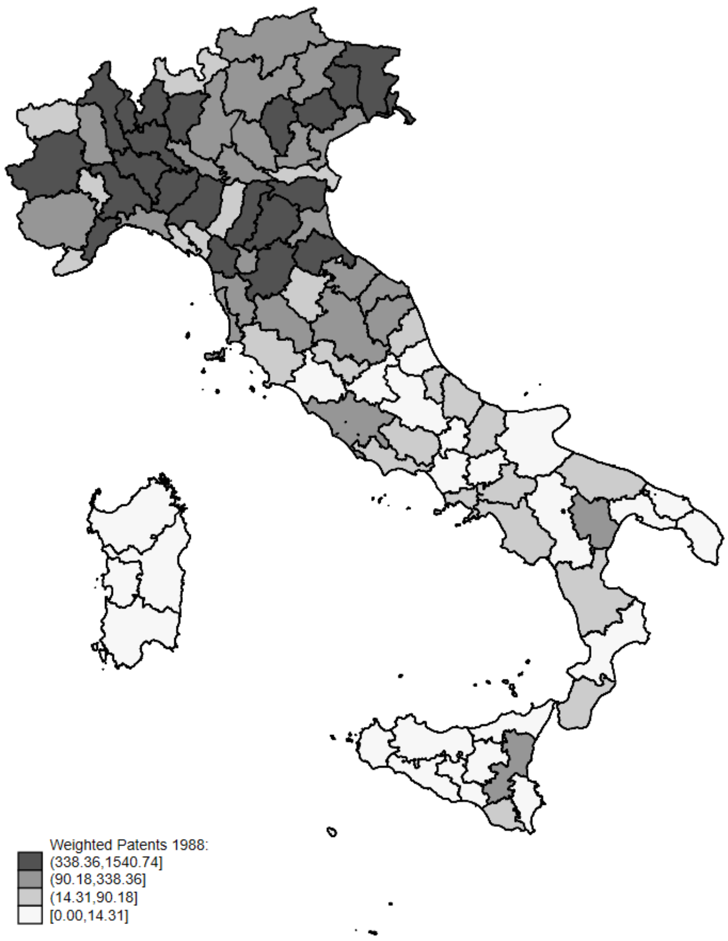
[17]Our calculation of the length of major Roman roads in each Italian NUTS-3 region is consistent with that of Licio (2020).

[18]Authors' elaboration from Nunn and Puga (2012).

Table 1: Descriptive Statistics.

Variables	Obs.	Mean	SD
Weighted Patent Fractional Count (per million people) 1983	89	127.39	235.50
Weighted Patent Fractional Count (per million people) 1988	89	251.58	332.15
Number of Inventors (per million people) 1988	89	21.01	24.95
Motorways lenght (km)	89	67.30	53.96
Major Roman Roads (km)	89	108.30	98.39
Surface (km^2)	89	2930.65	1594.01
Range (m)	89	758.77	510.56
Terrain Rugedness Index (*hundreds m*)	89	2.30	1.61

Figure 2: NUTS-3 Weighted Patent Fractional Count (per million people) 1988.



4 Identification Strategy

Following Agrawal et al. (2017), we consider a model where the innovative activity in each Italian region in 1988 is a function of the length of the motorways' system in 1983:

$$\log Innov_{i,1988} = \alpha + \beta(\log Motorways_{i,1983}) + \gamma(\log Innov_{i,1983}) + \varphi X_i + v_i \quad (1)$$

In the cross sectional model reported in Equation (1), which is consistent with a simple model where steady state innovation depends on motorways and innovation adjustment depends on its distance from the steady state, $Innov_{i,1988}$ refers to the natural logarithm of our innovation measure for region i in 1988, while $Motorways_{i,1983}$ is the length (logarithm) of the motorways' endowment for region i in 1983.[19] We also consider some alternative specifications of the model where we estimate the impact of the motorways network in 1983 on innovation performance in 1994, 2000, 2006 and 2012. Following a common practice in the innovation literature we also include in the model the lagged dependent variable as an input for future knowledge, in order to take into account the cumulative nature of the latter, as in Agrawal et al. (2017) and Dong et al. (2020).[20] One further advantage of controlling for the lagged dependent variable is that it should take into account most of the time invariant (or slowly moving) unobserved heterogeneity at the regional NUTS-3 level.[21] The empirical model includes also a set of inventors and geography control variables, X_i. In particular, we consider the number of inventors residing in each region as a proxy for the regional human capital and knowledge base, while geography controls include measures for surface, terrain asperity and elevation, whose inclusion should make the exclusion restriction associated to our instrument choice more likely to hold (see below).[22]

The β coefficient is the parameter of interest that reflects the impact of motorways endowment in 1983 on local innovative activity as of 1988, after controlling for innovative activity in 1983. Given the presence of the lagged dependent variable, Equation (1) can be interpreted in terms of the effect of motorways endowment on the growth rate of innovation between 1983 and 1988.

Equation (1) assumes that the analysis is conducted at an aggregate level, i.e. considering the sum of the patents originating in a given region, independently from patents sector or technology field. A more detailed analysis consider patents classified at a more disaggregated level according to their technology class, as in the following Equation, which includes also technology class fixed effects, μ_f:

[19] We provide an in depth explanation of the model in Appendix A.

[20] See also Aghion and Howitt (1990).

[21] Since the diffusion of knowledge is supported by face-to-face interactions favoured by the presence of highways, one might also expect to find a measure of market potential. Indeed, we believe that the presence of the lagged dependent variable should also control for this.

[22] Terrain asperity is an authors' elaboration from Nunn and Puga (2012).

$$\log Innov_{i,f,1988} = \alpha + \beta(\log Motorways_{i,1983}) + \\ + \gamma(\log Innov_{i,f,1983}) + \varphi X_{i,f} + \mu_f + v_{i,f} \tag{2}$$

where $Innov_{i,f,1988}$ refers to the natural logarithm of our innovation measure in region i for patents in technological field f as of 1988.

Another important issue that needs to be investigated within the evaluation of the economic impact of transport infrastructures is related to the possibility of displacement effects.

It might be that innovation is displaced from one region to another: in other words, an increase in the highways network in one region might just be simply diverting innovative activity from nearby regions (e.g. by attracting firms or inventors in the region from nearby ones) thereby generating a zero-sum game among regions. Therefore, in the spirit of Moretti and Wilson (2014) and Agrawal et al. (2017) we extend our main model of Equation (1) and we include a spatial lag in order to analyze if the impact of roads infrastructure investments generate spillover effects in nearby regions. Specifically, we estimate the following model:[23]

$$\log Innov_{i,1988} = \alpha + \beta(\log Motorways_{i,1983}) + \gamma(\log Innov_{i,1983}) + \\ + \theta SpatialMotorways_{i,1983} + \varphi X_i + v_i \tag{3}$$

where:

$$SpatialMotorways_{i,1983} = \sum_{i \neq j}^{I} w_{ij} \log Motorways_{j,1983} \tag{4}$$

The additional term $SpatialMotorways_{i,1983}$ represents, for each region i, a weighted average of the motorways stock in other regions j in 1983. Weights are built as a row normalized matrix of the inverse of the distances between any region i and j multiplied by the levels of innovative capacity, with elements $w_{ij} = (Patent_i * Patent_j)/Distance_{ij}$ such that $\sum_{i \neq j}^{I} w_{ij} = 1$.[24] Thus, the spatial term accounts for the the possibility that the infrastructural endowment in a certain region might have an impact on the performance of nearby regions. The θ coefficient associated to the spatial lag allows us to detect the nature of possible spillover effects.[25] A negative and significant sign of this parameter would suggest the presence of significant spatial displacement effects on regional innovation generated by road transport infrastructure investments in a particular region.

[23]As for Equation (1), also the Equation (3) is estimated not only at an aggregate level but also allowing variability between technological fields: $\log Innov_{i,f,1988} = \alpha + \beta(Motorways_{i,1983}) + \gamma(Innov_{i,f,1983}) + \theta SpatialMotorways_{i,1983} + \varphi X_{i,f} + \mu_f + v_{i,f}$.

[24]Distant regions receive a lower weight; moreover, the presence of the number of patents gives more weight to more innovative regions, so that our weight is a proxy of economic distance, according to Corrado and Fingleton (2012). In order to address possible endogeneity issues associated to the presence of the patent variable, we consider the innovative capacity at the beginning of the sample period, as suggested by Corrado and Fingleton (2012) and Bottasso et al. (2014).

[25]In the terminology of LaSage and Pace (2009), we estimate a Spatial X model.

Estimating the effect of road transport infrastructure investments on regional innovative capacity is a quite challenging task in terms of identification strategy since there might be simultaneity between regional technological evolution and transport infrastructure investments. In fact, such investments are typically not randomly allocated whenever policy makers tend to invest in lagging areas or in low-income and low-innovation regions, or when higher growth driven by local innovation fosters the demand for mobility and the construction of highways. Moreover, there might be omitted factors that drive both infrastructure and innovation. Possible correlation between unobservables, $v_{i,f}$, and the endowment of motorways of a given region in Equations (1), (2) and (3) would bring biased and inconsistent OLS estimates of the causal impact of motorways on innovative performance. To address such issue, we implement one of the approaches usually adopted in the applied literature on the economic impact of infrastructure as described by Redding and Turner (2015), namely the historical route instrumental variables one.[26]

In this work we use the ancient Roman roads network dating back to 117 AD as an instrument for the current motorways endowment of the Italian regions; moreover, for each region i, we build the instrument for the spatial lag presented in Equation (4) as a weighted average of Roman roads endowment in other NUTS-3 regions:[27]

$$SpatialRomanRoads_i = \sum_{i \neq j}^{I} w_{ij} \log RomanRoads_j, \tag{5}$$

This approach has been pioneered by Duranton and Turner (2011, 2012) that choose the routes of major expeditions of exploration between 1835 and 1850 and the major rail routes in 1898 as instruments for MSAs highways endowment. These variables have also been used in subsequent works, e.g. Duranton et al. (2014), Duranton (2015) and Agrawal et al. (2017).

Other ancient transport network measures have been proposed by the literature as instruments for current roads endowment. Baum-Snow et al. (2017) analyze how urban railroads and highways have influenced urban form in Chinese cities by using the 1962 Chinese transport network as instrument, while Martincus et al. (2017) consider the Inca roads built before 1530 as an instrument for the 2000s Peruvian transport infrastructure.

As far as the ancient Roman road network is concerned, Garcia-López et al. (2015) first use it as an instrument for the current transport system. Authors investigate the effect of highways on the suburbanization of Spanish cities by relying on an IV approach where the instrument is represented by Spanish historical roads, namely the

[26]The other approaches are the planned route instrumental variables and the inconsequential units approach.

[27]The weighting matrix has been constructed as the spatial matrix defined in Equation (4) using Roman roads instead of modern roads.

old Roman roads and the roads built by the Bourbons in the XVIII century.[28] Subsequent works that have used the roman road network within an IV approach include Percoco (2015), Holl (2016), Roca and Puga (2017), De Benedictis et al. (2018), and Garcia-López (2019).
The validity of the Roman roads network as an instrument for modern roads endowment has been largely discussed by the aforementioned studies, both in terms of relevance and exogeneity. Indeed historical transport networks should be relevant because modern networks are not built in isolation from them (Garcia-López, 2019) and this hypothesis has been tested by various studies. In particular a positive correlation between Roman roads and current Spanish highways has been shown by Garcia-López et al. (2015); Garcia-López (2019); Holl (2016); Holl and Mariotti (2018). Also Percoco (2015) and De Benedictis et al. (2018) have found a strong relationship between current and Roman roads network in Italy.[29]

Another condition that our instrument has to satisfy is the exclusion restriction, whereby it should affect regional innovation only through its effect on the current highway endowment and it should be independent from contemporaneous level of innovation activity at the NUTS-3 regional level. Indeed, the validity of the instrument requires its exogeneity conditional on controls and, according to previous literature, this requirement seems to be satisfied by our chosen instrument. In particular, as argued by Dalgaard et al. (2018), Roman roads are strongly predetermined and, more in general, almost any ancient transport network may be considered as exogenous because of the time that has elapsed since it was built (Duranton and Turner, 2012). Moreover, the literature has identified military reasons as the main purposes of Roman road construction, thus excluding a direct economic reason for their location (e.g. Garcia-López et al., 2015; De Benedictis et al., 2018). However, since geography may have influenced the construction of both Roman roads and modern motorways, in our empirical specifications we also control for a set of geographic characteristics in order to make our exclusion restriction more likely to hold, as in De Benedictis et al. (2018) and Garcia-López (2019).

5 Empirical Results

5.1 Main Results

We first estimate Equations (1) and (3) relying on a weighted measure of forward citation patent fractional counts aggregated at NUTS-3 regional level, without consider-

[28]For a previous application see Garcia-López (2012), where the author limits the analysis to the metropolitan area of Barcelona.

[29]Percoco (2015) uses Roman roads as an instrument for road accessibility, as measured by the presence of a motorway exit, while De Benedictis et al. (2018) use Roman roads as an instrument for modern roads endowment.

ing patent technological field. In column (1) of Table 2 we show basic OLS estimates, while in column (2) are included geographic and inventors controls. Parameter estimates show evidence in favor of a positive correlation between 1983 motorways stock and the weighted patent fractional count in 1988. In columns (3) and (4) we present estimates of Equation (3). The inclusion of the spatial lag does not affect the positive correlation between highways and innovation; the coefficient of $logMotorways_{i,1983}$ remains positive and significant, while the coefficient of $SpatialMotorways_{i,1983}$ is not significant.[30]

Table 2: The Impact of Motorways on Innovation: OLS Estimates and Reduced Form.

Dependent Variable: $\log Innov_{i,1988}$						
	(1)	(2)	(3)	(4)	(5)	(6)
$logMotorways_{i,1983}$	0.238***	0.252***	0.231***	0.249***		
	(0.0740)	(0.0728)	(0.0740)	(0.0740)		
$SpatialMotorways_{i,1983}$			0.699	0.269		
			(0.649)	(0.683)		
$logRomanRoads_i$					0.148**	0.144**
					(0.061)	(0.071)
$SpatialRomanRoads_i$						-1.232*
						(0.741)
Lagged Dep. Var.	YES	YES	YES	YES	YES	YES
Geography	NO	YES	NO	YES	YES	YES
Inventors	NO	YES	NO	YES	YES	YES
Observations	89	89	89	89	89	89
R-squared	0.684	0.704	0.686	0.704	0.682	0.692

Notes: All specifications are estimated by Ordinary Least Squares and include the lagged dependent variable. $\log Innov_{i,1988}$ is the weighted forward citation count of patents in region i in 1988. Inventor controls include the (log) number of inventors (per capita) in each NUTS-3 region in 1983. Geography controls include surface, terrain ruggedness and elevation. We add one to all patent, inventor and motorways before taking the log in order to preserve zero value observations. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

In order to interpret the positive correlation between road infrastructure and regional innovative activity in a causal way, we need to address possible endogeneity issues: as discussed above we rely on an instrumental variables approach.

First we estimate the reduced form of our model, i.e. a specification where regional innovative activity is let to depend on instruments, as well as on inventors and geography controls. Columns (5) and (6) of Table 2 reports OLS estimates. Results suggest that a denser Roman road network tends to significantly increase regional

[30]All estimated equations include the lagged dependent variable, whose coefficient ranges from 0.45 to 0.80 across the various specifications, consistently with Agrawal et al. (2017) results. Overall findings are broadly confirmed if we remove it.

innovation. Moreover, a higher stock of Roman roads in nearby regions is associated to lower levels of innovation, thus suggesting the possible existence of negative spillovers across regions.[31]

Table 3 presents IV estimates and shows that the coefficient of $logMotorways_{i,1983}$ is always statistically significant, even in the presence of control variables. As shown in columns (1) and (2), it has a magnitude of about 0.25 0.30, very similar to OLS estimates, while the inclusion of the spatial lag slightly increases its magnitude (columns (3) and (4)). Moreover, and consistently with reduced-form estimates, the spatial lag is negative and weakly statistically significant (at 10 per cent level of significance), thus denoting the existence of negative spillovers. This result suggest that a denser stock of motorways endowment in a region might negatively affect innovation activity in nearby ones, thus providing evidence in favor of displacement effects associated to road transport investments.

Table 3: The Impact of Motorways on Innovation: IV Estimates.

Dependent Variable: $\log Innov_{i,1988}$	(1)	(2)	(3)	(4)
$logMotorways_{i,1983}$	0.246**	0.297***	0.324**	0.405***
	(0.113)	(0.0982)	(0.133)	(0.128)
$SpatialMotorways_{i,1983}$			-2.532*	-2.694*
			(1.498)	(1.395)
Lagged Dep. Var.	YES	YES	YES	YES
Geography	NO	YES	NO	YES
Inventors	NO	YES	NO	YES
Observations	89	89	89	89
R-squared	0.684	0.703	0.644	0.664
F-statistic	31.30	24.62	10.09	7.365

Notes: All specifications are estimated by Two-Stage Least Squares and include the lagged dependent variable. $\log Innov_{i,1988}$ refers to the weighted forward citation count of patents in region i in 1988. Inventor controls include the (log) number of inventors (per capita) in each NUTS-3 region in 1983. Geography controls include surface, terrain ruggedness and elevation. We add one to all patent, inventor and motorways before to taking the log in order to preserve zero value observations. F is the First stage Kleinbergen-Paap Statistic. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

It is worth discussing the exclusion restrictions underlying our interpretation of IV results. In particular, in order to estimate Equation (1), we consider one excluded

[31]The estimation of the reduced form is insightful for various reasons. First, because it shows that the ancient Roman road network still has an effect on current regional innovation. Second, because the positive and significant effect of the Roman road network in the reduced form tells us that also current highways have a positive effect on regional innovation, given their positive correlation in the first stage equation (see below and Andrews et al. (2019) for an explanation). Finally, since the reduced-form is estimated with OLS, the estimates are not affected by possible weak instrument problems (see below).

instrument, namely the (log) length of ancient Roman roads in each region, to account for the possible endogeneity of current highways, while the estimation of Equation (3) requires an additional excluded instrument, i.e. $SpatialRomanRoads_i$ as defined in Equation (5). Therefore, in both cases we consider just-identified models, which prevents us from undertaking Sargan-type tests for the validity of excluded instruments. Nevertheless, on the basis of the discussion reported in the Identification Section, we argue that, conditionally on controls, the Roman road network is likely to be a valid instrument.

As far as the relevance of chosen instruments is concerned, we check first stage F-statistics reported at the bottom of Table 3.[32] Indeed, the robust F-statistic ("Kleibergen Paap rk Wald F statistic") exceeds all conventional critical values for the weak instrument test based on TSLS size. Moreover, F-statistics reported in columns (1) and (2) of Table 3 are also greater than Olea and Pflueger (2013)'s critical values at 10% of worst case bias, which are the appropriate critical values to consider in the non-homoschedastic case (see Andrews et al. 2019). [33]

Wrapping up, overall results obtained with different estimation techniques suggest that the regional endowment of highways has a positive and significant impact on innovation performance. In particular, estimates imply that an increase of 10% in the length of the motorways network in 1983 leads to an increase of about 2-3% in 1988 regional weighted patent fractional count, *ceteris paribus*.

5.2 Robustness Analysis

In order to verify the validity of our results we undertake a series of additional robustness checks. First, we account for the possibility that unobserved heterogeneity might be correlated across neighboring locations, thus leading to over-rejections of null hypothesis (Colella et al., 2019; Kelly, 2020). Following Colella et al. (2019), we estimate our basic specification of Equation (1) after allowing for arbitrary dependence of the errors across nearby observations and we compute standard errors corrected for cluster correlation in space.[34] In particular, we define a distance cutoff such that there are, on average, five regions in each spatial cluster. We believe that a 75 km cutoff is a reasonable value beyond which it should be safe to assume zero spatial correlation. We adopt both a uniform spatial decay kernel and Bartlett-

[32]First stage estimates show that the coefficients of $\log RomanRoads_i$ and $SpatialRomanRoads_i$ are always positive and statistically significant at 1% level in first stage regressions. Results are available upon request.

[33]For the case with more than one endogenous regressor and non-homoschedasticity (i.e. for the results displayed as columns (3) and (4) of Table 3), theoretical results are not yet clear cut (Andrews et al., 2019) and researchers still refer to the Tables reported by (Stock and Yogo, 2005) for the homoschedastic case.

[34]We use the Stata command *acreg*, that computes standard errors corrected for arbitrary cluster correlation in spatial and network settings. It implements a range of error correction methods for linear regression models, OLS and 2SLS.

type kernel and our main results are broadly confirmed (For details see Table B1 in Appendix B); however, evidence on the existence of spatial spillover is mixed and depends on assumptions of the spatial decay process.

Second, we replicate the analysis using two different measures for the dependent variable. In one specification we rely on an alternative measure of forward citation weighted patent fractional count which is based on the address of the applicants instead of the inventors' place of residence. Indeed, inventors and applicants (typically a firm) might not be located in the same region, so that we would end up attributing a patent to the region of residence of the inventor, although the invention was developed in a different region (where the inventor worked). It is reasonable to expect that such possibility is more likely where the highways network is more developed. Empirical results, displayed in Appendix C (Table C1), suggest that the impact of motorways is quite similar to our baseline estimates, thus suggesting that inventors commuting is not an issue in our data. We also replicate the analysis by using an unweighted patent fractional count measure as dependent variable. As shown in Appendix C (Table C2), estimated coefficients of $logMotorways_{i,1983}$ remain positive and significant, even though their magnitude is somewhat lower than in Table 3. It is worth nothing that, while results in Table C1 and in Table C2 clearly confirm the positive impact of highways on innovative activity, the spatial lag is no longer statistically significant, casting some doubts on the robustness of the negative spillover effects that we have uncovered in our baseline regressions.

In order to account for possible unobserved heterogeneity associated to different patents technology fields, we analyze the relation between the highways stock and weighted patent fractional count in each region i, as of 1988, in technological field f, as shown in Equation (2).
In particular, we consider five different technological classes, built on the basis of the WIPO systematic technology classification, namely Electrical Engineering, Instruments, Chemistry, Mechanical Engineering and Others (residual categories).[35] In Table 4 we report results obtained after replicating the analysis on observation units defined at region and technological field level. Estimated models include also technological field fixed effects and a specific control for the number of inventors in each field.[36] Results shown in Table 4 confirm our previous findings, even if parameter estimates are slightly smaller; in particular, the coefficients of the spatial lag variable are significantly negative.

The robustness analysis thus reveals a somewhat mixed evidence on the existence of spatial spillover effects, in contrast to the study by Agrawal et al. (2017), where no significant displacement effects are detected. A plausible explanation for

[35] The database is obtained by classifying patent fractional count according to technological classes. It is worth noting that not all regions in 1988 patented in all fields.

[36] As a consequence, inventors control variables include both the total number of inventors in each region and the number of inventors in each region in the specific field f.

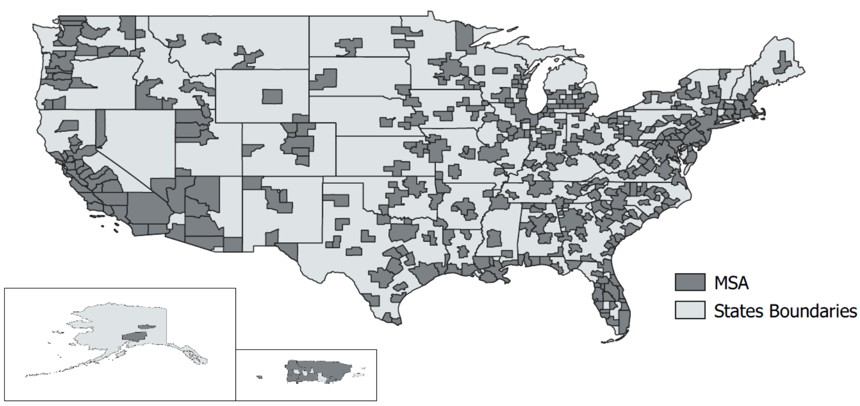
Table 4: The Impact of Motorways on Innovation by Region and Technological Field: IV Estimates.

Dependent Variable: $\log Innov_{i,f,1988}$ *by Region and Tech. Field*				
	(1)	(2)	(3)	(4)
$logMotorways_{i,1983}$	0.0328	0.202***	0.139	0.305***
	(0.118)	(0.0720)	(0.159)	(0.104)
$SpatialMotorways_{i,1983}$			-3.646**	-2.336**
			(1.552)	(1.062)
Lagged Dep. Var.	YES	YES	YES	YES
Field FE	YES	YES	YES	YES
Geography	NO	YES	NO	YES
Inventors	NO	YES	NO	YES
Observations	445	445	445	445
R-squared	0.388	0.539	0.295	0.512
F-statistic	22.99	21.17	9.729	7.202

Notes: Two-Stage Least Squares estimates. Lagged dependent variable is always included. $\log Innov_{i,f,1988}$ is the weighted forward citation count of patents in region i, in technological field f in 1988. Field FE refers to four technological fields classes. Inventor controls include the (log) number of inventors (per capita) in field f in each NUTS-3 region in 1983 and the total (log) number of inventors in each NUTS-3 region in 1983. Geography controls include surface, terrain ruggedness and elevation. We add one to all patents, inventor and motorways before to taking the log to preserve zero value observations. F is the First stage Kleinbergen-Paap Statistic. Robust standard errors, clustered at the NUTS-3 region level, in parentheses *** p<0.01, ** p<0.05, * p<0.1

this discrepancy could lie in the different characteristics of US MSAs if compared to Italian NUTS-3 regions.

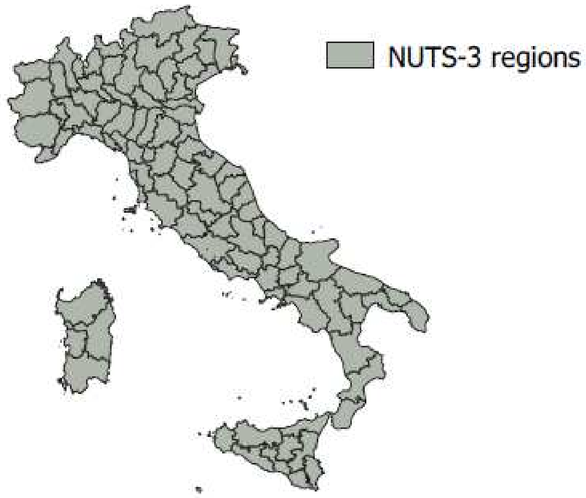
Figure 3: Metropolitan Statistical Area in the US.



Source: Authors' elaboration from US Census Bureau.

Indeed, as shown in Figure 3, US MSAs are more sparse across the country and not all of them share borders. In this context, the presence of road infrastructures facilitate ideas' circulation within the same MSA, but does not affect the probability

Figure 4: NUTS-3 Regions in Italy.



Source: Authors' elaboration from ISTAT.

that those ideas might be used by someone else in a different MSAs. On the other side, the geography of Italian regions (Figure 4) significantly differs from that of MSAs. Indeed, this administrative division covers the whole national territory and each region shares at least one border with another region. In this context, it seems more plausible that innovation gains associated to the motorways network in one region might be offset, at least in part, by losses in nearby ones. However, it is worth remarking that our results provide mixed evidence on displacement effects, since findings on spatial spillovers are not robust across different specifications. Moreover, it is also important to remember that in this study we primarily assess the local rather than the national impact of highways: however, as argued by Agrawal et al. (2017), conducting a precise evaluation of the effect of highways at national level is very challenging and would require a general equilibrium analysis.

5.3 Transmission mechanisms

As mentioned in the introduction, within the different channels through which transport infrastructure can affect innovation, we focus on the process of knowledge creation and diffusion, both within regions and across neighbouring ones. By lowering the costs of knowledge flows, higher highway stock may favor face-to-face interactions between inventors and may accelerate the circulation of ideas, thus generating important knowledge spillovers that are often constrained by space. In particular, we analyze whether more cross-regional travel possibilities promote interactions between inventors and better matching of ideas, thus generating an innovative process that is both qualitatively and quantitatively more significant.

Following Picci (2010), we develop a measure of innovative collaboration that rests on the usual notion of fractional counting of patents.[37]

In particular, consider $PFC_{i,a,1988}$ as the fraction of patent application a, assigned to region i in year 1988 according to the inventor place of residence.[38] Differently from Picci (2010), we consider not only the quantitative, but also the qualitative aspect of innovative collaborations, as proxied with the number of forward citations received by the patent, namely Cit_a. Hence, our forward citation weighted fractional measure of collaboration for each region i and patent a in 1988 is defined as:

$$Coauthor_{i,a,1988} = PFC_{i,a,1988} * Cit_a \tag{6}$$

For every patent application a and each region-pair $i-j$, the intensity of the collaboration between inventors is represented by the product of the forward citation weighted fraction of that application to the two regions:[39]

$$Coll_{i,j,a,1988} = Coauthor_{i,a,1988} * Coauthor_{j,a,1988} \tag{7}$$

Finally, the aggregate measure of collaboration between inventors of different regions is defined as the sum of Equation (7) for all patents:

$$Coll_{i,j,1988} = \sum_{a=1}^{A} Coll_{i,j,a,1988} \tag{8}$$

We then follow Dong et al. (2020) by relying on a cross-sectional structural gravity model framework (Anderson and Van Wincoop, 2003) to analyze the role of highways on the intensity of collaboration between pairs of regions:

$$\begin{aligned} \log Coll_{i,j,1988} = \alpha + \beta \log Coll_{i,j,1983} + \gamma \log Motorways_{i,j,1983} + \\ + \delta \log Distance_{i,j} + \varphi_i + \varphi_j + \epsilon_{i,j} \end{aligned} \tag{9}$$

In Equation (9), $Coll_{i,j,1988}$ is our measure of innovative collaboration between inventors located in regions i and j in 1988. $Motorways_{i,j,1983}$ is the sum of the length of motorways in the two regions i and j in 1983 (see Dong et al. (2020) for a rationale of this choice) and $Distance_{i,j}$ represents the travel distance (in kms of highways) between each pair of regions' centroids. We also include directional (i and j) fixed effects, φ_i and φ_j, in order to control for any unobservable omitted regional variables. [40]

[37] In order to assign patents to regions, we still follow the "inventor criterion": every time a patent can be ascribed to several inventors coming from different regions, patent attribution is performed through fractional counts.

[38] See Picci (2010) for a more in-depth explanation and some numerical examples.

[39] It is worth noting that, for a patent application a, if inventors located in region i do not collaborate with inventors located in region j, $Coll_{i,j,a,1988}$ assumes the value of zero.

[40] It is worth noting that region (i and j) fixed effects, φ_i and φ_j, account for most of inward and outward multilateral resistances as well as regional unobservable characteristics that may influence bilateral knowledge flows (Donaubauer et al., 2018). Moreover, following Bacchetta et al. (2012) we include a full set of gravity variables (e.g., bilateral travel distances and cross-regional motorways

Table 5: Transmission Mechanisms: Gravity Model Estimates.

Dependent Variable: $\log Coll_{i,j,1988}$		
	OLS	IV
	(1)	(2)
$\log RomanRoads_{i,j}$	0.0271**	
	(0.0134)	
$\log DistanceRR_{i,j}$	-0.187***	
	(0.0180)	
$\log Coll_{i,j,1983}$	0.451***	0.450***
	(0.0501)	(0.0503)
$\log Motorways_{i,j,1983}$		0.123*
		(0.0674)
$\log Distance_{i,j}$		-0.192***
		(0.0187)
Region i FE	YES	YES
Region j FE	YES	YES
Observations	7,832	7,832
R-squared	0.325	0.318
F-statistic		42.87

Notes: Reduced Form (OLS) and Two-Stage Least Squares estimates. $Coll_{i,j,1988}$ is the aggregate weighted forward citation measure of collaboration between inventors of regions i and j. $RomanRoads_{i,j}$ is the sum of the length of major Roman roads in the two regions i and j, while $Motorways_{i,j,1983}$ is the same measure calculated using motorways in 1983. $Distance_{i,j}$ and $DistanceRR_{i,j}$ are distances (in kms) between each pair of region centroids, calculated respectively on the basis of modern motorways and major Roman roads. We add one to the collaboration and roads measures before to taking the log to preserve zero value observations. F-statistics in column (2) is the first stage Kleinbergen-Paap statistic. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

As mentioned above, there might be simultaneity between the intensity of collaboration among pairs of regions, regional technological evolution and transport infrastructure endowment. Consequently, in Equation (9) there might be correlation between unobservables, $\epsilon_{i,j}$, and the regions-pairs' endowment of highways. In order to address such issue, we rely on the instrumental variable estimation approach. In particular, we use the sum of the length of major Roman roads in the two regions i and j ($RomanRoads_{i,j}$) as instrument for the 1983 sum of the length of motorways in each regions pairs. Moreover, $DistanceRR_{i,j}$ represents the travel distance (in

endowment) in order to proxy for bilateral travel costs. We also include a five-year lagged dependent variable, $Coll_{i,j,1983}$, in order to account for most of the time invariant unobserved heterogeneity at the $i-j$ pair level. Finally, $\epsilon_{i,j}$ is an error term.

kms of Roman roads) between each pair of region centroids and it is used as instrument for 1983 travel distance.

Table 5 presents results from the cross-sectional structural gravity model estimation. First we estimate the reduced form of Equation (9), i.e. an empirical specification where the measure of collaboration between pairs of regions is let to depend on above-mentioned instruments, including directional (i and j) fixed effects. Column (1) of Table 5 shows reduced form OLS estimates. Results suggest that the distance between pairs of regions deters knowledge flows; however, *ceteris paribus*, a denser Roman roads network among them positively affects cross-regional collaborations. Finally, column (2) of Table 5 presents IV estimates and shows that the coefficient of $logMotorways_{i,j,1983}$ is positive and statistically significant, with a magnitude of about 0.123. Consistently with reduced-form estimates, this result suggests that a denser stock of motorways in each pair of regions positively affect innovative collaborations, thus favouring knowledge flows and the circulation of ideas across regions.[41]

5.4 Heterogeneous Effects

We extend our analysis in order to explore possible heterogeneous effects of highways endowment on innovation. We first investigate potential heterogeneity associated to differentials in inventors density across regions. Indeed, geographic proximity favors the development of knowledge flows, learning processes and relations between inventors, which in turn might affect innovative performance. In this context, the motorways system may represent an important tool in favoring the creation of networks between scientists and organisations.

In order to capture the degree of inventors dispersion over the regional area, we construct a measure of inventor density, $(\frac{Inventors_{i,1983}}{Surface_i})$, and we split the sample in High-Density or Low-Density regions, depending on whether they are above or below the mean value of our density measure. We argue that the provision of highways would benefits relatively more those regions where interactions between inventors or researchers require large travelling distances.

In Table 6 we report results obtained by estimating the baseline model after splitting the sample according to inventors density (low or high) and using two different innovation measures: in columns (1) and (2) the dependent variable is our measure of weighted forward citation count of patents, while in columns (3) and (4) we use the unweighted count of patents. Results are consistent with those obtained by Agrawal et al. (2017) and show a null impact of motorways in High-Density regions, while for Low-Density ones a 10% increase in motorways endowment leads to an increase in patent fractional count that ranges from 1.2% (column (3)) to 3% (column (1)).

[41] Results obtained in Table 5 are also confirmed using another specification of the model where we use the geodetic distance, considered exogenous, instead of the travel distance.

Table 6: The Heterogeneous Impact of Motorways on Innovation (High/Low Density Regions): IV Estimates.

Dependent variable:	*Weighted* $logInnov_{i,1988}$		*Unweighted* $logInnov_{i,1988}$	
	Low-Density	High-Density	Low-Density	High-Density
	(1)	(2)	(3)	(4)
$logMotorways_{i,1983}$	0.307***	0.0774	0.117**	-0.0340
	(0.105)	(0.163)	(0.0584)	(0.122)
Lagged Dep. Var.	YES	YES	YES	YES
Geography	YES	YES	YES	YES
Inventors	YES	YES	YES	YES
Observations	71	18	71	18
R-squared	0.590	0.852	0.686	0.901
F-statistic	29.19	4.657	27.48	4.860

Notes: Two-Stage Least Squares estimates. Lagged dependent variable i always included. $\log Innov_{i,1988}$ in columns (1) and (2) the dependent variable is the weighted forward citation count of patents in region i in 1988, while in columns (3) and (4) it is unweighted count of patents in region i in 1988. Low (high) density NUTS-3 regions are characterized by inventor density $(\frac{Inventors_{i,1983}}{Surface_i})$ below (above) the sample mean. Inventor controls include the (log) number of inventors (per capita) in field f in each NUTS-3 region in 1983 and the total (log) number of inventors in each NUTS-3 region in 1983. Geography controls include surface, terrain ruggedness and elevation. We add one to all patent, inventor and motorways before to taking the log in order to preserve zero value observations. F is the First stage Kleinbergen-Paap Statistic. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

We further investigate whether roads differently affect innovation in industries characterized by a faster technological turnover. Since positive externalities associated to new knowledge need time to spread across regions (Caballero and Jaffe, 1993), motorways should play a greater role for knowledge diffusion in industries characterized by a faster process of *"creative destruction"*. In these industries the very high speed of technological turnover requires faster knowledge flows that might be favored by a more developed road network.

In order to identify industries characterized by a relatively higher technological turnover, we rely on Hall et al. (2001) that measure the obsolescence of knowledge by technology field. Following their classification, we consider Electrical Engineering as an "High-Tech Turnover" industry, while Instruments, Chemistry, Mechanical Engineering and residual categories as "Low-Tech Turnover" industries. We then split the sample according to this classification and we estimate Equation (2) for each sub-sample. Results obtained from IV estimation are presented in Table 7.

Findings obtained estimating the baseline model with two alternative measures of regional innovation performance, i.e the weighted forward citation patent count (columns (1) and (2)) and the unweighted innovation measure (columns (3) and (4)), are not consistent across different specification of the model, thus suggesting that the

Table 7: The Heterogeneous Impact of Motorways on Innovation (Speed of Technological Turnover): IV Estimates.

Dependent variable:	*Weighted* $logInnov_{i,f,1988}$		*Unweighted* $logInnov_{i,f,1988}$	
	High Tech T.	Low Tech T.	High Tech T.	Low Tech T.
	(1)	(2)	(3)	(4)
$logMotorways_{i,1983}$	0.150	0.201***	0.104	0.0161
	(0.132)	(0.0761)	(0.0674)	(0.0407)
Lagged Dep. Var.	YES	YES	YES	YES
Field FE	-	YES	-	YES
Geography	YES	YES	YES	YES
Inventors	YES	YES	YES	YES
Observations	89	356	89	356
R-squared	0.359	0.567	0.532	0.711
F-statistic	22.56	20.65	13.18	17.64

Notes: Two-Stage Least Squares estimates. Lagged dependent variable always included. $\log Innov_{i,f,1988}$ in columns (1) and (2) is the weighted forward citation count of patents in region i in technological field f in 1988, while in columns (3) and (4) is the unweighted count of patents in region i in technological field f in 1988. High Tech T. (turnover) patent field are Electrical Engineering and Communication technology, while low Tech T. ones include Instruments, Chemistry, Mechanical Engineering and residual technologies. Field FE are then computed only for low velocity subsample. Inventor controls include the (log) number of inventors (per capita) in field f in each NUTS-3 region in 1983 and the total (log) number of inventors in each NUTS-3 region in 1983. Geography controls include surface, terrain ruggedness and elevation. We add one to all patents, inventors and motorways before to taking the log in order to preserve zero value observations. F is the First stage Kleinbergen-Paap Statistic. Robust standard errors clustered at the NUTS-3 region level in parentheses *** p<0.01, ** p<0.05, * p<0.1

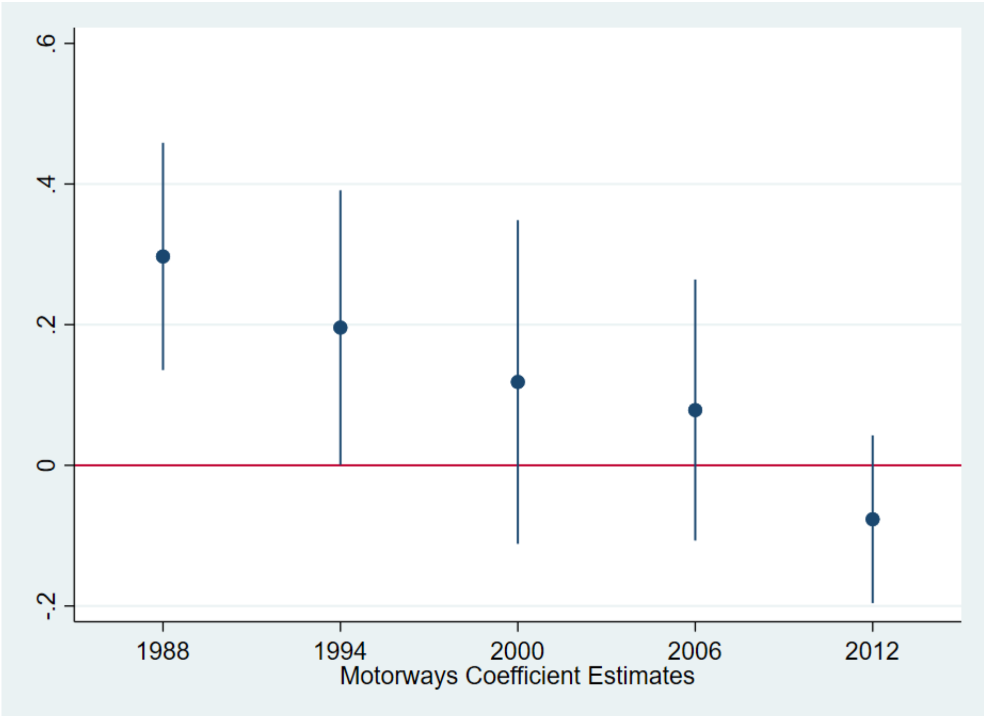
relationship between motorways and innovation may not differ in industries characterised by different speed of technological turnover. These findings are not in line with those obtained by Agrawal et al. (2017) for US MSAs, but this difference might be associated to the different composition of industrial sectors between Italy and the United States in the sample period.

5.5 ICT and Roads

As we explained in the introduction, in this study we focus on the relationship between regional innovative capacity observed in 1988 and the 1983 regional motorways stock, since we believe that these years of observations are not affected by the 1990s Internet revolution. Since the ICT has brought about revolutionary changes in the way people work, communicate, learn, spend time and interact (Jorgenson and Vu, 2016), it is interesting to investigate whether road infrastructures built in 1983 still have an important role in shaping knowledge flows when new communication technologies have been made available.

In order to analyze the long term effect of highways, we explore the impact of the 1983 regional highways endowment on regional innovation performance observed some years later. In particular, we estimate the baseline model for five different years, namely 1988 (as in Table 3), 1994, 2000, 2006 and 2012.

Figure 5: The Impact of 1983 Motorways Endowment on Innovation Over Time.



Notes: IV estimates for the motorways coefficient in 1988, 1994, 2000, 2006 and 2012. The estimated Equation is $\log Innov_{i,t} = \alpha + \beta Motorways_{i,1983} + \gamma Innov_{i,t-5} + \varphi X_i + v_i$. $Innov_{i,t}$ is the forward citation patent fractional count based on inventor address. Lagged innovation is dated 1983, 1989, 1995, 2001 and 2007 respectively. Confidence intervals are reported at 90% level.

Figure 5 shows the time path of our coefficient of interest ($Motorways_{i,1983}$) estimated with the IV method. In 1994 the coefficient slightly decreases in magnitude with respect to its 1988 value, but remains positive and significant. This may reflect the fact that the Internet revolution is still in its early phases, especially in Italy. The transition to digital media appears not yet complete and the role of road infrastructure remains persistent. On the other hand, ten years after the ICT revolution the impact of 1983 motorways endowment on regional innovation in year 2000 is no longer statistically significant. The same result is also evident from 2006 onward.[42]

Although it seems natural to attribute the declining importance of the 1983 motorways network for regional innovation to the ICT revolution, it is important to discuss a possible complementary explanation. Indeed, over the same period congestion in the Italian motorways network increased substantially. By way of example, the number of vehicle kms travelling over Italian motorways over the period 1990-2007

[42]This result differs somewhat from Agrawal et al. (2017), who also find a declining effect over time of the 1983 highways network on subsequent regional innovation, which however remained statistically significant.

raised by more than 60%, while the motorways network increased by just 11% over the whole period 1983-2007. Such an important increase in traffic levels over the motorways network might have significantly attenuated its role as a driving force for regional innovation. Unfortunately, the data on vehicle kms at regional level are not available, so that we cannot control for the different degrees of congestion over time.

6 Conclusions

In this work we assess the impact of motorways endowment on regional innovative performance. We estimate an empirical model that links the Italian NUTS-3 regional innovative performance in 1988 to the stock of regional motorways infrastructure 5 years earlier. We address the endogeneity of the stock of highways with an IV approach, using the ancient Roman roads system dating back to 117 as the excluded instrument. Main results suggest that the 1983 highways network had a positive and significant impact on 1988 regional innovative performance; in particular an increase of 10% in the length of motorways is associated to a 2-3% increase in the number of patents.

Interestingly, we find that this sizeable economic effect tends to decline over time and persists only until the first half of the 1990s, probably because the onset of the internet revolution have made highways less crucial for knowledge diffusion, or because of the increasing congestion which took place on Italian motorways in the most recent period.

Our results are robust to a series of sensitivity checks, such as the consideration of different patent metrics (weighted and unweighted for future citations), of different criteria to attribute the patent to a specific region (e.g. depending on the applicant or the inventor region), the use of standard errors robust to spatial correlation and the consideration of the region or the region-by-technological patent class as unit of analysis. Moreover, according to our estimates of a gravity model, it seems that one possible economic mechanisms driving our findings is related to the improved knowledge diffusion process associated to denser motorways networks, that favor collaborations between inventors living in different regions, as shown by Dong et al. (2020) for the case of high-speed trains in China.

We also show the existence of significant heterogeneous effects, as motorways seem to display a much stronger effect in low density regions, where local interaction requires longer distances to be travelled. Finally, we also provide some evidence that could be useful for the debate over the spatial reorganization versus net effect of road infrastructures hypothesis ((Redding and Turner, 2015)). Results weakly point towards the existence of negative spillovers, which however are not fully robust across empirical specifications.

References

Abramovsky, L. and Simpson, H. (2011), 'Geographic proximity and firm–university innovation linkages: evidence from great britain', *Journal of economic geography* **11**(6), 949–977.

Acemoglu, D. and Akcigit, U. (2012), 'Intellectual property rights policy, competition and innovation', *Journal of the European Economic Association* **10**(1), 1–42.

Aghion, P. and Howitt, P. (1990), A model of growth through creative destruction, Technical report, National Bureau of Economic Research.

Agrawal, A., Galasso, A. and Oettl, A. (2017), 'Roads and innovation', *Review of Economics and Statistics* **99**(3), 417–434.

Anderson, J. E. and Van Wincoop, E. (2003), 'Gravity with gravitas: A solution to the border puzzle', *American economic review* **93**(1), 170–192.

Andrews, I., Stock, J. H. and Sun, L. (2019), 'Weak instruments in instrumental variables regression: Theory and practice', *Annual Review of Economics* **11**, 727–753.

Bacchetta, M., Beverelli, C., Cadot, O., Fugazza, M., Grether, J.-M., Helble, M., Nicita, A. and Piermartini, R. (2012), *A practical guide to trade policy analysis*, Geneva, CH: World Trade Organization.

Baldwin, R. E. and Martin, P. (2004), Agglomeration and regional growth, *in* 'Handbook of regional and urban economics', Vol. 4, Elsevier, pp. 2671–2711.

Baptista, R. (1998), 'Clusters, innovation and growth: a survey of the literature', *The dynamics of industrial clusters: international comparisons in computing and biotechnology* pp. 13–51.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A. and Zhang, Q. (2017), 'Roads, railroads, and decentralization of chinese cities', *Review of Economics and Statistics* **99**(3), 435–448.

Bennett, R. J., Bratton, W. A. and Robson, P. J. A. (2000), 'Business advice: the influence of distance', *Regional Studies* **34**(9), 813–828.

Black, D. and Henderson, V. (1999), 'A theory of urban growth', *Journal of political economy* **107**(2), 252–284.

Bom, P. R. and Ligthart, J. E. (2014), 'What have we learned from three decades of research on the productivity of public capital?', *Journal of economic surveys* **28**(5), 889–916.

Bottasso, A., Conti, M., Ferrari, C. and Tei, A. (2014), 'Ports and regional development: a spatial analysis on a panel of european regions', *Transportation Research Part A: Policy and Practice* **65**, 44–55.

Bronzini, R. and Piselli, P. (2016), 'The impact of r&d subsidies on firm innovation', *Research Policy* **45**(2), 442–457.

Caballero, R. J. and Jaffe, A. B. (1993), 'How high are the giants' shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth', *NBER macroeconomics annual* **8**, 15–74.

Colella, F., Lalive, R., Sakalli, S. O. and Thoenig, M. (2019), 'Inference with arbitrary clustering'.

Corrado, L. and Fingleton, B. (2012), 'Where is the economics in spatial econometrics?', *Journal of Regional Science* **52**(2), 210–239.

Dalgaard, C.-J., Kaarsen, N., Olsson, O. and Selaya, P. (2018), Roman Roads to Prosperity: Persistence and Non-Persistence of Public Goods Provision, CEPR Discussion Papers 12745, C.E.P.R.
URL: *https://ideas.repec.org/p/cpr/ceprdp/12745.html*

De Benedictis, L., Licio, V., Pinna, A. M. et al. (2018), The long-term effects of the historical roman road network: trade costs of italian provinces, Working Paper CRENoS 201801, Centre for North South Economic Research, University of Cagliari and Sassari.

Donaubauer, J., Glas, A., Meyer, B. and Nunnenkamp, P. (2018), 'Disentangling the impact of infrastructure on trade using a new index of infrastructure', *Review of World Economics* **154**(4), 745–784.

Dong, X., Zheng, S. and Kahn, M. E. (2020), 'The role of transportation speed in facilitating high skilled teamwork across cities', *Journal of Urban Economics* **115**, 103212.

Duranton, G. (2015), 'Roads and trade in colombia', *Economics of Transportation* **4**(1-2), 16–36.

Duranton, G., Morrow, P. M. and Turner, M. A. (2014), 'Roads and trade: Evidence from the us', *Review of Economic Studies* **81**(2), 681–724.

Duranton, G. and Turner, M. A. (2011), 'The fundamental law of road congestion: Evidence from us cities', *American Economic Review* **101**(6), 2616–52.

Duranton, G. and Turner, M. A. (2012), 'Urban growth and transportation', *Review of Economic Studies* **79**(4), 1407–1440.

Feldman, M. P. (1999), 'The new economics of innovation, spillovers and agglomeration: Areview of empirical studies', *Economics of innovation and new technology* **8**(1-2), 5–25.

Fernald, J. G. (1999), 'Roads to prosperity? assessing the link between public capital and productivity', *American Economic Review* **89**(3), 619–638.

Ferrari, C., Bottasso, A., Conti, M. and Tei, A. (2019), *Economic Role of Transport Infrastructure: Theory and Models*, Elsevier, Amsterdam.

Fujita, M. and Thisse, J.-F. (1996), 'Economics of agglomeration', *Journal of the Japanese and international economies* **10**(4), 339–378.

Fujita, M. and Thisse, J.-F. (2002), 'Agglomeration and market interaction'.

Fujita, M. and Thisse, J.-F. (2003), 'Does geographical agglomeration foster economic growth? and who gains and loses from it?', *The Japanese Economic Review* **54**(2), 121–145.

Garcia-López, M.-A. (2012), 'Urban spatial structure, suburbanization and transportation in barcelona', *Journal of Urban Economics* **72**(2-3), 176–190.

Garcia-López, M.-À. (2019), 'All roads lead to rome… and to sprawl? evidence from european cities', *Regional Science and Urban Economics* **79**, 103467.

Garcia-López, M.-À., Holl, A. and Viladecans-Marsal, E. (2015), 'Suburbanization and highways in spain when the romans and the bourbons still shape its cities', *Journal of Urban Economics* **85**, 52–67.

Ghani, E., Goswami, A. G. and Kerr, W. R. (2016), 'Highway to success: The impact of the golden quadrilateral project for the location and performance of indian manufacturing', *The Economic Journal* **126**(591), 317–357.

Guillain, R. and Huriot, J.-M. (2001), 'The local dimension of information spillovers: a critical review of empirical evidence in the case of innovation*.(dialogue)', *Canadian Journal of Regional Science* **24**(2), 313.

Hall, B. H., Jaffe, A. B. and Trajtenberg, M. (2001), The nber patent citation data file: Lessons, insights and methodological tools, Technical report, National Bureau of Economic Research.

Hall, B. H., Jaffe, A. and Trajtenberg, M. (2005), 'Market value and patent citations', *RAND Journal of economics* pp. 16–38.

Harhoff, D., Scherer, F. M. and Vopel, K. (2003), 'Citations, family size, opposition and the value of patent rights', *Research policy* **32**(8), 1343–1363.

Hervas-Oliver, J.-L. and Albors-Garrigos, J. (2008), 'The role of the firm's internal and relational capabilities in clusters: when distance and embeddedness are not enough to explain innovation', *Journal of Economic Geography* **9**(2), 263–283.

Holl, A. (2016), 'Highways and productivity in manufacturing firms', *Journal of Urban Economics* **93**, 131–151.

Holl, A. and Mariotti, I. (2018), 'Highways and firm performance in the logistics industry', *Journal of Transport Geography* **72**, 139–150.

Inoue, H., Nakajima, K. and Saito, Y. U. (2017), 'The impact of the opening of high-speed rail on innovation', *Research Institute of Economy. Trade and Industry (RIETI)* .

Izushi, H. (2003), 'Impact of the length of relationships upon the use of research institutes by smes', *Research policy* **32**(5), 771–788.

Jaffe, A. B., Trajtenberg, M. and Henderson, R. (1993), 'Geographic localization of knowledge spillovers as evidenced by patent citations', *the Quarterly journal of Economics* **108**(3), 577–598.

Jorgenson, D. W. and Vu, K. M. (2016), 'The ict revolution, world economic growth, and policy issues', *Telecommunications Policy* **40**(5), 383–397.

Kelly, M. (2020), 'Direct standard errors for regressions with spatially autocorrelated residuals'.

Licio, V. (2020), 'When history leaves a mark: a new measure of roman roads', *Italian Economic Journal* pp. 1–35.

Lin, Y. (2017), 'Travel costs and urban specialization patterns: Evidence from china's high speed railway system', *Journal of Urban Economics* **98**, 98–123.

Love, J. H. and Roper, S. (2001), 'Outsourcing in the innovation process: locational and strategic determinants', *Papers in Regional Science* **80**(3), 317–336.

Martincus, C. V., Carballo, J. and Cusolito, A. (2017), 'Roads, exports and employment: Evidence from a developing country', *Journal of Development Economics* **125**, 21–39.

McCormick, M., Huang, G., Zambotti, G. and Lavash, J. (2013), 'Roman road network (version 2008)', *DARMC Scholarly Data Series, Data Contribution Series* **5**.

Melo, P. C., Graham, D. J. and Brage-Ardao, R. (2013), 'The productivity of transport infrastructure investment: A meta-analysis of empirical evidence', *Regional Science and Urban Economics* **43**(5), 695–706.

Minerva, G. A. and Ottaviano, G. I. (2009), 'Endogenous growth theories: Agglomeration benefits and transportation costs', *Handbook of Regional Growth and Development Theories. Edward Elgar, Cheltenham* .

Moretti, E. and Wilson, D. J. (2014), 'State incentives for innovation, star scientists and jobs: Evidence from biotech', *Journal of Urban Economics* **79**, 20–38.

Nunn, N. and Puga, D. (2012), 'Ruggedness: The blessing of bad geography in africa', *Review of Economics and Statistics* **94**(1), 20–36.

Olea, J. L. M. and Pflueger, C. (2013), 'A robust test for weak instruments', *Journal of Business & Economic Statistics* **31**(3), 358–369.
URL: *https://doi.org/10.1080/00401706.2013.806694*

Pakes, A. and Griliches, Z. (1980), 'Patents and r&d at the firm level: A first report', *Economics letters* **5**(4), 377–381.

Percoco, M. (2015), 'Highways, local economic structure and urban development', *Journal of Economic Geography* **16**(5), 1035–1054.

Picci, L. (2010), 'The internationalization of inventive activity: A gravity model using patent data', *Research Policy* **39**(8), 1070 – 1081.
URL: *http://www.sciencedirect.com/science/article/pii/S0048733310001319*

Redding, S. J. and Turner, M. A. (2015), Transportation costs and the spatial organization of economic activity, *in* 'Handbook of regional and urban economics', Vol. 5, Elsevier, pp. 1339–1398.

Roca, J. D. L. and Puga, D. (2017), 'Learning by working in big cities', *The Review of Economic Studies* **84**(1), 106–142.

Romer, P. M. (1990), 'Endogenous technological change', *Journal of political Economy* **98**(5, Part 2), S71–S102.

Smith, K. (2005), 'Measuring innovation'.

Stock, J. H. and Yogo, M. (2005), 'Testing for weak instruments in linear iv regression, in dwk andrews and jh stock, eds., identification and inference for econometric models: Essays in honor of thomas j. rothenberg. cambridge: Cambridge university press'.

Talbert, R. J. (2000), *Barrington Atlas of the Greek and Roman World: Map-by-map Directory*, Vol. 1, Princeton University Press, Princeton.

Trajtenberg, M. (1990), 'A penny for your quotes: patent citations and the value of innovations', *The Rand journal of economics* pp. 172–187.

Wang, X., Xie, Z., Zhang, X. and Huang, Y. (2018), 'Roads to innovation: Firm-level evidence from people's republic of china (prc)', *China Economic Review* **49**, 154–170.

Xu, H. and Nakajima, K. (2017), 'Highways and industrial development in the peripheral regions of china', *Papers in Regional Science* **96**(2), 325–356.

Yamasaki, J. (2017), 'Railroads, technology adoption, and modern economic development: Evidence from japan'.

Appendix A

In this section we strictly follow Agrawal et al. (2017) in explaining the theoretical innovation model underlying Equations 1 and 3.
The deterministic innovation level in a region, K_t^*, is related to the level of motorways, R_t, trough the relation $K_t^* = A * R_t^{\alpha}$.
Authors argue that the adjustment rate depends on how far a region is from the deterministic level of innovation. The innovation adjustment rate is defined as $K_{t+j} = {K_t^*}^{1-\gamma} K_t^{\gamma}$, with $0 < \gamma < 1$. Authors show that the level of innovation at time $t+j$ is equal to $K_{t+j} = BR_t^{\beta} K_t^{\gamma}$, where $\beta = \alpha(1-\gamma)$ and $B = A^{1-\gamma}$. Taking the log of this latter Equation, Equation (1) is obtained. The parameter of interest, β, describes the rate at which knowledge creation responds to motorways endowment.

Appendix B

Following Colella et al. (2019), we take into account the possibility that standard errors might be spatially correlated. Authors propose a variance-covariance matrix estimator that allows us to obtain cluster-robust inference in a TSLS setting with arbitrary dependence across observations. Their approach considers a circle around each observational unit that spatially bounds distance dependence, allowing for different decay processes. In Table B1 we provide results obtained by estimating our baseline model with the Colella et al. (2019) estimator; in particular, we introduce the spatial correction with a threshold of 75 km. This means that the error of each region is assumed to be correlated with errors of other regions located within a radius of 75 km. Following this hypothesis, there are, on average, five regions in each spatial cluster. We conduct the analysis both using a uniform decay kernel and Bartlett-type kernels. In the first case, the matrix used for the computation of the variance-covariance matrix is binary; in the second case, we allow for weights in the matrix to linearly decrease as the distance increases, with values very close to one for nearby regions and almost zero for regions close to the distance cutoff. Results in Table B1 are robust to spatial correlation and confirm the positive and significant impact of 1983 motorways on 1988 innovative capacity of Italian regions. Moreover, while not significant in column (3), we detect negative and significant spatial spillovers in column (4).

Table B1: The Impact of Motorways on Innovation: IV Estimates.

Dependent Variable: $\log Innov_{i,1988}$				
	(1)	(2)	(3)	(4)
$logMotorways_{i,1983}$	0.297***	0.297***	0.405***	0.405***
	(0.108)	(0.099)	(0.134)	(0.132)
$SpatialMotorways_{i,1983}$			-2.694	-2.694*
			(1.746)	(1.482)
Lagged Dep. Var.	YES	YES	YES	YES
Geography	YES	YES	YES	YES
Inventors	YES	YES	YES	YES
Kernel	Uniform	Bartlett	Uniform	Bartlett
Observations	89	89	89	89
R-squared	0.703	0.703	0.664	0.664
F-statistic	14.26	20.66	6.22	6.31

Notes: Two-Stage Least Squares estimates. Lagged dependent variable always included. $\log Innov_{i,1988}$ refers to the weighted forward citation count of patents in region i in 1988. Inventor controls include the (log) number of inventors (per capita) in each NUTS-3 region in 1983. Geography controls include surface, terrain ruggedness and elevation. We add one to all patent, inventor and motorways before to taking the log in order to preserve zero value observations. Columns (1) and (3) are based on a uniform decay kernel, while columns (2) and (4) on a Bartlett kernel. F is the First stage Kleinbergen-Paap Statistic. Standard errors corrected for cluster spatial correlation in parentheses *** p<0.01, ** p<0.05, * p<0.1

Appendix C

In this section, we asses the validity of the results obtained in Table 2 and Table 3 (main text) by performing a set of robustness tests. In particular, we replicate our baseline estimates by changing the dependent variable according to two different criteria.

First, we focus on the possibility that ascribing patents to each region according to inventor address might lead to misleading results. Indeed, the presence of a specific motorway may increase the patent activity in a specific region i even if some inventors might live outside it, so that the real impact of the motorway is to favor inventors commuting. In order to analyze this issue, we built our innovation measure based on applicant address (instead of inventor address). In fact, the inventor plays a central role in conceiving the invention and contributes to its industrialization process; he might have entered into a contract with the applicant (usually the organization for which he-she works) so that the applicant assumes the rights to deal with the invention. In this case, the use of patents fractional count computed on the basis of applicant address reduces the problem of inventors working in regions other than the region of residence. Estimates are shown in Table C1 and confirm the positive relationship of innovative capacity with the provision of motorways infrastructure at regional level. Indeed, the estimated coefficient of $logMotorways_{i,1983}$ is quite similar to those reported in Table 3.

Table C1: The Impact of Motorways on Innovation Activity (Forward Citation Weighted Patent Fractional Count based on Applicant Address): OLS and IV Estimates.

	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Dependent Variable: *Applicant* $\log Innov_{i,1988}$				
$logMotorways_{i,1983}$	0.418***	0.329*	0.429***	0.342*
	(0.103)	(0.182)	(0.102)	(0.179)
$SpatialMotorways_{i,1983}$			-1.013	-1.186
			(0.624)	(0.748)
Geography	YES	YES	YES	YES
Applicants	YES	YES	YES	YES
Observations	89	89	89	89
R-squared	0.549	0.545	0.556	0.552
F-statistic		18.74		8.773

Notes: *Applicant* $\log Innov_{i,1988}$ is the forward citation weighted count of patents based on applicant address. All specifications include the lagged dependent variable. Applicant controls include the (log) number of applicants (per capita) in each NUTS-3 region in 1983. Geography controls include surface, terrain ruggedness and elevation. We add one to all patent, applicants and motorways before to taking the log in order to preserve zero value observations. F is the First stage Kleinbergen-Paap Statistic. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Second, we replicate our baseline analysis using an unweighted patent fractional count as dependent variable in order to account for the raw quantitative measure of innovative activity at regional level. Comfortingly, our previous results are confirmed as shown in Table C2, even if in this specification the coefficient of $\log Motorways_{i,1983}$ shows values that are smaller in magnitude with respect to those shown in Table 2 and Table 3 in the main text. Moreover, the coefficient of $SpatialMotorways_{i,1983}$ is no longer statistically significant.

Table C2: The Impact of Motorways on Innovation (Unweighted Patent Fractional Count): OLS and IV Estimates.

	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Dependent variable: *Unweighted* $\log Innov_{i,1988}$				
$logMotorways_{i,1983}$	0.0967**	0.102*	0.0928**	0.130**
	(0.0396)	(0.0582)	(0.0397)	(0.0619)
$SpatialMotorways_{i,1983}$			0.396	-0.742
			(0.413)	(0.497)
Geography	YES	YES	YES	YES
Inventors	YES	YES	YES	YES
Observations	89	89	89	89
R-squared	0.783	0.782	0.784	0.771
F-statistic		24.28		8.549

Notes: *Unweighted* $\log Innov_{i,1988}$ is the unweighted count of patents in region i in 1988. Lagged dependent variable is always included. Inventor controls include the (log) number of applicants (per capita) in each NUTS-3 region in 1983. Geography controls include surface, terrain ruggedness and elevation. We add one to all patent, applicants and motorways before to taking the log in order to preserve zero value observations. F is the First stage Kleinbergen-Paap Statistic. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1