The definitive, peer-reviewed and edited version of this article is published and can be cited as

Duran-Fernandez, R. and G. Santos (2014), 'Road Infrastructure Spillovers on the Manufacturing Sector in Mexico', *Research in Transportation Economics*, Vol. 46, pp. 17-29. DOI: 10.1016/j.retrec.2014.09.002

Road Infrastructure Spillovers on the Manufacturing Sector in Mexico

Roberto Duran-Fernandez[1*] and Georgina Santos[2]

[1]Transport Studies Unit, University of Oxford, UK
Tel: +52 (55) 5249 5060
E-mail address: r.duran.fernandez@gmail.com
[*]Corresponding author

[2]School of Planning and Geography, Cardiff University, UK, and Transport Studies Unit, University of Oxford, UK
Tel: +44 (0) 29 208 74462
E-mail address: SantosG@Cardiff.ac.uk

ABSTRACT

This paper investigates the effects of road infrastructure on industrial activity in Mexico from a quantitative perspective. It addresses three main issues. First, it investigates the existence of a relationship between road infrastructure and the industrial average product of labour. Second, it studies the determinants of such a relationship and quantifies the magnitude of the impact of road infrastructure on the average product of labour. Third, it analyses the spatial effects of road infrastructure in Mexico. The findings can be summarised as follows. First, road infrastructure has a positive and significant effect on the industrial average product of labour. However, not all the elements (roads or groups of roads) of the road system have the same effect. Second, we find that the actual magnitude of the effects of accessibility on the average product of labour depends on the physical attributes of the roads, as well as the peculiarities of the road network. Finally, we show that the regional gaps in the average product of labour across the country can be partially attributed to differences in infrastructure endowments. From a methodological perspective, the main contribution of this paper is the development of a comprehensive methodology for the analysis of some of the benefits of road infrastructure in Mexico. This methodology can be applied as a tool in the planning, implementation, and evaluation of actual infrastructure policy in this country.

Key words: road infrastructure, industrial activity, average product of labour, road infrastructure spatial effects, accessibility, attraction-accessibility measures, regional production function, Mexican manufacturing sector

JEL codes: R11, R12, R3, R4, R42, R53, N66

1 INTRODUCTION

This paper investigates the impact of road infrastructure on industrial activity in Mexico. It addresses three issues. First, it investigates the existence of a relationship between road infrastructure and the industrial production, second, it studies the determinants of such a relationship, and third it analyses its causality nature. The paper estimates a regional production function that includes, as determinants, some metrics of the value of road infrastructure. These metrics are based on a set of attraction-accessibility measures (AMs) for Mexico. The analysis uses data of the manufacturing sector from the National Economic Censuses (NECs, Institute of Statistics, Geography, and Information, INEGI 2004).

The results of this paper can be summarised in the following points: First, road infrastructure has a positive and significant effect on industrial production in Mexico. However, not all the elements (roads or groups of roads) of the road system have the same effect. Second, the actual magnitude of the effects of accessibility on industrial product depends on the physical attributes of the roads, as well as the peculiarities of the road network. Finally, we show that the regional gaps in per worker industrial product across the country can be partially attributed to differences in infrastructure endowments.

The paper is organised as follows: Section 2 presents a brief literature review, Section 3 introduces the empirical approach taken in this study, Section 4 describes the data, Section 5 presents the empirical analysis, Section 6 discusses the results, Section 7 focuses on a causality analysis, Section 8 estimates the magnitude and elasticities of the effects that are identified, and Section 9 concludes.

2 LITERATURE REVIEW

Aschauer (1989) started a debate in the literature about the effects of public infrastructure on economic activity. The empirical strategy that it follows consists of estimating an aggregate production function of the US economy that includes public capital as an argument. This methodology is known in the literature as the 'production function approach'. Aschauer (1989) finds a positive and significant effect of public infrastructure on production, raising several questions that initiated a boom in the study of the effects of public infrastructure on the economy. Although, from a theoretical point of view, few authors would doubt that infrastructure has impacts on production, the magnitudes of the estimated effect in that paper have been questioned.

Serious attempts to review Aschauer's work focus on the estimation of aggregate production functions at state level in the US, such as Evans *et. al.* (1994), Garcia-Mila (1996), Hotz-Eakin (1994), and Kelejian *et. al.* (1997). These papers follow the same empirical strategy using aggregate production data of the 48 contiguous states in the US. These studies tend to generate smaller estimates, and thus solve some of the econometric problems found in Aschauer (1989). However, they fail to estimate robust results in the sense that they are significantly sensitive to the econometric specification and estimation technique.

The analysis of local level data and the incorporation of explicit spillover effects across regions give an important insight on the cause of the divergence between the results presented above. Boarnet (1996) presents an empirical extension of a two-city location model to analyse the spillover effects of transport infrastructure at local level. Using disaggregate information at county level for California, he finds that transport infrastructure has positive and significant effects on output, however he also identifies negative spillovers making the direction of the aggregate effect ambiguous.

The linkage between economic activity and transport infrastructure has not been fully studied for the Mexican case and the literature only presents a limited number of examples. The literature for Mexico has followed the international mainstream, focusing on the estimation of aggregate production-functions and industry specific cost-functions. Castañeda *et. al.* (2000) use data of the manufacturing sector to estimate a production function at national level and find a positive and significant effect of public infrastructure. Ramírez (2004) estimates an aggregate production function using a VAR model and finds similar results. Feldstein *et. al.* (1995) and Shah (1998) on the other hand, estimate a national level cost-function, and find negligible effects of transport infrastructure on the economy. Costa-i-Font *et. al.* (2005), Looney *et. al.* (1980) and Fuentes *et. al.* (2003) find evidence of positive and significant effects of public infrastructure investment using different econometric techniques.

Another methodology that has been used consists of using a neo-classical production function to derive an empirical expression in order to estimate the parameters of an out-of-equilibrium production function. This is useful for analysing the dynamics of the effect of public capital, and determining the convergence rate of the economy, conditional on infrastructure stock. Chiquiar (2003), Fuentes *et. al.* (2003), and Mallick *et. al.* (1994) follow this approach and find significant and positive effects of infrastructure on economic activity.

The main problem that these studies face is the lack of primary data on public infrastructure stock. They estimate the value of this variable using indicators, which most of the time are not reliable. The problem is particularly important for state or municipal level analysis since available data does not allow the calculation of disaggregate figures for infrastructure stock. For solving this problem, authors have used physical measures as proxies for the value of public infrastructure such as length of the road and rail systems (Fuentes 2003).

An outstanding methodology to solve the data availability problem is proposed by Deichmann *et. al.* (2004). In this paper, the authors propose the use of an accessibility index at municipal level as a simple measure of market integration in order to capture the effects of the transport network. This approach has also been used to analyse data from Brazil, India and Indonesia (Deichmann *et. al.* 2005, Lall 2001, Lall *et. al.* 2004, and Lall *et. al.* 2005). Deichmann *et. al.* (2004) find evidence that accessibility induced by transport network infrastructure has a positive and significant effect on productivity using individual firm data from an industrial survey in Mexico.

3 EMPIRICAL APPROACH

The empirical analysis estimates a production function using cross-section regional data from the manufacturing sector in Mexico at regional level, as shown in Equation 1:

Equation 1 $$y_{i,s} = f(z_{i,s}, \varphi_s, x_s) + u_s + \varepsilon_{i,s}$$

In Equation 1, $y_{i,s}$ is the per worker production[1] of a sub-sector *i* in the manufacturing sector, in location *s;* $z_{i,s}$ is a vector of private factors of production; φ_s is a vector of socioeconomic characteristics in region *s;* x_s is the value of transport infrastructure; and u_s and $\varepsilon_{i,s}$ are independent and identically distributed random variables. Vector $z_{i,s}$ is expressed in terms of productive inputs per worker. The error terms u_s and $\varepsilon_{i,s}$ can be interpreted as a regional disturbances of per worker production and as a sector-specific disturbance respectively. They are assumed to be independent across regions and industrial activities. Mean independence

[1] For the purpose of this exercise, the per worker production is estimated as the ratio of Gross Value Added (GVA) to the total number of employees.

with respect to the arguments of the production function is also assumed. The model considers private stock of capital, labour, and intermediate consumption as private factors of production ($z_{i,s}$). Vectors φ_s and x_s include human capital and infrastructure metrics respectively. These variables will capture any effect on per worker production that has not been internalised by private input factors.

3.1 Accessibility and the Value of the Road Network

We use three attraction-accessibility measures (AM) as metrics for the value of road infrastructure. These metrics are characterised by a domestic, an international, and a regional accessibility index. The indexes measure the market potential of a location at different geographic scales and depend on the characteristics of the road infrastructure such as its structure and level of service. Equation 2 introduces the functional form of the accessibility indexes used in this paper.

Equation 2

$$A_i = \sum_j \frac{p_j}{d_{i,j}^{\theta}}$$

In this equation, p_j is the population in j, $d_{i,j}$ is the travel time between locations i and j, and θ is the elasticity of the market potential with respect to impedance.[2] In the domestic index i and j are any two regions in Mexico. In the international index i and j are regions in Mexico and the USA respectively. Finally, the regional index is the average accessibility for all the urban settlements in a particular region.[3]

The analysis also includes a *network curvature index* as a control for the quality of the roads at local level. This variable is defined as the ratio between the length of the optimal routes

[2] Population data is from the National Population Census (INEGI 2000). Travel times were estimated under the North American Geographic Information System (GIS) Road Network Model (Duran 2014). The value of the elasticity is 0.73 after Duran and Santos (2014).

[3] Mexican regions follow the regionalisation proposed by Bassols-Batalla (1993, 2002). The regions in the USA are defined by the Bureau of Economic Analysis in Johnson and Kort (2004).

between i and j in a region and the linear distance between them. This index is a measurement of road quality in any region under the assumption that higher quality roads will tend to trace direct linear paths. Finally, the regression considers as additional controls the mean distance of a location i to a border port of entry to Central America and the minimum distance to any of the 17 international cargo ports in the country.

'

3.2 Regional Hierarchies and Transport Infrastructure

Tabushi and Thisse (2005) and Boarnet (1998) propose that the overall effects of transport infrastructure are uneven across locations. The empirical analysis explicitly acknowledges this possibility, introducing a hierarchical ranking that allows variations of the estimated effects depending on these hierarchies. This approach assumes that locations of the same hierarchy might exhibit a homogenous economic structure and in principle should respond symmetrically to symmetric variations of technological and transport parameters. In particular, we use the Shannon's Entropy Index (*H*) as a hierarchical score (Dendrinos and Mullally 1985). The index follows the functional form presented in Equation 3 (Shannon 1963):

Equation 3

$$H_i = -\sum_j \pi_{i,j} \ln(\pi_{i,j})$$

where $\pi_{i,j}$ is the ratio of production of the j^{th} industrial activity with respect to the total, and H is non-dimensional and continuous in $\pi_{i,j}$. Intuitively, if each industrial activity has the same weight in the economy (*i.e.* $\pi_{i,j} = n^{-1}\ \forall j$) then H is a monotonic increasing function of n (*i.e.* the number of industries). In this case, the hierarchical score depends exclusively on industrial heterogeneity. For two production centres with the same industrial range, the most even (least specialised) will have a higher score.

H is adjusted to obtain a better approximation for the hierarchical ranking of a region. We assume that the strength of economic interaction among the agents in a functional region decreases with population dispersion. Therefore, we estimate an Adjusted Shannon Entropy Index (H^*) as the product of a regional Herfindahl population and H. Equation 4 presents the specification for the Adjusted Shannon Entropy Index (H^*). In this expression, H_r is the Shannon Entropy Index in region r, and the ratio that follows H_r is the regional Herfindahl population index, where N_k is total population in urban settlement k and $p_{i,k}$ is the urban settlement k in region i.

Equation 4

$$H_i^* = H_r \cdot \frac{\sum_i p_{i,k}^{\ 2} - 1/N_k}{1 - 1/N_k}$$

H^* is estimated using the gross production manufacturing sector data aggregated at regional level. Mexican regions are classified in three groups after this index. Group 1 includes all the regions with a score below the average, Group 2 includes all the regions with a score above the average but excludes the outliers that are two standard deviations above this value, and Group 3 includes the outlier regions with a score two standard deviations above the mean. In general, this classification associates rural regions to Group 1, urban areas and their hinterlands to Group 2, and major metropolitan areas to Group 3.

We expect that road infrastructure will have homogenous effects on regions of similar hierarchy. The empirical approach that we take is to estimate the model presented in Equation 1 separately for each of the groups described above. This approach recognises that the parameters of the aggregated production function for each of the groups might be different, given the difference in their industrial profile.

4 DATA DESCRIPTION

The analysis uses a cross-section database of the manufacturing sector, generated from the National Economic Census (INEGI 2004) in Mexico. This dataset is an exhaustive survey, which presents data for every economic unit in the country, defined for the manufacturing sector as an individual production place (INEGI 2004a, 2004b). Each observation in the original dataset has aggregate data at the municipal level for all the economic activities according to the 6-digit level North American Industrial Classification (NAIC). Almost 65 percent of the observations in the sample contain information of only one or two economic units; however, strict disclosure laws forbid INEGI to reveal the actual number of firms in each cohort. The data is aggregated at regional level, following the regionalisation of the Mexican territory proposed by Bassols-Batalla (1993, 2002).[4]

We use this information to build a database of the manufacturing sector in Mexico at regional level. The database is composed of 11,533 observations, which include information for 291 manufacturing activities in 135 regions. On average, there are 85.4 manufacturing activities per region.

The database includes information on gross production, gross value added (GVA), stock of private capital, cost of intermediate consumption such as energy and raw materials, total number of workers, paid wages, and investment for each of the manufacturing activities in each region. The estimation of the model, as presented in Equation 1, defines per worker production as the ratio of GVA to the total number of employers in the sector. Private capital stock and intermediate consumption are used as private input factors. The estimation of the model includes two additional regional controls: the first is the regional average of schooling years, as a proxy for human capital in the region, and the second is the regional Adjusted Shannon Entropy Index, as a proxy for the effects of agglomeration economies.

The three accessibility indexes used in this paper were calculated in Duran-Fernandez and Santos (2014). The rest of the geographic variables –curvature and proximity measures- were estimated using the North American GIS Road Network Model (Duran-Fernandez 2014). The

[4] A discussion of this methodology, as well as a list of the regions and their geographic location, is presented in Appendix A.

data for Mexico was produced by the INEGI and the USA data was produced by the USA Bureau of Transportation Statistics (BTS).

5 REGRESSION ANALYSIS

The parameters of the model were estimated under a Generalised Least Square (GLS) regression. The model was estimated using four different samples. The first considers all the observations in the dataset. The other three sub-samples take into account the hierarchical groups described in Section 3.2. The first sub-sample (Group 1) comprises 5,603 observations from 86 regions with an estimated Adjusted Shannon Entropy Index below the mean. The second sub-sample (Group 2) includes 4,421 observations from 42 regions with a hierarchical ranking above the mean but below two standard deviations. Finally, the third sub-sample (Group 3) comprises 1,509 observations from seven regions with a hierarchical score that is two standard deviations above the mean value of H^*.[5]

5.1 General Results

Table 1 shows the results of the estimation of a version of the model that just considers private factors of production, (capital per worker and intermediate consumption per worker), excluding the regional controls of the model. The estimated parameters are highly significant for both the complete dataset and each of the hierarchical groups. In Table 2 we present the results of the estimation of the model that includes the regional controls. The coefficients of the private factors of production are, again, highly significant. A number of interesting features can be derived from the comparison of these two versions of the model.

First, the value of the coefficients of the private factors and their overall significance does not vary significantly when the regional variables are included. In equilibrium, the demand for private factors per worker of a competitive firm is a fixed ratio, which depends on the factors' relative prices and technology. The estimation shows that the inclusion of the regional controls

[5] Due to the fact that Group 3 only comprises seven regions, the model cannot include all the regional controls. Thus, the model only considers schooling, domestic, and international accessibility.

is not highly correlated with this ratio, suggesting that these variables have no effect on the relative prices of productive inputs and technology.

Second, the inclusion of the regional variables has an effect on the explanatory power of the model. However, they show divergences depending on the hierarchical group. The model shows that at national level, most variations in the per worker output among regions are explained by differences in the use of private factors. This result is driven by the results of the low hierarchy regions in Group 1. For this group, the variations in the use of private factors account for 82 percent of the interregional explanatory power of the model, while the regional variables only account for 2.3 percent. The results, however, are somewhat different for the other hierarchical groups. For the middle-hierarchy and outlier regions in Groups 2 and 3, the regional controls raise the interregional explanatory power of the model by 11.4 and 31.4 percent, respectively.

Third, the estimated variance of regional per worker output is considerably lower in the model with regional controls. It drops by 21 percent at national level and by 12 percent and 51 percent for Groups 1 and 2, respectively. These results illustrate that part of the unexplained variation in the output per worker at regional level can be attributed to the regional variables rather than to random variations.[6]

Finally, the regional controls have an important effect on the spatial distribution of the estimated regional residuals of the output per worker (u_s in Equation 1). In the model with no regional controls, this variable exhibits a positive and highly significant spatial autocorrelation.[7] The inclusion of geographic controls changes this picture: the original pattern described by the residuals disappears. Furthermore, under the second model the residuals

[6] For Group 3 the regional controls explain almost 100 percent of domestic variations in the output per worker. As shown in Table 2, the variance of the model with regional controls is zero. This can be attributed to the relatively large number of regional controls included in the model in comparison to the number of regions in that group.

[7] Moran's I for the regional residuals $\hat{u}$ is equal to +0.0474 (z=8.22). The spatial weights for this estimation use actual travel times between any two regions in the country.

describe a random pattern captured by a non-significant Moran's *I*. [8] This result indicates that the regional variables can explain the systematic clustering pattern described by per worker output not associated to the production factors.

Table 1 GLS Regression with No Geographic Controls

(Standard Errors in Parenthesis)

GLS	All Regions		Hierarchical Groups					
			Group 1		Group 2		Group 3	
Private Factors								
ln(K/L)	0.1368	*	0.1350	*	0.1468	*	0.1153	*
	(0.0062)		*(0.0086)*		*(0.0105)*		*(0.0177)*	
ln(IC/L)	0.5900	*	0.6055	*	0.5791	*	0.5120	*
	(0.0069)		*(0.0098)*		*(0.1160)*		*(0.0205)*	
Constant	0.8372	*	0.7187	*	0.9075	*	1.6444	*
	(0.0283)		*(0.0375)*		*(0.0475)*		*(0.0916)*	
Variance								
σ_{υ}^2	0.165		0.162		0.147		0.110	
σ_{ε}^2	0.764		0.777		0.769		0.692	
ρ□ [/1]	0.045		0.042		0.035		0.025	
N	11,533		5,603		4,421		1,509	
Groups	135		86		42		7	
R^2 (Overall)	0.6262		0.6175		0.5910		0.4954	
Wald χ (2)	16,118.36	*	8,062.88	*	6,113.43	*	1,553.01	*

[/1] Fraction of variance due to regional effect u_i

Significance level *1% ** 5% *** 10%

Note: K, stock of private capital; IC, intermediate consumption; L, labour

Source: Own calculation

[8] Moran's *I* for the regional residuals $\hat{u}$ is equal to –0.011 (z=-0.5423).

Table 2 GLS Regression with Geographical Controls
(Standard Errors in Parenthesis)

GLS	All Regions		Hierarchical Groups					
			Group 1		Group 2		Group 3	
Private Factors								
ln(K/L)	0.1370	*	0.1362	*	0.1437	*	0.1114	*
	(0.0062)		*(0.0086)*		*(0.0105)*		*(0.0177)*	
ln(IC/L)	0.5828	*	0.5987	*	0.5805	*	0.5213	*
	(0.0070)		*(0.0098)*		*(0.0117)*		*(0.0205)*	
Accessibility and road network structure								
Domestic	-0.0012		-0.0016		-0.0006		-0.0030	
	(0.0012)		*(0.0017)*		*(0.0015)*		*(0.0025)*	
International	0.0166	*	0.0116	**	0.0168	*	0.0142	*
	(0.0030)		*(0.0051)*		*(0.0035)*		*(0.0048)*	
Regional	0.0020	**	0.0014		0.0021		**	
	(0.0008)		*(0.0011)*		*(0.0011)*			
Network Curvature	-0.1799	*	-0.2179	*	-0.0568			
	(0.0665)		*(0.0802)*		*(0.1445)*			
Geographic controls								
Port Pacific	0.0007		-0.0050		0.0113			
	(0.0055)		*(0.0072)*		*(0.0084)*			
Port Gulf	0.0105		0.0045		0.0121			
	(0.0069)		*(0.0095)*		*(0.0102)*			
South	-0.0274		-0.0091		-0.0346			
	(0.0167)		*(0.0226)*		*(0.0251)*			
Other Variables								
Schooling	0.0423	**	0.0463		0.0534	***	-0.0404	
	(0.0213)		*(0.0309)*		*(0.0303)*		*(0.0504)*	
H	0.0242		0.0464		0.0082			
	(0.0205)		*(0.0778)*		*(0.0428)*			
Constant	0.2132		0.3748		-0.1002		1.3382	**
	(0.2084)		*(0.2689)*		*(0.3367)*		*(0.6269)*	
Variance								
σ_υ^2	0.131		0.142		0.072		0.000	
σ_ε^2	0.763		0.777		0.767		0.692	
ρ□ [/1]	0.029		0.032		0.009		0.000	
N	11,476		5,603		4,364		1,509	
Groups	133		86		40		7	
R^2 (Overall)	0.6480		0.6310		0.6101		0.5221	
Wald χ (11) [/2]	16,818.21	*	8,264.00	*	6,424.80	*	1,642.20	*

[/1] Fraction of variance due to regional effect u_i
[/2] χ (3) for G3
Significance level *1% ** 5% *** 10%
Note: K, stock of private capital; IC, intermediate consumption; L, labour
Source: Own calculation

5.2 Private Factors of Production

The elasticity of the dependent variable with respect to private capital is not statistically different between the models estimated for Groups 1 and 2[9]. It is higher for the model estimated for Group 1 than for that estimated for Group 3, which in turn is higher than the elasticity estimated for the model restricted to Group 2, both at a significance level of 10 percent.[10] The elasticity of per worker output with respect to intermediate consumption decreases with hierarchy. The elasticity for Group 3 is lower than the elasticities for Groups 1 and 2. These differences are significant at one percent level. In addition, the elasticity for Group 1 is higher than that for Group 2; however, the difference is not highly significant.

This pattern suggests that regions with higher hierarchical ranking operate with technologies that are less labour-intensive. The urban economics literature has proposed that capital abundant economies tend to have higher degrees of urban concentration. Henderson (1988) suggests that capital abundant economies tend to have higher degrees of urban concentration. Because regions with high values for the adjusted industrial diversity in our model have higher urbanisation rates[11], the results are supported by this theoretical observation.

5.3 Accessibility and Geographic Controls

The most important result is that the effects of international and regional accessibility are positive and highly significant for the whole sample, highlighting the importance of physical access to the USA market on the output per worker at national level. Another important result is that the network curvature has a negative and significant effect on production. The effects of domestic accessibility are negligible for the whole sample. The geographic controls that indicate the distance to maritime port facilities as well as the distance to the southern border are also negligible.

---

[9] To test the hypothesis that the coefficient estimated for one group G_i (with i=1,2,3) was not significantly different to the coefficient estimated for another group G_j (with j different to i and j = 1,2,3) we carried out three χ^2 tests, each of them under a model restricted to G_i and G_j respectively.

[10] However, the hypothesis of equal coefficients for G2 and G3 cannot be rejected under the model restricted to Group 3.

[11] The factor of correlation between adjusted industrial diversity and urbanisation rate in our model is 0.58. The urbanisation rate was taken from the National Population Census 2000 and computed as the ratio of urban population to total population.

However, the general pattern found at national level is not replicated when the model is estimated conditional on the hierarchical ranking of the regions. In fact, some of the results of the estimations for each hierarchical group show that the external effects of transport infrastructure are asymmetric.

6 DISCUSSION

6.1 International Accessibility and International Trade Flows

Firms with higher accessibility to the USA market face lower transport costs in comparison with those with lower accessibility to the USA market. This is reflected in the positive spillovers on the output per worker that the model identifies. The results show a positive and significant effect of international accessibility on output per worker for the manufacturing sector. A direct consequence of this effect is that regions with high international accessibility would be expected to exhibit large international trade flows. However, the unavailability of data prevents us from verifying this claim. Having said that, INEGI has information on the *maquiladora* sector, which enables us to analyse indirectly the international trade flows of the manufacturing sector and its relationship with international accessibility.[12]

The *maquiladora* industry is located in the border regions, which exhibit high international accessibility levels. However, there is an important concentration of its activities in central and western Mexico. The industry has no significant presence at all in the south and southeast of the country. This pattern matches with great accuracy the spatial distribution of international accessibility. In fact, the simple correlation between *maquiladora* production and the average level of this variable for the regions in a given state is 0.65.

12 The maquiladora in Mexico is an export orientated manufacturing firm, which operates under a special government programme. In this programme, the industry is entitled to a preferential customs treatment. The programme allows firms to temporarily import intermediate goods, such as materials, machinery, and equipment on a duty-free basis from the USA and Canada. Up to 99 percent of the production is exported to the USA.

The export orientation of the sector and its large weight on the manufacturing sector make the *maquiladora* an ideal instrument to investigate the relationship between manufacturing trade flows and accessibility. The analysis of the sector suggests that firms located in regions with high international accessibility develop strong commercial links with the USA market. This observation confirms indirectly that international accessibility is capturing the positive spillovers originated by the proximity of firms to the international markets.

6.2. Domestic Accessibility, Toll-Roads, and the Internalisation of Spillovers

The positive and significant coefficient of international accessibility can be attributed to a positive impact on costs for firms located in regions with highly articulated international markets. However, the analysis cannot identify any positive and significant spillover effects on the output per worker for domestic accessibility. This is a puzzling result because the theoretical mechanism that explains the positive effect of international accessibility must also apply to the domestic index. We propose two possible explanations. First, the actual domestic market potential is smaller than that estimated by the domestic accessibility index. The second one is that any positive spillover has already been internalised by the cost structure of the firms and therefore it is not reflected in the model results.

The first option is unlikely, at least for the high hierarchy regions. The domestic accessibility index was modelled after the parameters of a gravity model based on interstate freight flows (Duran-Fernandez and Santos 2014). Due to the fact that the higher hierarchy regions strongly dominate the economy of their states, the index must be a good approximation to the actual domestic freight flows for those regions. Moreover, this possibility does not explain why we do identify a positive effect for the international index.

A simpler and more straightforward explanation is that the spillovers of domestic routes have been internalised. Accessibility would present significant spillovers on production only if its benefits on production were larger than the transport costs. This is clearly the case when firms are granted free access to road infrastructure. However, if any positive spillover has been internalised by the cost structure of the firms, higher accessibility is not necessarily associated to a higher per worker output.

The NEC contains some basic information about transport costs for firms. Unfortunately, it only presents data on distribution costs paid to logistic suppliers, data that is not useful to estimate tolls paid or the total transport costs at aggregate level. Nevertheless, we can make important inferences on the dependence of domestic routes analysing the Mexican road system.

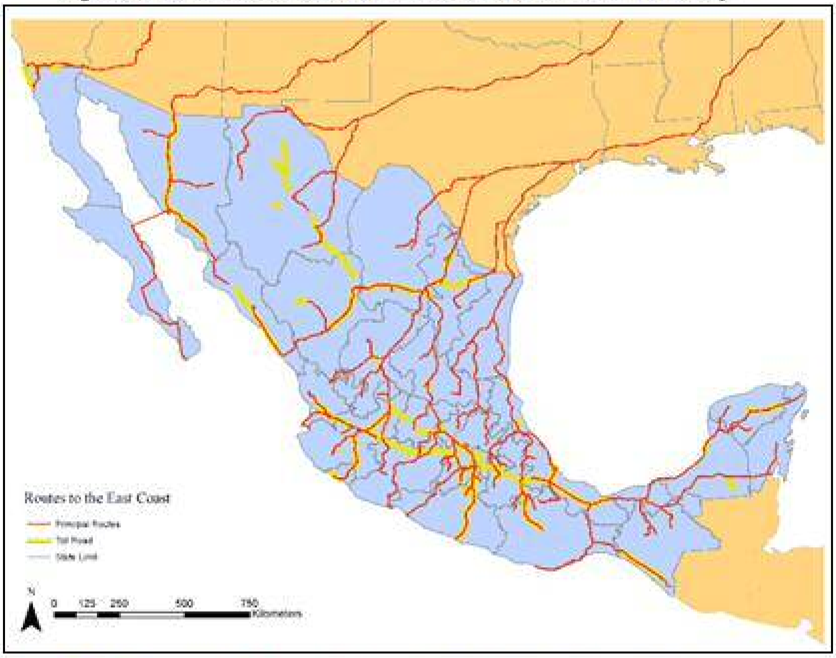
The North American GIS Road Network Model (Duran-Fernandez 2014) allows us to estimate the contribution of a particular link of the road network to accessibility. In particular, we estimated the accessibility gains attributed to the toll-subnetwork for every region as $\Delta A_i = A_i - A_i'$, where A_i is the actual accessibility and A'_i is the accessibility that we would observe if the toll-subnetwork did not exist. Under this context, ΔA can be interpreted either as the direct benefits of the toll-subnetwork or a measure of dependence on this subsystem.

The estimation of ΔA shows that regions with higher domestic accessibility are the ones where the contribution of the toll-subnetwork is larger. On the other hand, the higher the international accessibility, the lower the contribution of the toll-subnetworks to this variable. This pattern is clearly identified by the estimated simple correlations between ΔA_i and A_i, which are 0.49 and -0.03 for the domestic and international index, respectively.

A'_i, would be the observed accessibility if the demand for the use of these roads were zero. This hypothetical situation would arise if the marginal benefit of travelling on toll roads were lower than the tolls. Thus, the contribution of the toll-subnetwork to accessibility can be interpreted as an upper bound to the tolls faced by firms using the subsystem. Under this assumption, firms located in regions with high domestic accessibility will potentially face larger transport costs, given their higher dependence on the toll-subnetwork. Domestic freight is expected to be concentrated in high domestic accessibility regions, due to the fact that accessibility is a measure of market potential. Therefore, a monopolistic concessionaire, not subject to appropriate regulation, is able to extract high rents from an important fraction of the users of the domestic routes. At aggregate level, a positive spillover on production derived from higher domestic accessibility levels will be internalised by the cost structure of the transport sector.

On the other hand, the analysis shows that firms in high international accessibility regions are not very dependent on toll roads. As in the case of domestic accessibility, international freight is expected to be concentrated in high international accessibility regions. However, given that freight is less dependent on toll roads (given the existence of alternative free routes), a monopolistic concessionaire cannot extract monopolistic rents, and therefore, at aggregate level, the presence of the toll-subsystem does not neutralise the spillovers on production derived from higher accessibility to international markets. This can be indirectly illustrated in an analysis of international routes. For example, Figure 1 shows the optimal routes that link Mexico's regions to New York City. These routes, in their Mexican section, are the same as the ones that link Mexico to the Southwest, the Great Lakes, and the Mideast, the main origins and destinations of freight to and from the USA, according to the BTS. This figure shows that the estimated routes do not rely on the toll sections of the road network. This is particularly important for the regions in the Central Plateau, in the states of Guanajuato, San Luis Potosi, Tamaulipas, and Zacatecas. These states are linked to the USA market almost entirely by free roads of the secondary network. Due to the fact that road infrastructure does not have an additional cost to the firms located in these regions international accessibility generates positive and significant spillovers. The quantitative analysis has identified these effects at aggregate level.

Figure 1 International Trade Corridors to/ from New York City



Source: Own elaboration (Digital Cartography from the Municipal Geo-statistical Framework and the Topographic Digital Dataset, INEGI)

Table 3 presents the results of a regression analysis, which provides additional evidence for the internalisation process based on this variable. We regress intermediate consumption per worker against the domestic and international accessibility gains attributed to the toll-subnetwork, A'_i. The estimation assumes an error structure similar to the one presented in the estimation of the regional production function (Equation 1) with two kinds of disturbance: a region and an industry specific residual.

The results of the regression show that the intermediate consumption per worker, which is the unit cost of production excluding labour costs, is systematically higher in the regions that exhibit the higher benefits from the toll-subsystem. This result supports the hypothesis that positive spillovers derived from a better access to the domestic market are internalised by higher costs in the form of tolls.

Surprisingly, the international accessibility gains attributed to the toll-subnetwork do not have a significant effect on unit intermediate costs. This outcome supports the claim that, given the structure of the toll-subnetwork and the geometry of the main international trade corridors, any internalisation effect at international level is not large enough to be captured by the analysis. Alternatively, the positive spillovers are very large relative to the tolls, which only internalise a fraction of them.

Table 3 Effect of Accessibility Improvements Attributed to the Toll-Subnetwork on Intermediate Consumption per Worker
(Standard Errors in Parenthesis)

	Manufacturing Sector	
Domestic	0.0442	*
	-0.0149	
External	-0.0193	
	-0.0278	
Constant	3.6464	*
	-0.1366	
Variance		
σ_{υ}^{2}	0.606	
σ_{ε}^{2}	1.301	
ρ /1	0.178	
N	11,654	
Groups	133	
R^2 (Overall)	0.0136	
Wald χ (11)	8.98	*

/1 Fraction of variance due to regional effect u_i
Significance level *1% ** 5% *** 10%
Source: Own calculation

6.3 Accessibility and Parameters' Robustness

An alternative explanation for the non-significant impact on domestic accessibility might be linked to the misspecification of the accessibility index. This possibility arises because the elasticity of the accessibility index was estimated using inter-state freight flows. If the actual elasticity of interregional trade flows were different from the estimated value the real market potential for a region could be different from the one estimated by the model.

Duran-Fernandez and Santos (2014) show that the behaviour of the accessibility model used in the econometric assessment is robust to different elasticities, definitions for economic weights, and specifications. Nevertheless, in order to push the analysis further, we estimate the model in Equation 1 using different versions of domestic and international accessibility. Each version of these variables is assessed under different elasticity assumptions in a range between 0.1 and 1.5. Table 4 presents the results of this exercise. The coefficient of domestic accessibility is non-significant for any elasticity within this range and the coefficient of international accessibility is positive and significant for all the elasticities. This suggests that the non-significance of this coefficient is robust.

Table 4 Sensitivity Analysis of Domestic and International Accessibility
GLS Regression
(Standard Errors in Parenthesis)

θ	Domestic	International	
0.10	-0.0048	0.0155	*
	(0.0032)	*(0.0033)*	
0.20	-0.0034	0.0110	*
	(0.0021)	*(0.0022)*	
0.30	-0.0025	0.0102	*
	(0.0017)	*(0.0020)*	
0.40	-0.0020	0.0106	*
	(0.0014)	*(0.0021)*	
0.50	-0.0017	0.0118	*
	(0.0013)	*(0.0022)*	
0.60	-0.0014	0.0135	*
	(0.0012)	*(0.0025)*	
0.70	-0.0012	0.0158	*
	(0.0012)	*(0.0028)*	
0.80	-0.0011	0.0187	*
	(0.0012)	*(0.0033)*	
0.90	-0.0010	0.0222	*
	(0.0012)	*(0.0038)*	
1.00	-0.0009	0.0265	*
	(0.0012)	*(0.0044)*	
1.10	-0.0008	0.0314	*
	(0.0011)	*(0.0051)*	
1.20	-0.0008	0.0370	*
	(0.0012)	*(0.0059)*	
1.30	-0.0007	0.0431	*
	(0.0012)	*(0.0068)*	
1.40	-0.0007	0.0497	*
	(0.0012)	*(0.0078)*	
1.50	-0.0007	0.0566	*
	(0.0012)	*(0.0089)*	

Significance level *1%
Note: θ is the elasticity of time respect to traffic flow
Source: Own Calculation

6.4 Asymmetric Effects of Road Infrastructure

The econometric results show that international accessibility has positive and significant effects on production per worker. This result indicates that regions benefit symmetrically from the proximity to international markets, regardless of their hierarchy. Nevertheless, the estimation presents some evidence of asymmetric effects of regional accessibility. In particular, the effects of *network curvature* are significant for regions in Group 1, while regional accessibility has a positive effect on regions in Group 2.

Due to the fact that the hierarchical grouping classifies regions according to similarities in their industrial characteristics, these asymmetries can be attributed to actual variations in production technologies. For example, regions in Group 1 are mainly specialised in the manufacturing of petroleum and coal products (81 percent of domestic production). The input factors of the sector can be obtained only from specific locations (oil production centres) and its production is not locally distributed. Another important characteristic of the sector is that it depends on pipelines, not road infrastructure, as a channel for distribution and supply. Therefore, an efficient geometry and high quality of regional roads, rather than the overall articulation of the regional market, has a positive impact on production per worker. This effect is captured by the significant coefficient of network curvature and the negligible impact of regional accessibility.

On the other hand, regions in Group 2 are specialised in the production of transport equipment (43 percent of domestic production). These sectors can potentially obtain their productive inputs in the region's markets and production can be partially absorbed by them. Hence, the articulation of the regional markets, which is captured by regional accessibility, has a positive effect on the output per worker.

7 CAUSALITY ANALYSIS

7.1 Possible Sources of Endogeneity

Endogeneity can arise under three circumstances. First, policy makers invest more in regions where output per worker is higher. Second, policy makers may use infrastructure investment as a counter-cyclical fiscal instrument, investing more in regions with negative disturbances on production per worker. Third, agents subject to negative disturbances on output per worker may persuade policy makers to invest more in their regions. In any of these cases, the infrastructure component of accessibility would be endogenous.

The historic investment trends in Mexico have interesting implications for endogeneity. The main structure of the National Road System in Mexico (NRS) was built during the 1950s and 1960s. In that period, Mexican economic policy was focused on the development of an indigenous industrial sector. During this period, the country investment in public roads focused on the articulation of the domestic market, linking the main production centres in the country with Mexico City. This radial structure of the road network survives up to this date, and no major programs to change its structure have ever been implemented (Bassols-Batalla 1993, Garcia Martinez and Takako 1992). In the early 1990s a major investment programme took place (Solidaridad, President's Office 1994). The programme focused on the improvement of unpaved local roads and the upgrade of the level of service of some corridors in the existing network. However, the basic radial structure of the network was not modified. This development illustrates that international trade was a minor decision variable during the period when the structure of the road network was planned. Thus, the infrastructure that today links Mexico to the USA arose mainly as a by-product of the radial network designed 50 years ago (Bassols-Batalla 1993).

These circumstances justify the use of international accessibility as an exogenous variable. For the case of the roads that link the domestic and the regional markets, the possibility that policy makers in the 1950s and 1960s based their infrastructure plans in the regional disturbances on production per worker that the country would experience in the early 2000s is remote. The early 1990s infrastructure programme represents a more serious challenge, because this scenario is not impossible at all. Nevertheless, even if this were the case the historic set up of the road infrastructure indicates that at least the structure of the network is exogenous

A real problem that we face is the possibility that the investment criterion on infrastructure in Mexico has been based on productivity. If this were the case, our quantitative analysis would face self-selection problems, because output per worker is a measure of productivity. If the spatial distribution of productivity has been stable in the last 50 years, it could be possible that policy makers invested largely in highly productive regions that were still highly productive in the early 2000s. In that case, the relationship between the production per worker and the value of the road network would be spurious.

A flaw of this argument is that in a developing country, investment in poor low productive regions will typically command political support. In such case, the configuration of the transport network that arises is related to political variables rather than productivity, a situation that does not lead to a self-selection bias. Another problem of this argument is linked to the stability of the spatial distribution of productivity. The opening of the economy to international trade in the late 1980s changed the spatial distribution of the industrial activities in Mexico. Therefore, even if a productivity criterion was applied in the 1950s and 1960s it does not necessarily associate large levels of infrastructure to high productive regions in the early 2000s.

On top of that, there is statistical evidence that suggests that the productivity criterion has not been applied in Mexico. First, the fact that accessibility has asymmetric effects on production per worker according to the hierarchy of the region indicates that policy makers have not followed a pure productivity investment criterion; otherwise, the estimated correlation would be positive, independently of the hierarchy. In fact, the simple correlation between these two variables is not significantly different from zero (-0.01). The output per worker and hierarchy are positively correlated, showing that high hierarchy regions are more productive on average[13]. Policy makers might have followed a combined investment criterion based on hierarchy and productivity. The problem here is that the hierarchical score –*i.e.* the Adjusted Shannon Entropy Index- is not correlated with domestic accessibility at all. Therefore, the output per worker data does not provide evidence of any productivity criterion being applied to infrastructure investment policies.

[13] The correlation between the hierarchical ranking *H* and product per worker is 0.32 and significant at a one percent level.

A final possibility that is compatible with the observed data is that policy makers followed a productivity criterion but they were unaware of the effects of investment on the value of the road network. Therefore, even if investment is concentrated into highly productive regions, this policy does not necessarily increase the market potential of those regions. Thus, the value of the road network measured through accessibility could be treated as an exogenous variable. Despite the data matching this description, the lack of information on historic infrastructure investment flows at regional level prevents us from verifying this possibility in a rigorous way.

Given these facts, the endogeneity problem is centred on two concerns. The first is that the early 1990s infrastructure programme was targeted to regions with high production per worker. The second is that the upgrade of some corridors of that programme is correlated to the same disturbances on output per worker that affected the manufacturing sector in Mexico in the early 2000s, the period that we are analysing. In order to formally rule out these cases, we present two extensions to the quantitative analysis. The first identifies the regions and observations that are more likely to be affected by this problem, given the mechanisms that have been proposed as a source of these problems. Then the model is estimated excluding these observations. The second approach proposes a set of instruments and estimates the model under Generalised Least Square Instrumental Variables (GLSIV).

7.2 Restricted Samples

In the dataset, that most observations represent regional manufacturing activities with a minimum weight on the domestic economy. Despite being aggregated at regional level, the industrial disaggregation of the dataset generates very small observation units. Individually, the firms that integrate these industries have a very small weight on the domestic and regional economy. Therefore, it is extremely unlikely that they were able to influence national policies on infrastructure investment. The scale of operation is not directly related with the production per worker: it is perfectly possible to observe small-scale industries with high production per worker, as well as large-scale ones with low production per worker. However, it is more likely that large-scale industrial sectors push policy makers to make larger investments in their regions.

An important sector where this mechanism could apply is the oil industry. The oil related industry represents 8.5 percent of the domestic manufacturing GVA in Mexico. Its share could be as large as 80 percent of the GVA of the manufacturing sector in some regions. The output per worker in oil regions is slightly higher than the national average but the difference is statistically significant.[14] Moreover, refinery and the petrochemical industry are activities reserved exclusively to the state-owned monopoly Pemex. In 2004, taxes on oil revenues represented almost 30 percent of the federal government revenue. This picture could lead to a situation in which transport investment decisions are biased towards these regions.[15]

Oil regions are defined as those where oil extraction, refinery, and petrochemical activities represent more that 15 percent of the manufacturing GVA. Under this criterion, we identified eight of them. All other industrial activities in these regions were excluded from the sample, including those activities large enough to influence infrastructure policy at national and regional level. Only oil related activities were kept in the sample. Further details are presented in the Appendix.

The aforementioned cases account for 1,174 observations that are excluded from the sample to estimate a new version of the model. Assuming that the remaining observations do not have any effect on infrastructure policy, the estimation of the model should be consistent, conditional on the restricted sample. If the excluded observations originate a self-selection bias in the original model, the coefficients estimated under the restricted model will be significantly different. If they do not originate a self-selection bias, the coefficients estimated

[14] Productivity in oil and non-oil regions is 4.0 and 3.9, respectively. The differences between means is significant (t=3.4).

[15] Nevertheless, the accessibility differences between oil and non-oil regions do not support this possibility. Accessibility in oil and non-oil regions is 308.3 and 303.6, respectively. The differences between means is non-significant (t=0.1).

under the restricted model will not be significantly different and we will be able to conclude that the original model was consistent.

In the first column of Table 5 we present the results of the estimation under the restricted sample. The results for the infrastructure variables are not different from those from the original estimation: international and regional accessibility have positive and significant effects, domestic accessibility is not significant, and the network curvature index is negative and significant. The significance levels do not change for the rest of the variables either, with the exception of industrial diversity, which becomes significant under the restricted sample.[16] Finally, the estimated variance does not present any substantial variations and the residuals do not exhibit spatial autocorrelation either.

[16] The oil regions have on average lower industrial diversity and high productivity. Their exclusion from the sample pulls the coefficient upwards.

Table 5 GLS Regression with Geographical Controls

(Standard Errors in Parenthesis)

	Restricted Sample GLS		Instrumental Variables IVGLS	
Private Factors				
ln(K/L)	0.1308	*	0.1370	*
	(0.0065)		*(0.0062)*	
ln(IC/L)	0.5721	*	0.5809	*
	(0.0075)		*(0.0071)*	
Accessibility and road network structure				
Domestic	-0.0001		-0.0033	**
	(0.0013)		*(0.0015)*	
International	0.0099	*	0.0144	*
	(0.0021)		*(0.0032)*	
Regional	0.0023	**	0.0065	*
	(0.0009)		*(0.0021)*	
Network Curvature	-0.1690	*	-0.1699	
	(0.0687)		*(0.1131)*	
Geographic controls				
Port Pacific	0.0030		0.0014	
	(0.0059)		*(0.0057)*	
Port Gulf	0.0082		0.0052	
	(0.0073)		*(0.0072)*	
South	-0.0081		-0.0064	
	(0.0169)		*(0.0181)*	
Other Variables				
Schooling	0.0462	**	0.0193	
	(0.0229)		*(0.0290)*	
H	0.0284	**	0.0370	***
	(0.0228)		*(0.0217)*	
Constant	0.3128		0.3333	
	(0.2141)		*(0.3056)*	
Variance				
$\beta\sigma_{\upsilon}^{2}$	0.133		0.126	
σ_{ε}^{2}	0.761		0.763	
ρ /1	0.030		0.027	
N	10,302		11,476	
Groups	124		133	
R^2 (Overall)	0.6390		0.6414	
Wald c (11) /2	13,967.64	*	16,819.39	*

/1 Fraction of variance due to regional effect u_i

/2 □□(3) for G3

Significance level *1% ** 5% *** 10%

Note: K, stock of private capital; IC, intermediate consumption; L, labour

Source: Own calculation

7.3. Instrumental Variables

In order to rule out any possible endogeneity of the infrastructure variables, we carry out an instrumental variable analysis. The objective of this approach is to find a set of instruments correlated to the explanatory variables but independent from output per worker. We use as instruments historic demographic data for Mexico and information on the physical geography of the different regions. The variables that are instrumented are domestic, international, and regional accessibility, and network curvature. Further details about the data used to build the instruments for this analysis is presented in the Appendix.

The historic demographic data was taken from the National Censuses 1950 and 1960.[17] The analysis uses the following instruments: population density in 1950, population density in 1960, the interaction of these two variables, and the percentage of indigenous population in 1960. In addition, it includes the distance to Mexico City, as well as the distance to the closest border point with the USA, the mean altitude of each urban centre in a region, as well as the standard deviation of this variable.

The demographic instruments attempt to model the historic development of the National Road System. In the 1950s, the network was built up by regional roads linking only close urban centres. It was not until the 1960s that the communications among the principal cities were fully articulated at national level. Population density has been related to urbanisation levels (Berry 1967); thus, the inclusion of the historic population densities as explanatory variables of accessibility must capture the evolution of the infrastructure component of these variables.

The regions with large indigenous population share have been historically located in areas of difficult access in the mountainous areas of the states of Chiapas, Chihuahua, Oaxaca, and Puebla, as well as in the rain forest of the Yucatan Peninsula. However, some regions in the

[17] Data was provided by Prof. Donald J Treiman from the Department of Sociology, University of Wisconsin, and verified against the original source.

central plateau also have a large indigenous population. Historically, indigenous regions have been economically deprived and underrepresented in the political arena. These characteristics should have an impact on the value of the infrastructure stock of these regions. The '*percentage of indigenous population*' in a region is included to capture at least part of these effects.

The inclusion of the '*distance to Mexico City'* variable attempts to capture the development of the road network into a radial structure with Mexico City in the centre. The inclusion of the '*distance to closest border point*' variable attempts to isolate the fixed geographic component of international accessibility. Finally, the mean and the standard deviation of the altitude of the urban centres in a region quantify the evenness of its landscape, a variable that must be correlated with curvature.

The long time span between the instruments, all of which correspond to the 1950s and 1960s, and the dependent variable, which corresponds to 2004, makes any dependence between them extremely unlikely. In addition, the deep economic and social transformation that Mexico has experienced in the last 40 years (1960s to 2000s) has had a significant impact in the socioeconomic geography of the country. To assume that this new spatial arrangement has been determined by population densities in the 1950s and 1960s or by the other controls included in the first stage regression is a position that has been widely criticised in the development literature (Gleaser, *et. al.* 2004, Murphy, *et. al.* 1991).

We estimated the model presented in Equation 1 using Instrumental Variables Generalised Least Squares (IVGLS). All the infrastructure variables were instrumented. In general, the first stage regression shows that the instrumental variables are highly significant for all the instrumented variables. In the second column of Table 5 we present the results of the estimation under the restricted sample. Table 6 presents the results of the first stage regression.

The estimated coefficients for the private factors do not change significantly in comparison to the GLS results presented in Section 5. In general, the coefficients maintain the signs. However, the analysis shows three differences with respect to the canonical model. The first is the loss of significance of network curvature; the second is that the negative coefficient of domestic accessibility becomes significant at five percent level. Finally, industrial diversity becomes significant.

If the original estimation had been subject to endogeneity problems, the difference between the GLS and the IVGLS coefficients could be interpreted as their biases. The Hausman Specification Test proposes a statistic based on the difference between the two estimates to test for endogeneity. The result of the Haussmann Specification Test does reject the hypothesis that the GLS and the IVGLS estimations are systematically different (χ^2_{11}=8.62). The test does not provide a direct proof of the absence of endogeneity in the model; however, if this were the case the test would have identified systematic differences in the estimated coefficients, accounting for the bias of the GLS estimates. Therefore, the test provides strong evidence against endogeneity of the infrastructure variables.

Under a non-endogeneity scenario, the differences in the significance level of domestic accessibility and curvature can be directly attributed to the loss of efficiency of the IVGLS regression in comparison to the original model. Therefore, the instrumental variable analysis does not suggest the need to make any changes in the interpretation of the results of the canonical model.

Table 6 Instrumental Variable Regression. First Stage Regressions

(Standard Errors in Parenthesis)

	Domestic		International		Regional		Network Curvature	
Instruments								
Historic Demography								
Density 1950	0.5296	*	-0.1137	*	0.1373	*	-0.00380	*
	(0.0149)		*(0.0046)*		*(0.0379)*		*(0.00042)*	
Density 1960	-0.1686	*	0.0885	*	0.2718	*	0.00334	*
	(0.0113)		*(0.0035)*		*(0.0289)*		*(0.00032)*	
Interaction Density 1950x1960	-0.0003	*	0.0000	*	-0.0006	*	0.00000	*
	(0.0000)		*(0.0000)*		*(0.0000)*		*(0.00000)*	
Indigenous	2.6690	*	-0.6985	*	3.3268	**	0.15447	*
	(0.5525)		*(0.1713)*		*(1.4060)*		*(0.01558)*	
Distance to:								
Mexico City	-0.0165	*	0.0058	*	0.0095	*	0.00000	
	(0.0003)		*(0.0001)*		*(0.0009)*		*(0.00001)*	
Northern Border	-0.0043	*	-0.0138	*	-0.0004		-0.00005	*
	(0.0002)		*(0.0001)*		*(0.0006)*		*(0.00001)*	
Orography								
Mean Altitude	2.5628	*	0.2245	*	-0.7118	*	-0.00478	**
	(0.0678)		*(0.0210)*		*(0.1726)*		*(0.00191)*	
Std Dev. Altitude	-0.0171	*	-0.0019	*	-0.0011		0.00073	*
	(0.0004)		*(0.0001)*		*(0.0010)*		*(0.00001)*	
Explanatory Variables in the Model								
ln(K/L)	-0.0141		-0.0191	**	0.0217		0.00090	
	(0.0259)		*(0.0080)*		*(0.0660)*		*(0.00073)*	
ln(IC/L)	-0.0559	***	0.0011		0.3536	*	-0.00124	
	(0.0293)		*(0.0091)*		*(0.0745)*		*(0.00083)*	
Schooling	1.0587	*	0.0384		5.2649	*	-0.07904	*
	(0.0882)		*(0.0273)*		*(0.2244)*		*(0.00249)*	
H	-2.2850	*	0.0739	**	-7.9640	*	0.00216	
	(0.0976)		*(0.0303)*		*(0.2483)*		*(0.00275)*	
South	0.3054	*	0.4013	*	-0.3057		-0.00015	
	(0.0786)		*(0.0244)*		*(0.2001)*		*(0.00222)*	
Port Pacific	-0.9181	*	-0.0703	*	-0.7142	*	-0.00075	
	(0.0232)		*(0.0072)*		*(0.0592)*		*(0.00066)*	
Port Gulf	-0.4719	*	-0.4444	*	-0.7450	*	-0.00280	*
	(0.0373)		*(0.0116)*		*(0.0950)*		*(0.00105)*	
Constant	45.3056	*	53.5689	*	-9.7012	*	1.88412	*
	(0.9859)		*(0.3057)*		*(2.5089)*		*(0.02780)*	
Wald□□ (15)	88,186.00		172,526.00		6,272.00		15,634.00	

Significance level *1% ** 5% *** 10%

Note: K, stock of private capital; IC, intermediate consumption; L, labour

Source: Own calculations

8. MAGNITUDE OF THE SPILLOVERS ON THE AVERAGE PRODUCT OF LABOUR

We define the magnitude of the spillovers of a variable q on output per worker in a region j as the product of the coefficient of variable q in the estimated production function ($\hat{\beta}_q$) and the value of q in the region. The aggregate magnitude of the spillovers of variable q on per worker production (S_q) is defined as the weighted average of individual spillovers for each observation in the sample, using as weight (W) the share of manufacturing GVA of each observation in the manufacturing GVA for the whole country, as presented in Equation 5. In this exercise, variable q is defined as the variable that exhibits significant coefficients in the estimated production function. Thus, q is international accessibility, regional accessibility, and network curvature in the different estimations.

Equation 5
$$S_q = \sum_{j=1} W_j \hat{\beta}_q q_j$$

Table 7 shows the results of the average impact of international and regional accessibility on production per worker as well as the loss due to the regional network curvature. The first column presents the value of S_q for q= {international accessibility, regional accessibility, network curvature}. These effects are expressed as an increment on the natural logarithms of production per worker. The exponential of this value can be interpreted as a multiplier effect, given a demand for private factors. If the multiplier is larger than one, it is a positive spillover, if it is less than one it is a loss on output per worker.

The total multiplier effect of international accessibility on output per worker at national level is 2.05. This multiplier implies that if accessibility increased by one percent for all the regions, the improvement in the output per worker would be 0.72 percent. This can be easily seen as the elasticity of production per worker with respect to international accessibility. The results show that the elasticity is significantly lower for the low hierarchy regions in Group 1, and more or less the same for regions in Groups 2 and 3. Hence, given the current accessibility endowments, a proportional improvement on this variable has larger effects on the high hierarchy regions.

We extend this exercise to estimate the overall effect of regional accessibility. The results show that at national level, the elasticity of the output per worker with respect to regional accessibility is 0.02. For the regions in Group 2, the elasticity of regional accessibility is larger: 0.07. Finally, the elasticity of output per worker with respect to the network curvature is estimated at -0.09 at national level, and -0.31 for Group 1.

The local effects of regional accessibility and curvature are larger than the domestic effects. These results indicate that the impact of infrastructure on output per worker at regional level is not as large as the effects of international accessibility. However, it is worth bearing in mind that in many regions road infrastructure is not well developed. Therefore, investment in regional roads can potentially generate major improvements in both regional accessibility and network curvature and, as a result, have positive effects on output per worker.

Table 7 Estimated Magnitudes of the Effects of Accessibility and Curvature Network and Curvature by Hierarchical Rank

	Δln (y)	Multiplier	Implied Elasticity
National Average (N=11,476, Weight=100%)			
International Accessibility	0.71972 *(0.1713)*	2.05	0.72
Regional Accessibility	0.01823 *(0.0436)*	1.02	0.02
Network Curvature	-0.08523 *(0.1387)*	0.92	-0.09
Group 1 (N=5,603, Weight=27.7%)			
International Accessibility	0.52215 *(0.0672)*	1.69	0.52
Regional Accessibility	n.s.	n.s.	n.s.
Network Curvature	-0.30776 *(0.0318)*	0.74	-0.31
Group 2 (N=4,364, Weight=27.6%)			
International Accessibility	0.83041 *(0.1409)*	2.29	0.83
Regional Accessibility	0.06604 *(0.0610)*	1.07	0.07
Network Curvature	n.s.	n.s.	n.s.
Group 3 (N=1,509, Weight=44.7%)			
International Accessibility	0.77377 *(0.1262)*	2.17	0.78
Regional Accessibility	n.s.	n.s.	n.s.
Network Curvature	n.s.	n.s.	n.s.

Source: Own calculation
n.s. Note Significant

9 FINAL REMARKS

Transport infrastructure has important effects on the economy, and it has been identified as one of the fundamental determinants of regional development. However, the empirical research that has tackled this linkage faces one common problem: finding a suitable scalar measure that captures the effects of transport infrastructure on the economy. Another important problem, pointed out by the theory, is that these effects can be uneven across locations, a situation that could mislead the estimation of its overall impact.

In this paper, we have used three AMs to model the value of transport infrastructure and its impacts on production per worker for the manufacturing sector. We also took into account the geometry of the road network, and classified each region in Mexico according to a hierarchical ranking, based on the Shannon Entropy Index. The ranking explicitly introduces the possibility that the aggregate effects of the road network on the economy depend on the hierarchy of a region, which is a direct measure of its industrial profile.

The main conclusion of the econometric analysis is that the geographic landscape of the production per worker in the manufacturing sector in Mexico depends in an important way on road infrastructure endowment. Road infrastructure in general has a positive and significant impact on regional variations in output per worker. Accessibility to international and regional markets has large and significant spillovers on output per worker, while a poor quality of roads at regional level can be associated to losses on output per worker. The study does not find any significant effect of domestic accessibility on output per worker, a result that we attribute to the toll-subnetwork.

The analysis shows that the effect of international accessibility is even, regardless of the hierarchical ranking of a region. However, at regional level we have identified asymmetric effects that can be directly associated to the industrial profile of the different regions.

Endogeneity and self-selection threaten the validity of the econometric assessment. However, the analysis presented in Section 7 suggests that these problems are not present in our model. To formally address these problems, we estimated two alternative versions of the model: one excludes observations from the dataset, which were likely to be subject to self-selection bias, and the other one uses instrumental variables. The results of the model that excludes observations are similar to the ones obtained in the original analysis. The results of the model with instrumental variables were subject to the Haussmann Specification Test, which showed that the estimates were not systematically different from the estimates computed with the original model. These results provide strong evidence on the consistency of the results reported in the econometric analysis.

Nevertheless, there is a possibility that is not ruled out by the analysis. Two of the main components of accessibility, infrastructure and land use patterns, ultimately depend on the historic spatial distribution of land use as well as on the historic set-up of the road network. On one end of the spectrum, land use patterns could depend entirely on the historic configuration of transport infrastructure. In this case, the effects of accessibility we have identified would be attributed uniquely to infrastructure. At the other end of the spectrum, the current set-up of the road network could depend entirely on the historic structure of land use, so this structure would affect current production patterns. In this case, the positive effects of accessibility would be explained ultimately by the historic spatial distribution of land use. However, the idea that historic events entirely determine the current level of productivity, measured as the output per worker, have been widely criticised from a theoretical perspective in the development literature. Unfortunately, an actual empirical test that rules out this possibility would require truly random data from a social experiment that in practice, would be impossible to collect. Given this restriction, this research has used the best available data and standard quantitative methods to study the relationship between transport infrastructure and the economic geography of Mexico.

Acknowledgements and disclaimer

The authors are grateful to two reviewers and to the attendees to the 11th Meeting of the World Conference on Transport Research Society, Berkeley, June 24-28 2007, where an earlier version of this paper was presented. This study was financed by the Mexican Federal Government through the National Council of Science and Technology (CONACYT). Any opinions, findings, conclusions and recommendations expressed in this paper are those of the authors alone and should not be attributed to any other person or entity.

REFERENCES

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

Bassols-Batalla, A. (1993) "Formación de regiones económicas" UNAM, Mexico City, Mexico.

Bassols-Batalla, A. (2002) "Geografía socioeconómica de México" Trillas, Mexico City, Mexico.

Berry, B. (1967) "Geography of market centres and retail distribution" Prentice-Hall, NJ, USA.

Boarnet, M. (1996) "The direct and indirect economic effect of transportation infrastructure" WP No. 340, Department of Urban and Regional Planning and Institute for Transportation Studies, University of California at Irvine, USA.

Boarnet, M. (1998) "Spillovers and the locational effects of public infrastructure" Journal of Regional Science (38:3) pp. 381-400.

Castañeda Sabido, A., P. Cotler, Gutierrez, O. (2000) "The impact of infrastructure on Mexican manufacturing growth" Economía Mexicana, Nueva Época (9:2) pp. 143-64.

Chiquiar, D. (2004) "Why Mexico's regional income convergence broke down?" Journal of Development Economic (77) pp. 257-275.

Costa-i-Font, J. and Rodriguez-Oreggia E. (2004) "Is the Impact of Public Investment Neutral across the Regional Income Distribution? Evidence from Mexico" WP No. 113, Colleccio d'Economia, Universitat de Barcelona, Spain.

Deichmann, U., Fay, M. and Lall, S. (2004) "Economic structure, productivity, and infrastructure quality in southern Mexico" Annals of Regional Science (38:3) pp. 361-85.

Deichmann, U., Kaiser, K., Lall. S. and Shalizi, Z. (2005) "Agglomeration, transport, and regional development in Indonesia" WP No. 3470, World Bank Policy Research, World Bank, DC, USA.

Dendrinos, D.S. and Mullally H. (1985) "Urban evolution studies in the mathematical ecology of cities" Oxford University Press, Oxford, UK.

Duran Fernandez, R. (2014) "A GIS Model of the National Road Network in Mexico" *Research in Transportation Economics, **THIS** ISSUE.*

Duran Fernandez, R. and Santos, G. (2014) "A Regional Model of Road Accessibility in Mexico: Accessibility Surfaces and Robustness Analysis" *Research in Transportation Economics, THIS ISSUE.*

Evans, P. and Karras, G. (1994) "Are government activities productive? Evidence from a panel of US states" The Review of Economics and Statistics (76:1) pp. 1-11.

Feltenstein, A. and Ha J. (1995) "The role of infrastructure in Mexican Economic Reform" World Bank Economic Review (9:2) pp. 287-304.

Fuentes, N.A. and Mendoza Cota J. E. (2003) "Infraestructura pública y convergencia regional en México, 1980-1998" Comercio Exterior (53:2) pp. 178-87.

Garcia Martinez, B and Takako, S. (1992) "Las carreteras de México 1891-1991" Edited by Secretaría de Comunicaciones y Transporte, Mexico City, Mexico.

Garcia-Mila, T., McGuire, T. and Porter, R. (1996) "The effect of public capital in state-level production functions reconsidered" The Review of Economics and Statistics (78:1) pp. 177-180.

Gleaser, E. La Porta, R. Lopez de Silanes, F. and A. Shleifer (2004) "Do institutions cause growth?" WP No. 10568, National Bureau of Economic Research, USA.

Holtz-Eakin, D. (1994) "Public-sector capital and the productivity puzzle" The Review of Economics and Statistics (76:1) pp. 12-21.

Johnson, K, and Kort, J. (2004) "Redefinition of the BEA economic areas" Internet publication from the Bureau of Economic Analysis http://www.bea.gov/SCB/PDF/2004/11November/1104Econ-Areas.pdf

Kelejian, H. and Robinson, D. (1997) "Infrastructure productivity estimation and its underlying econometric specifications: a sensitivity analysis" Regional Science (76:1) pp. 115-131.

Lall, S. and Chakravorty, S. (2005) "Industrial location and spatial inequality: Theory and evidence from India" Review of Development Economics (9:1) pp. 47-68.

Lall, S. and Yepes, T. (2004) "Location, concentration, and performance of economic activity in Brazil" WP No. 3268, World Bank Policy Research, World Bank, DC, USA.

Lall, S., Shalizi, Z. and Deichmann, U. (2001) "Agglomeration economies and productivity in Indian industry" WP No. 2663, World Bank Policy Research, World Bank, DC, USA.

Looney, R. and Frederiksen P. (1980) "The Regional Impact of Infrastructure Investment in Mexico" Regional Studies (15:4) pp. 285-96.

Mallick, R. and Carayannis, E. (1994) "Regional Economic Convergence in Mexico: An Analysis by Industry" Growth and Change (25) pp. 325-334.

Murphy, M., Shleifer, A. and Vishny, R. (1991) "The allocation of talent: implications for growth" Quarterly Journal of Economics (106) pp. 503-530.

Ramirez, M. D. (2004) "Is Public Infrastructure Spending Productive in the Mexican Case? A Vector Error Correction Analysis" Journal of International Trade and Economic Development (13:2) pp. 159-178.

Shannon, C. and Warren, W. (1963) "The mathematical theory of communication" University of Illinois Press, ILL, USA.

Tabushi, T. and Thisse, J. (2005) "Regional specialization, urban hierarchy, and commuting costs" International Economic Review (47:4) pp. 1295-1317.

Appendix

In this Appendix, we present a description of the activities that were excluded from the restricted sample analysis. We show an exhaustive definition of oil-related industries according to the six digits NAIC criterion (Table A1), as well as some basic statistics of oil regions (Table A2), and a summary of the observations that were excluded due to their weight in their regional economy (Table A3).

Table A1 **Definition of Oil Activities**

NAIC6	Industrial Activity
211110	Oil and Gas Extraction
213111	Drilling Oil and Gas Wells
213119	Support Activities for Oil and Gas Operations
324110	Petroleum Refineries
324120	Asphalt Paving, Roofing, and Saturated Materials Manufacturing
324121	Asphalt Paving Mixture and Block Manufacturing
324122	Asphalt Shingle and Coating Materials Manufacturing
324190	Other Petroleum and Coal Products Manufacturing
324191	Petroleum Lubricating Oil and Grease Manufacturing
324199	All Other Petroleum and Coal Products Manufacturing
325110	Petrochemical Manufacturing
486110	Pipeline Transportation of Crude Oil
486210	Pipeline Transportation of Natural Gas
541360	Geophysical Surveying and Mapping Services

Source: North American Industrial Classification 2002

Table A2 **Gross Value Added of the Petroleum Sector**

Real MEX$ Million 2004 Prices

		Petroleum Sector		Refinery and Petrochemical	
Code	Region	**Total**	**Share /1**	Total	Share /1
R1124	Central and North Chiapas	123,012	87%	22,187	16%
R1129	Tehuantepec Isthmus	32,238	65%	27,075	55%
R1115	Veracruz	3,980	18%	324	1%
R1105	Bravo Bajo Matamoros	22,053	37%	1,449	2%
R1131	Las Huastecas	11,054	27%	2,499	6%
R1044	Valle del Mezquital y Tula	6,765	33%	6,765	33%
R1010	Cd. del Carmen	248,116	97%	2	0%
R1103	La Chontalpa y Cardenas	40,297	90%	1,362	3%
R1124	Central and North Chiapas	100,392	71%	433	0%
R1129	Tehuantepec Isthmus	5,162	10%	2	0%
R1115	Veracruz	3,656	16%	0	0%
R1105	Bravo Bajo Matamoros	19,892	33%	712	1%
R1131	Las Huastecas	8,546	21%	9	0%
R1044	Valle del Mezquital y Tula	0	0%	0	0%
R1010	Cd. del Carmen	247,650	97%	464	0%
R1103	La Chontalpa y Cardenas	38,904	87%	31	0%

/1 GVA of the Sector divided by Industrial GVA

Source: National Economic Census 2004. INEGI

Table A3 **Restricted Sample. Excluded Activities**

Code	Region	NAIC		GVA Region / National	GVA Sector / Region	Employment Region / National	Employment Sector / Region
Sector share 25-30% (N=8)							
R1081	Puebla	336110	Automobile Manufacturing	2.5%	27%	2.8%	3%
Sector share 30-40% (N=11)							
R1129	Tehuantepec Isthmus	324110	Petroleum Refineries	1.6%	38%	1.2%	4%
R1016	Sierra Mojada y Cuatro Cienagas	325920	Explosives Manufacturing	0.0%	30%	0.0%	5%
R1030	Conchos y Ojinaga	315229	Clothes Manufacturing	0.0%	35%	0.0%	13%
R1001	Calvillo	315229	Clothes Manufacturing	0.0%	30%	0.0%	21%
R1085	Queretaro	315221	Underwear and Nightwear Manufacturing	0.0%	34%	0.0%	7%
R1045	Ameca	312142	Tequila Industry	0.1%	37%	0.1%	3%
R1044	Valle del Mezquital y Tula	324110	Petroleum Refineries	0.7%	32%	0.5%	5%
R1044	Valle del Mezquital y Tula	327310	Cement Manufacturing	0.7%	34%	0.5%	2%
R1083	Teziutlan	315229	Clothes Manufacturing	0.1%	36%	0.3%	33%
R1064	Cuernavaca	325412	Pharmaceutical Preparation Manufacturing	0.8%	33%	0.9%	2%
R1079	Izucar de Matamoros	311311	Sugarcane Mills	0.0%	30%	0.1%	5%
Sector share 40-50% (N=4)							
R1013	Parras	313240	Weft Knit Fabric Mills	0.0%	41%	0.0%	11%
R1089	Charcas	336320	Vehicular Lighting Equipment Manufacturing	0.0%	42%	0.1%	20%
R1053	Atlacomulco	311930	Flavoring Syrup and Concentrate Manufacturing	0.2%	40%	0.3%	0%
R1058	Lazaro Cardenas	331111	Iron and Steel Mills	0.2%	43%	0.2%	7%
Sector share 50-60% (N=1)							
R1019	Tecoman	327310	Cement Manufacturing	0.1%	56%	0.1%	0%
Sector share +50% (N=1)							
R1025	Casas Grandes	336320	Vehicular Lighting Equipment Manufacturing	0.4%	91%	0.3%	75%

Source: National Economic Census 2004