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

Duran-Fernandez, R. and G. Santos (2014), 'A Regional Model of Road Accessibility in Mexico: Accessibility Surfaces and Robustness Analysis', *Research in Transportation Economics*, Vol. 46, pp. 55-69. DOI: 10.1016/j.retrec.2014.09.005

A Regional Model of Road Accessibility in Mexico: Accessibility Surfaces and Robustness Analysis

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 introduces an empirical accessibility model for Mexico based on land transport infrastructure. The model assesses an *attraction-accessibility measure* derived from a gravity framework. The measure is estimated on a regional basis and can be interpreted as the market potential of a region. We introduce three versions of the model: the first measures the potential restricted to the domestic market, the second considers the external sector, and the third is essentially a regional accessibility index, which captures the domestic market potential at regional level. We carry out an exhaustive analysis to test the robustness of the accessibility model and find that its behaviour is robust with respect to several criteria.

Key words: accessibility model, Mexico, transport infrastructure, attraction-accessibility measure, gravity model, market potential, regional accessibility index

JEL codes: O18, R12, R40, R49

1 INTRODUCTION

Accessibility is a fundamental concept in the analysis of transport systems. It can be defined as a measure of individual freedom to take part in activities given an environment (Weibull 1980). From this perspective, accessibility can be understood as the potential opportunities for socio-economic interactions attainable by an individual (Hansen 1959). Accessibility has been used in the context of transport studies as a measure of the benefits provided by a transport and land-use system (Ben-Akiva and Lerman, 1979). It has also been used in development studies. Improvements in transport systems and infrastructure lead to an overall reduction in transport costs and travel times. This promotes physical mobility and accessibility (Preston and Raje 2007). Therefore, from a development perspective an improvement in accessibility reduces social exclusion by increasing the capabilities of individuals.

This paper introduces an empirical accessibility model for Mexico based on land transport infrastructure. The model assesses an *attraction-accessibility measure* derived from a gravity framework. The measure is estimated on a regional basis and can be interpreted as the market potential of a region. We introduce three versions of the model: the first one measures the potential restricted to the domestic market, the second one considers the external sector, and the third one measures the domestic market potential at regional level. We also test the robustness of the accessibility model and conclude that it is robust with respect to different specifications, parameters, albeit, within a range, and the variables that measure the activity level of each location. The paper is organised as follows: Section 2 presents a literature review, Section 3 introduces the empirical framework, Section 4 describes the data, Section 5 presents the estimates of the model's parameters and Section 6, a robustness analysis, Section 7 describes the accessibility results for Mexico and Section 8 concludes.

2 LITERATURE REVIEW

Banister and Berechman (2000) propose two possible definitions of accessibility. According to these authors, accessibility can be interpreted as the '*ease of access between spatial opportunities*' (pp. 175), following Weibull's original formulation. The underlying implication of this definition is that accessibility is determined by the time and travel costs associated with a trip to a particular location. An alternative definition they also propose identifies accessibility as '*the potential attainment of a set of transportation choices*' (pp. 175). This concept is based on the assumption that travellers are time-constrained utility-maximising agents that face a number of travel options. There are several approaches to measuring accessibility, depending on the definition that is assumed.

Accessibility is a complex concept, as it is determined by multiple factors. In an extensive literature review, Geurs and Wee (2004) classify the determinants of accessibility into four main components. The first component is a land-use factor, which comprises the demand and supply of opportunities for each destination, given the spatial structure of a land-use system. The second component refers to the subjective judgements, based on individual utility, which determines the amount and type of activities of an agent. The third component is the temporal constraint that individuals face to perform their activities. Finally, the fourth component is transport infrastructure, which includes the structure (connectivity) and geometry of the transport network, as well as the physical attributes (number of lanes and width of roads), and the level of service (quality of infrastructure) of their links.

Geurs and Wee (2004) suggest that under ideal circumstances, an accessibility measure should consider all these components. Nevertheless, the authors recognise that in practice accessibility measures focus only on a subset of these components. They classify the different methodologies to measure accessibility in four major categories, according to the subset of components on which they focus: *infrastructure measures*, *location indexes*, *utility models,* and *individual-based accessibility*. The first two categories correspond to the first definition and the remaining two correspond to the second definition by Banister and Berechman (2000).

2.1 Infrastructure-Based Accessibility

Geurs and Wee (2004) classify infrastructure-based accessibility measures as those purely based on the performance of transport infrastructure. Among these measures, we can find simple indexes such as travelling speeds and congestion levels. The most important shortcoming of these measures is that they fail to incorporate land-use components such as the spatial distribution of destinations. Nevertheless, they have been widely used in transport and urban planning in both Europe and the USA (Ewing 1993, Ypma 2000).

2.2 Location-Based Accessibility

The second category of accessibility measures according to Geurs and Wee (2004) are *location-based* and they include *connectivity measures* and *potential models*.

Connectivity measures are quantities that reflect the distance between an agent and a target. Examples of these measures are the minimum impedance to a destination, and the average distance to a group of facilities. Urban planning and geography research have developed interesting applications based on connectivity. One of these applications is the isochronic measures, which present in a contour map the cumulative opportunities that can be reached within a given travel impedance.

On the other hand, *potential models* are one of the most important tools available in the literature. They estimate the impedance costs[1] of a trip and weigh them using the level of activity of the location that generates it. This allows the estimation of an accessibility measure for any route on the network. It is also possible to aggregate the estimated accessibility for every route to obtain a measure of accessibility for the whole transport network in question (Banister and Berechman 2000).

Potential models can trace their origins to the early social physics tradition. They rely on the assumption that socio-economic interactions depend positively on the level of activity of a particular location, and negatively on impedance costs. The approach is closely related to gravity models, which have been used to analyse a number of socio-economic processes, such as trade, migration, and traffic flows (Bigman and Deichmann 2000).

An *attraction-accessibility measure* (AM) is the functional form of potential models. Weibull (1976) presents an axiomatic framework for the formulation of these measures. His axioms impose general requirements on AM, which ensure consistent mathematical formulations for these models. Any measure that satisfies these axioms is called a *standard attraction-accessibility measure* (SAM). Equation 1 presents the general specification for this measure. In this expression, x_j is the activity level or attraction in zone j and $d_{i,j}$ is the impedance between zones i and j.

Equation 1

$$f_i(d_{i,j}, x_j) = \sum_j g(d_{i,j}) x_j$$

The most important advantage of potential models is that they can be estimated empirically. They incorporate both a transport system and a land-use component. However, strictly speaking, they lack explicit behavioural foundations. Despite this, some theoretical

[1] Impedance costs are defined either as generalised travel costs, or as time costs, or simply as the physical distance between any two locations.

frameworks provide a sort of behavioural justification to potential models. For example, AMs can be interpreted as a generalisation of a gravity model. These models can be derived as a particular case of a utility maximisation problem, under assumptions of monopolistic competition (Anderson 1979; Bergstrand 1985). In the same way, Fujita, Krugman, and Mori (1999) propose a microeconomic model where a market potential measure, based on real wages, determines the location of new production centres in a land-use system.

2.3 Utility-Based Accessibility

Utility accessibility measures, which incorporate subjective judgements of individual's travel choices, are an alternative to potential models. This is the third category of accessibility measures identified by Geurs and Wee (2004). Their objective is to provide a direct measure of the benefits obtained by individuals from a transport and land-use system in terms of utility or a pecuniary base, instead of measuring the level of an *ad hoc* potential interaction.

Geurs and Wee (2004) indicate that utility-based models can be divided in two groups. The first group of models are based on an approach which was originally developed by Domencich and McFadden (1975), Williams (1977), and Ben-Akiva and Lerman (1979). This approach assumes that expected utility is maximised given a random utility choice model. Therefore, utility depends on the opportunities of interaction that the individual faces. These opportunities are given by a fixed choice set. A result of the model is that the marginal expected utility of an opportunity is equal to the choice probability of that opportunity. This condition is exploited to derive an accessibility measure that resembles the general formulation of Weibull's SAM. The most important contribution of this type of model is that it identifies a correspondence between expected maximum utility and consumer's surplus (Miller 1999). This property can be used to conceptually relate changes in accessibility to welfare variations derived from a transport and land-use system.

The second group of models is based on a double constrained entropy model. Entropy models derive three balancing factor measures for the trips generated in a particular zone, the trips attracted to this zone, and trips that take place between any two zones. These measures are derived from the transport-user measure in Williams (1976). The balancing factor treats accessibility as a state vector of the interaction between transport and the land-use system (Martinez 1995). The aim of this type of measures is to provide a strict definition of accessibility in economic terms.

2.4 Individual-Based Accessibility

A limitation of utility models is that they do not consider time and space constrains explicitly. For example, in the expected utility model, activities are drawn from a fixed set, which constitutes only an implicit recognition of time budgets. Geurs and Wee (2004) classify those accessibility measures that incorporate spatial and temporal constrains into a fourth class. These types of models, referred to as *individual-based accessibility*, are based on the space-time geography framework developed by Hägerstrand (1970). This approach pays special attention to the factors that constrain individual choices, analysing the interrelation between activities in space and time. The most important application is the space-time prisms, which is a tool to delimit all the possible locations for individual physical movements in space with respect to time. The shortcoming of this approach is also the lack of a behavioural element.

Utility and individual-based models have been successfully synthesised in the work by Miller (1999). Miller's model derives a standard *attraction-accessibility measure* from the maximisation of individual utility. The model is based on a space-time utility function to model the individual decision process. It comprises a time constraint, which represents the time available for activity participation. In addition, it considers the spatial restrictions of the transport system in terms of travel time. Finally, it takes into account the attraction level of an activity in terms of utility.

Miller's accessibility is an ideal measure that incorporates all the components of accessibility proposed by Geurs and Wee's programme. However, from a practical perspective it is very difficult to apply this methodology to an empirical framework. An empirical version of this model would be highly intensive in data on travel behaviour and time budgets at individual level, a shortcoming shared by utility and person based models in general. In fact, the data necessary to calibrate these models is not typically available in standard travel surveys, not even for developed countries (Thil and Horowitz 1997). The unavailability of data of this complexity explains why empirical research based on these methods has been restricted to relatively small regions and population cohorts.

The application of new technologies, such as geographic positioning systems (GPS), for the collection of more detailed and complete data on the flows of goods and passengers, will improve the availability and detail of data in the future. This will enable the application of the tools mentioned above to an empirical context, which will constitute an improvement on present methods.

Unfortunately, the present state and structure of data hinders that possibility, especially for the case of developing countries, where even standard travel surveys are rare. In these countries, potential market measures constitute by far the best available analytic tool for the analysis of transport systems and infrastructure.

3 EMPIRICAL FRAMEWORK

This section introduces an empirical framework for the estimation of a computable *attraction-accessibility measure* for Mexico based on land transport. Given the availability and structure of the Mexican data, the analysis needs to be done from the location perspective of accessibility, described in Section 2.2 above. We propose the estimation of the classic potential model as an empirical metric for market potential. The classic model is described by Equation 2.

Equation 2

$$A_i(\mathbf{x}, \mathbf{W} \mid \theta) = \sum_{j \neq i} \frac{x_j}{d_{i,j}^{\theta}}$$

Accessibility is defined as a function A_i of vector **x** and matrix **W**, with x_j and $d_{i,j}$ as their respective components, given a parameter θ. The elements of **x** can be interpreted as the level of socio-economic activity taking place in location j. The components of **W** represent an impedance measure between location i and j. Finally, the parameter θ can be interpreted as the elasticity of the market potential of a particular link with respect to impedance.

Accessibility depends positively on the socio-economic size of each node in the transport network, and negatively on the impedance between the nodes. Intuitively, this variable can be interpreted as the market potential of a particular location, or the opportunities for economic interaction that an agent faces. The intuition is clearer from the perspective of a gravity model: accessibility is equal to the sum of all the interaction flows that affect a particular location, given its relative size. Equation 3, where $T_{i,j}$ represents the interaction flows between i and j, illustrates this point. These flows can be interpreted directly as traffic through the road network so that the decay parameter θ is the elasticity of this variable with respect to impedance. Finally, we focus on the industrial sector in Mexico. Therefore, the study concentrates on freight only.

Equation 3 $$\sum_j T_{i,j} = x_i A_i(\mathbf{x}, \mathbf{W} \mid \theta)$$

<u>**3.1 Travel Times as Impedance Measure**</u>

We propose a time-based impedance measure as an approximation to the generalised cost, so that **W** is the set of optimal solutions that minimise the travel time between any two locations.[2]

Ideally, a generalised-cost measure would include a time-based component and a distance-based component, which would capture the cost of fuel and vehicle use, tolls[3], and any other distance-related cost. In the case of the Mexican road network, there are no readily available detailed data on tolls for the different sections of the network or on the value of time. The value of time for this exercise would need to be estimated based on average wages in the transport sector, available from the National Economic Census (NEC), combined with data on the volume and type of freight moved on each route, not available in any of the Mexican government statistics. Therefore, the use of travel times as impedance measure is a good compromise.

Table 1 presents the cost structure of the freight industry in Mexico for 2004. These figures show that tolls represent five percent of the total running costs of the sector respectively. This magnitude suggests that, as long as the costs are uniformly distributed across the domestic routes, their exclusion will not significantly modify route choice.[4] Moreover, recent studies on travel demand and travel costs show that the estimated elasticity of traffic respect to variable transport costs in Mexico is quite low[5] (Crotte *et. al.* 2008). Despite Crotte *et. al.* (2008) not

[2] This impedance measure depends on the structure and geometry of the transport network, as well as on the physical attributes and level of service of their roads. It acknowledges that improvement of transport infrastructure leads to a reduction of travel time and cost, and a subsequent improvement of the accessibility to markets (Rietveld 1989).
[3] Tolls may be related to distance driven or they may not.
[4] The uniformity assumption seems to be valid for the case of insurance cost. The price of insurance policies offered in Mexico depends solely on the type of vehicle and on whether the freight in question is classified as a dangerous product. This indicates that insurance costs do not vary significantly across routes.
[5] The authors estimate the elasticity of *private traffic*, defined as *km* travelled by private cars, with respect to petrol price at -0.156.

explicitly considering tolls, their result strongly suggests that minor increments in travel cost do not affect travel behaviour in an important way. Under these assumptions, optimal routes[6] estimated under time-based impedance will be essentially the same as those based on generalised-cost measures. Consequently, a time-based accessibility model would behave in a similar way to a more detailed model based on complete generalised costs.

3.2 Domestic and International Market Potential

Total market potential has two differentiated components. The first is the potential for domestic markets. This element can be characterised by an accessibility model restricted to a subset of elements i, j, such that they are geographically located within the national territory. The use of land transport to model accessibility to domestic markets in Mexico is justified by the importance of this mode of transport relative to other alternatives. In fact, 85 percent of all domestic freight in Mexico relies on land transport (North American Transportation Statistics, NATS 2006).

The second component of total accessibility is the potential for international markets. This element is particularly important for Mexico. After joining the North American Free Trade Agreement (NAFTA) in 1994, the country transformed itself into an export-oriented economy. Nowadays, the external sector[7] represents 60 percent of the gross domestic product (GDP). Almost 70 percent of these trade flows take place with the USA, by far Mexico's largest trade partner. Mexico and the USA share a border of more than 3,000 km. Therefore, it is not surprising that road transport ranks as the most important mode of international transport, with 71 percent of the international trade between both countries taking place by road.

The importance of international freight indicates that international market potential must be explicitly considered in the accessibility model. The characterisation of international

[6] An optimal route is defined as the route that minimises travel times between any two points i and j.

[7] It includes the total value of exports and imports.

accessibility follows the same function presented in Equation 2. However, for this case the model is restricted to a subset of elements $\{i, j\}$ such that i is located in Mexico and j in the USA. The international accessibility function depends on a vector **x,** which represents the weight of international location j and is equivalent to the vector **x** used to estimate domestic accessibility. The components of **W** also represent an impedance measure between domestic and international locations i and j.

As in the estimation of domestic accessibility, the model uses a time-based impedance measure. The limitations of this approach, which have already been discussed for the domestic case, also apply to the estimation of this index. Despite the fact that there is information for the USA that would enable us to estimate generalised costs, it would still be impossible to do so for Mexico. Moreover, due to the fact that the model needs to work on the same type of parameters, we need to stick to the time-based impedance for the USA case as well. [8]

Finally, an additional fixed component of impedance, which can be incorporated into the model, is the cost of crossing an international border. Detailed data about average crossing delays through the most important international land gateways is publicly available. This data can be directly used, without the need to make any further assumptions, enriching the accuracy of the accessibility model. This approach has been previously used in accessibility models where border delays used to play an important role such as, for example, in pre-Maastricht Europe (Keeble and Walkers 1988; Fürst and Schürmann 2000; Schürmann and Talaat 2000).

[8] The exclusion of tolls costs in the USA road network is likely to have negligible effects on the estimated accessibility: the average cost per kilometre for lorries is only USD 0.75 and the total length of all the tolled motorways in the 48 contiguous states represents less that 0.1 percent of the extension of the network in the country (Bureau of Transportation Statistics, BTS 2008). Moreover, the cost structure of the sector and the fact that high productivity levels in the USA imply higher values of time indicate that the time component of the generalised-cost dominates other costs.

3.3 Geo-Statistical Framework

In an ideal situation, accessibility would be modelled for all the individual settlements, defined as villages, towns, cities, or metropolitan areas. However, this would be an incredibly challenging enterprise from a computational perspective.[9] Therefore, the selection of a suitable geo-statistical framework for the estimation of the accessibility model requires a compromise between computing capability and geographic detail.

An alternative is the use of a geo-statistical framework based on a low tier of government: for example the *municipio* in Mexico and the county in the USA. The problem of this approach is that most of the time these geopolitical units comprise an arbitrary territory, which does not reflect real socio-economic interactions. For example, the metropolitan area of Mexico City is formed by at least 22 municipalities, which would be treated as independent geo-statistical units. Individually, they do not necessarily capture the magnitude of the metropolitan area that they constitute.

In the European context, a common approach is the use of regional data based on the third level of the Nomenclature of Territorial Units for Statistics (NUTS). For some countries, the equivalent NUTS3 is a defined political or administrative unit but this is not always the case. A NUTS3 approach in the North American context would rely on the definition of economic regions, not associated to any tier of government. For the case of Mexico, Bassols-Batalla (1993, 2002) has proposed a criterion for the regionalisation of its territory in 135 functional regions.[10] For the case of the USA, the Bureau of Economic Analysis (BEA) has identified 177 economic regions in the 48 contiguous states.[11]

[9] Mexico has 4,028 urban settlements and the United Sates, 34,611.This would require the estimation of a 38,639 x 38,639 origin-destination matrix. Moreover, with the exception of population, there is no publicly available socio-economic data at this aggregation level for either country.

[10] Bassols-Batalla (1993, 2002) defines a region as a geographical area characterised by a particular economic structure, which exhibits a high degree of homogeneity, and has developed particular internal and external links in comparison to other regions.

[11] The BEA defines a region as a group of counties within a labour market in which the proportion of resident workers who commute to a given central county exceeds the proportion who commute to alternative central counties (Berry el. al. 1968; Johnson *et. al.* 2004). This methodology leads to a list of economically and socially integrated geographical units, in the same spirit as Bassols-Batalla's regionalisation of Mexico.

3.4 Regional Accessibility

A functional region describes the relationship between a hinterland and a central or nodal city. Therefore, a regional model of accessibility is useful for the analysis of the socio-economic interactions between these hinterlands. However, it fails to identify the regional interactions taking place in it. In order to provide a more complete picture of Mexico's market potential and its dependence on road infrastructure, we estimate a variant of the accessibility function to characterise the domestic market potential for each region in Mexico.

Equation 4 defines the accessibility between any two locations *i* and *j* located in region *r*. The regional accessibility is defined as the weighted average accessibility within the region, where the elements ω of matrix **Ω** are taken as weights. Just as the standard accessibility model, this function depends on the structure of the road network in a particular region and on its *Road Type* (RT).

Equation 4

$$A_r^*(\mathbf{x}, \mathbf{W}, \mathbf{\Omega} \mid \theta) = \sum_{i \in R} \sum_{j \neq i} \omega_i \frac{x_j}{d_{i,j}^{\theta}}$$

4 DATA DESCRIPTION

The estimation of the accessibility index, as presented in Section 3 above, requires two groups of variables. It requires, first, a vector of socio-economic variables **x,** which captures the weight of location *j* as source and destination of socio-economic interaction, and second, the estimation of the components of the impedance matrix **W**, characterised as the typical time required to travel between any pair of locations $\{i, j\}$.

4.1 Socio-economic Variables

In order to assess the robustness of the model, several versions of the accessibility function are estimated using different definitions for vector **x**. In particular, we centre our attention on five variables: population, non-farm workforce, manufacturing workforce, non-farm personal income, and total wages and salaries. In principle, these variables should be proportional to the size of the regional markets that they represent, making them good proxies for the estimation of market potential. However, there are some potential sources of divergence between the different versions of accessibility. In particular, geographic inequalities in income distribution could be reflected in systematic biases between an accessibility measure based on pecuniary variables and a measure based on demographic indicators.

For Mexico, the demographic data was taken from the 2000 National Population Census (INEGI 2000c). For the USA, the source of this data was the 2000 USA Census. Data about income and wages for Mexico was taken from the NEC 2003. For the USA, that data was taken from the Bureau of Economic Analysis (BEA) 2001.[12] A summary of the data is presented in Table 2.

4.2 Geographic Information System Variables

The optimal travel times in the accessibility model were estimated under the North American Geographic Information System (GIS) Road Network Model (Duran Fernandez 2014).[13] For the domestic accessibility model, we estimate the minimum travel time between any two of the 135 regions of the country. For each region, the urban settlement with the highest population was defined as a node on the road network. For most of the regions, this settlement was typically the closest one to its centroid.[14] For those regions where two or more settlements could be ranked as the one with the highest population, we selected as the node the one closest to the centroid. For the international accessibility model, we estimate the minimum travel time between any of the 135 Mexican regions, and any of the 177 economic regions of the 48 contiguous states in the USA. For the USA the city with the largest population within a region was defined as a node of the network.

[12] All the pecuniary variables are expressed in USD using the February 2008 exchange rate published by the Central Bank of Mexico.
[13] We use the Network Analyst utility of ArcMap 9.1 The algorithm used by this utility finds the minimum cost between any two nodes, considering the length and average speed of each section on the network.
[14]Centroid is defined as the geographical centre of a region.

The regional accessibility index requires the estimation of matrix **W** for all the urban settlements located in a particular region. The geographic coordinates of each of the 4,028 urban settlements in Mexico were taken from the Integrated Geo-statistical System Browser of Mexico. Then, for each region, we estimated an origin-destination (OD) matrix of minimum times, which considers as nodes all the urban settlements within their boundaries. This is also based on the North American GIS Road Network Model.

5 ESTIMATION

A gravity model can be interpreted as a generalisation of a potential model. In a gravity model framework, we estimate the interaction between two specific locations, while in a potential model approach the measure is extended to all possible or relevant destinations. Under this interpretation, we can assume that the impedance decay parameter θ in the accessibility model is equivalent to the elasticity of the interaction flows with respect to the impedance costs of a simple gravity model (Fotheringham 1981). This assumption is extremely useful because it allows the estimation of the parameter from actual data on relevant interaction patterns (Bigman and Deichmann 2000). The alternative to this empirical approach is the calibration of this value in function of *ad hoc* properties that the accessibility model should fulfil. Some examples of this approach for the European Union include Fürst *et. al.* (2000) and Schürman *et. al.* (2000). Deichmann *et. al.* (2004) uses this approach for the Mexican case.

The accessibility model presented in this section is estimated following a combined approach. First, we use data from the Origin and Destination of Passengers and Freight Survey in Mexico (Institute of Statistics, Geography, and Information, INEGI 1999) to estimate an inter-state gravity model. The gravity model is estimated under several variations of the original specification. These results are used to identify an interval for the value of the decay parameter. In a second step, the accessibility model is calculated for several values of the parameter within this interval. The exercise allows the study of the statistical properties of the accessibility model and an assessment of its robustness by comparing the behaviour of its different versions.

5.1 Interstate Gravity Model for Freight Flows

The gravity equation is an analytic tool widely used for modelling bilateral flows between different geographic entities. Equation 5 presents a general version of the gravity model. In this expression $T_{i,j}$ is the flow from origin i to destination j, K is a constant, M_i and M_j are relevant socio-economic variables of the two locations –also known as economic masses, given their resemblance with mass in Newton's Gravity law-, and $f(\cdot)$ is a function that depends negatively on transport costs $d_{i,j}$ between the two locations. The classic version of the model considers for this function $d_{i,j}$ raised to a power θ, where $\theta<0$.

Equation 5

$$T_{i,j} = K\ M_i M_j f(d_{i,j} \mid \theta)$$

The estimation of Equation 5 is straightforward using different econometric techniques. It can be transformed in a linear model taking logarithms on both sides of the equation. However, Anderson and Wincoop (2003) argue that this specification is not correct due to the fact that it does not take into account multilateral resistance terms. The solution proposed by the authors is to explicitly consider importer and exporter fixed effects. Another modification is the inclusion of a remoteness index that measures the average distance of a region from all trading partners. The final specification is presented in Equation 6, where fe and r represent a geographic effect specific to j, and the remoteness index respectively, and ε is a stochastic variable. The empirical literature has used different socio-economic definitions as economic variables M, such as gross production, gross value added, population, and workforce among others. Travel costs are usually measured as the physical impedance between locations i and j.

Equation 6

$$\ln(T_{i,j}) = \gamma_1 \ln(M_i) + \gamma_2 \ln(M_j) - \theta \ln(d_{i,j}) + \beta_1 fe_i + \beta_2 fe_j + \beta_3 r_{i,j} + K + \varepsilon_{i,j}$$

The empirical literature from international trade presents several examples of estimation of gravity models. The objective of these studies has been centred on the analysis of the determinants of trade, so that the impedance variable plays only a secondary role as a control in the estimation. In general, the mean elasticity of impedance with respect to trade flows has been estimated at 0.9 with a 90 percent of estimates lying between -0.28 and -1.55. These results have been estimated using meta-analysis techniques on an exhaustive survey of existing literature (Disdier 2006). In North America this elasticity has been estimated at -1.52 for Canada (McCallum 1995) and -0.77 for the USA (Wolf 2000), using provincial and state trade flows respectively. In Mexico, this type of assessment has only been carried out using data on international trade flows. The estimated elasticities lie between -0.9 and -1.4 (Lopez-Cordova 2002; Montenegro 2006; Soloaga 1996).

5.2 Estimation of Interstate Freight Flows

The dataset used in this estimation is a module of the NEC 1999. For each of the 31 states and the Federal District the database contains the lorry freight flow for their eleven main domestic destinations. The rest of the freight flow is aggregated in a single category labelled as *others*. On average, this flow represents only 14 percent of state freight. For each state, the freight label as *others* is equally allocated among the rest of the states. Freight flows S are used to build an interstate OD matrix. For each possible origin and destination combination $\{i, j\}$ we define total freight flow T as $T_{i,j} = S_{i,j} + S_{j,i}$. The OD matrix is symmetric under this definition; therefore, the analysis uses only the entries of the upper triangle to avoid including repeated observations.

The remoteness index is included as a geographic control, following the work of Anderson-Wincoop (2003). It is calculated as described by Equation 7. In this expression N_j is the population of state j, $d_{i,j}$ is the travel time between state i and state j –defined as the weighted average of the travel time between the major metropolitan areas- and A is a standardisation.

Equation 7

$$r_i = \frac{1}{A}\sum_j \frac{N_j}{d_{i,j}}$$

In order to assess the robustness of the model, different estimations were carried out using each of the following economic variables: total population, workforce (total, industrial, services, and both), and gross state product (total, industrial, services, and both). The estimation of the model shows that the estimated elasticity of distance with respect to freight flow is remarkably stable independently of the chosen economic variables. As impedance measure, we consider the weighted average travel time between the major metropolitan areas of a state (Table 3).

5.3 Results and Selection of Canonical Parameters

The model is estimated using ordinary least squares (OLS) with robust standard errors. To explore the robustness of the model, we estimate it using nine different definitions for the socio-economic variables M_i and M_j: total gross product, gross production of the secondary sector, gross production of the tertiary sector, non-farm gross production, total population, total workforce, workforce of the secondary sector, workforce of the tertiary sector, and non-farm workforce.

Table 4 presents the results of the OLS estimation of the decay parameter of the gravity model. The first four columns (under the heading '*canonical*') present the coefficients, standard errors, and the adjusted R^2 for the model that includes state controls. The last four columns (under the heading '*no dummies*') present the results of the same model but omitting dummy variables to control for states.

In the '*canonical*' estimates column there are two sets of results: the first does not assume any restriction on the coefficient of the socio-economic variables M_i and M_j (*i.e.* $\gamma_1 \neq \gamma_2$), while the second assumes that they are equal (*i.e.* $\gamma_1 = \gamma_2$). It is important to note that the estimated parameter is remarkably stable across the different definitions for the economic variables. The results indicate that state specific effects work as a good control for any omitted variable, generating an unbiased estimation of the parameters of the model. In the '*no dummies*' estimates column, we also present two sets of results that correspond to the unrestricted and restricted models respectively. These '*no dummies'* results exhibit more variations in comparison to the model that includes the state controls. However, the estimation continues to be stable. In fact the average elasticity estimated with the model that excludes the dummies for the states (-0.8) is very close to that estimated with the model that includes them (-0.7).

Finally, the adjusted R^2 statistic indicates that explanatory power of the '*canonical'* models is larger than any of the '*no dummies'* model. This property indicates that the decay parameter estimated under the '*canonical'* model would predict the actual interstate freight flows in the dataset more accurately. The stability of the decay parameters of the '*canonical'* model as well as the higher explanatory power that it provides, suggest that the extrapolation of this statistic into the accessibility model, represent a superior choice in comparison to other options.[15]

5.4 Poisson Regression

A criticism of the Anderson-Wincoop methodology, which is followed in this exercise, concerns the probability distribution that is assumed for the empirical version of the gravity model. Santos and Tenreyro (2006) argue that due to Jensen's inequality, an OLS estimation of the logarithm of $T_{i,j}$ would lead to biased estimates. As a solution to this problem, the authors apply a Poisson model, using pseudo-maximum likelihood estimation. Here, we follow Santos and Tenreyro (2006) and carry out a Poisson analysis of the gravity model.

[15] The original freight flows dataset contains information for only the main freight origin and destination for each state. As mentioned above, the volume of freight labelled as '*other*' represents, on average, only 14 percent of the state freight. This volume was equally allocated to the states with missing data. The OLS regressions reported in Table 3 were run under those assumptions. In order to explore how this methodology affects the results we re-estimated the model using only the hard-data of the original freight flows dataset. We found very similar results to those reported in Table 3: a stable decay parameter with respect to all the different definitions for the socio-economic variables, and very close average values for the parameter computed using the '*canonical*' and '*no dummies*' models: -0.86 and -0.87, respectively.

An important problem in the estimation of the Poisson model is that it drops all the observations with negative values. An alternative solution is to perform a linear transformation on $T_{i,j}$ such that the values of the explanatory variables are positive. Unfortunately, the results of the estimation are not always neutral to the linear transformation. Therefore, we estimated two versions of the model. The first drops all the negative observations while the second uses as explanatory variable the linear transformation ln[100$T_{i,j}$+1].

The elasticity estimated using this methodology is roughly half the one estimated under the OLS regression (-0.3). In spite of this difference, the general patterns described in the OLS analysis are kept: the parameter is remarkably stable across the nine definitions of socio-economic variables M_i and M_j, and on average, the elasticity for the hard-data estimation is slightly higher in absolute value, and the variability of the estimated parameter is higher when the state controls are omitted.

In summary, the estimation of the absolute value of the decay parameter under the OLS regression was always between 0.67 and 1.03, while the absolute value of the Poisson parameter took values between 0.22 and 0.35. The difference between the decay parameters estimated under the OLS and Poisson models, as argued by Santos and Tenreyro (2006), could be the result of bias due to an incorrect assumption on the probability distribution. However, we note that this is not the case under an assumption of log-normality of the explanatory variable $T_{i,j}$. Under this assumption, the Santos-Tenreyro bias does not arise. In this case, the difference between the two models can be easily attributed to the linear transformation of the data, the dropping of observations, and the additional restrictions on the specification of the model, rather than to an estimation bias.

The probability distribution of the logarithm of $T_{i,j}$ has a minor negative bias (0.6) and its kurtosis is slightly higher than that expected for a normal distribution (7). However, this behaviour can be attributed to measurement errors and does not rule out the log-normality assumption. On the other hand, under the Poisson assumption the variance of $T_{i,j}$ should be equal to its mean, a condition not fulfilled by the data. Our data does not conclusively support either the log-normality or the Poisson distribution. However, it is worth noting that the histogram of both the logarithm of the observed and estimated explanatory variable $T_{i,j}$ is closer to a normal distribution than to a Poisson process, as shown by Figures 1A and 1B. Also, it is clear that the log-normality assumption does not impose more restrictions than the Poisson one. Therefore, we will assume that the canonical value of the decay parameter is the one estimated under the OLS regression, a value that is close to the results obtained by Wolf (2000) for the USA. Another argument in favour of the OLS regression is that the 99 percent confidence interval for the decay parameter estimated with the OLS regression is always negative, generating coherent estimators from a theoretical perspective.

Finally, it is worth mentioning that most of the literature on gravity surveyed by Disdier (2006) estimate the absolute value of this parameter above 0.5 (in fact, none of the studies for North America estimates the parameter below this threshold). This result favours the selection of the OLS model as the canonical case. In any case, the analysis of the accessibility model presented in Section 6 shows that differences in the estimates of the parameter in the range found between the OLS and the Poisson model are not translated into divergences in the behaviour of the corresponding accessibility model.

5.5 Extrapolation to the Accessibility Model

We have already presented several arguments in favour of choosing the OLS estimate that includes state controls as canonical or base value. Among those arguments, we recall the stability of the estimation with state controls, and the similarity of the decay parameter with the results of previous literature carried out in other North American countries. This value is -0.73 and it is assumed to follow a normal distribution with an estimated standard deviation of 0.11. The use of this value as an input in the accessibility model faces an important consideration about the convenience of extrapolating results, based on interstate freight flows, to an interregional context. The problem arises with the implicit assumption that the decay parameter is independent from the geographical scale. Unfortunately, the interstate freight flow database used in the gravity model is the only source that allows this kind of empirical exercise. Therefore, it is not possible to test empirically if the decay parameter is sensitive to geographical scale.

Nevertheless, there are two arguments in favour of the extrapolation of these results to the regional accessibility model. The first is that the literature for Mexico shows that the estimated decay parameter in an international trade context is not considerably different from our estimate; as explained in Section 5.1, the elasticities lie between -0.9 and -1.4 (Lopez-Cordova 2002; Montenegro 2006; Soloaga 1996). The other argument is that for most of the states in Mexico the economic activity is concentrated in a few metropolitan areas. Therefore, the interstate freight reported in the Origin and Destination Freight Survey must be driven by actual intercity flows. Although the structure of data does not allow us to test this claim the notion that the estimated state parameter is close to the actual interregional value is not counterintuitive.

Finally, a word should be said about the assumption that the decay parameter is constant across the country. Fotherigham (1981) shows that the calibration of the decay parameter for pairs of cities in the USA presents an important degree of variation. From the perspective of the econometric estimation of the model, these differences are absorbed by the residual of the regression and are assumed to be independently distributed. However, Fotherigham suggests that this is not the case, and they depend in fact on the interurban spatial structure. If this were true, the gravity approach would lead to inconsistent estimates for accessibility. Nevertheless, this is a general criticism to the potential approach to accessibility rather than a specific problem with the present study. The literature proposes some alternative approaches to surpass this issue. One of the most promising methodologies would be the application of micro-behaviour accessibility models, such as the one presented by Miller (2000). However, the application of these types of models is not possible given the present state and structure of data in Mexico.

6 ROBUSTNESS ANALYSES

The results of the interstate gravity model in Section 5 show that the absolute value of the decay parameter lies between 0.67 and 1.03, if it is estimated using an OLS regression, or between 0.22 and 0.33, when a Poisson model is used instead. The different values of this estimation depend on the socio-economic variables M_i, the restrictions imposed on the model, and on the inclusion of state controls.

In this section, we analyse how variations in the decay parameter are translated into the accessibility model. The methodology consists in the estimation of two versions of the model $A_i(\mathbf{x}, \mathbf{W}, \mid \theta)$, evaluated for two different values of the parameter θ. Then we estimate the linear correlation between the two versions of the model to determine if their behaviour experiences any significant change.

The accessibility model is essentially non-dimensional. For this reason, the effect of the parameters on its absolute magnitude is not important (Bigman and Deichmann 2000). In fact, the relevant information that the model provides depends only on the values that it takes for each observation (in this case, a region), relative to the rest of the system. If the two versions of the model present a high and positive correlation, then the relative value of accessibility for every region with respect to the system is not significantly sensitive to the parameters.

Equation 9 summarises the statistic that is assessed in this exercise. Function Ψ is defined as the sample correlation coefficient between two different versions of the model, where i represents a region. The parameters are chosen from a set Θ such that all its elements are in the interval (0, 1.5]. All the possible values for the decay parameter found in the gravity estimation are located in this interval.

Equation 9 [16]

$$\Psi(\mathbf{x},\mathbf{x}',\mathbf{W},\hat{\theta},\hat{\theta}') = \frac{\sum_i A_i(\mathbf{x},\mathbf{W}\mid\hat{\theta})A_i(\mathbf{x}',\mathbf{W}\mid\hat{\theta}') - n\overline{A}\cdot\overline{A}'}{(n-1)s_A s_{A'}}$$

$$\hat{\theta},\hat{\theta}' \in \Theta$$

Initially we estimate Ψ for **x=x′**, defining this vector as population. We select 100 random values for $\hat{\theta}$ and $\hat{\theta}'$ (*i.e.* the decay parameters of models A and A' respectively) from the set Θ, such that Ψ is estimated at 10,000 different points. The exercise is carried out separately for the domestic and the international versions of the model.

The results are summarised in Figure 2A and 2B. On these graphs, the darkest area corresponds to high correlation values where Ψ is close to one. For the domestic accessibility model (Figure 2A), the correlation Ψ is never below 0.7 in the interval Θ, and the larger the difference between the two parameters, the lower the correlation is. It is worth recalling that according to the OLS gravity model, the decay parameter presented values between 0.67 and 1.03. For that interval the correlation is always higher than 0.99. Also, the correlation Ψ in the region where the mean values of the Poisson and the OLS parameters intercept is always higher than 0.95.[17] This takes place for both the domestic and the international versions of the accessibility model. These results show that variability in the decay parameter of the model within these limits does not affect the overall behaviour of the model at regional level.

Figure 2B presents the results of the international version the model. For this case, the estimated correlation Ψ is lower in comparison to the first exercise. This result is not surprising because the typical value of $d_{i,j}$ for the international optimal routes is larger than in

---

[16] In Equation 9, $\overline{A}$, $\overline{A}'$ are the means of *A* and *A'*, respectively. *A'* is the accessibility model evaluated in ***x'***. Finally, $\mathbf{s}_t$ is defined as the standard deviation of variable ***t*** and ***n*** is the number geographic units (regions).

[17] This region is defined by the interval $0.67 < \hat{\theta},\hat{\theta}' \leq 1.0$.

the domestic model. The estimated average value of the surface Ψ in the region $0.67 < \hat{\theta}, \hat{\theta}' \leq 1.0$ is 0.96, with values lying between 0.8 and 1.0. All the estimated values of the decay parameter under the OLS gravity model can be found in this region, as well as the estimates of the same parameter for the USA, which, as mentioned in Section 5.1, Wolf (2000) estimates at 0.77. The values taken by Ψ in the intersection of the OLS and the Poisson intervals are considerably lower, with an estimated average of 0.77 and values as low as 0.55. Nevertheless, given the results of previous literature, it is unlikely that the true value of the decay parameter will be as low as in the Poisson interval, especially for international optimal routes.

Finally, we perform a similar analysis at the regional model. The results of this exercise show that the correlation Ψ is always higher than 0.9 for values of the decay parameter located in the interval Θ. In comparison to the domestic and the international versions of the model, regional accessibility presents the most stable behaviour with respect to θ.

We can extend this methodology to explore the sensitivity of the model with respect to vector **x.** We consider five possible definitions for this vector: population, non-farm workforce, manufacturing workforce, gross value added, and total wages. These variables can be built at regional level for both Mexico and the USA. The exercise can be applied to the domestic and the international version of the model. Unfortunately, it cannot be extended to the regional model because **x** can be represented only by demographic variables, due to the unavailability of other indicators at town level.

For each of the domestic and international models there are ten possible permutations for the correlation Ψ depending on vector **x**. In each permutation, we select 1,000 values of θ from the set Θ to evaluate the correlation in $\Psi(\mathbf{x}, \mathbf{x}', \mathbf{W}, \hat{\theta}, \hat{\theta})$. This statistic provides information on the stability of the accessibility model with respect to the definition of **x**, given a value for $\hat{\theta}$.

The results of this exercise show that for every permutation (**x, x`)** the correlation between the resulting accessibility models is always higher than 0.9. In general, the correlation between the models based on demographic and monetary variables is slightly lower when compared to the other cases. Nevertheless, this value is high in absolute terms: for example, the correlation between the population and the non-farm gross value added model is close to one for the relevant interval of θ. The results of this exercise show that the behaviour of the accessibility model is remarkably stable, independently of the type of variable **x** that is used to assess its value.

A final concern about the robustness of the model is its functional specification. Several functions can be used as valid *attraction-accessibility measures* and there is no *a priori* reason to select one over another. It would be impossible to test the robustness of the classic model presented in this study in comparison to all other possible alternatives. Therefore, we limit ourselves to the estimation of the decay exponential model, often cited in the literature. Equation 10 presents the functional form of this alternative.

Equation 10

$$\widetilde{A}_i(\mathbf{x}, \mathbf{W} \mid \varphi) = \sum_j x_j \exp(-\varphi d_{i,j})$$

Mathematically, this function shares many of its properties with the classic model. However, from a conceptual point of view it presents a significant difference with the classic model: the elasticity of the individual potential between i and j is proportional to impedances, while in the classic model this value is assumed to be constant.

We estimate an empirical value for $\widetilde{A}_i$ to calculate the correlation of this specification with the classic model. To be consistent with this specification, the decay parameter of $\widetilde{A}_i$ is estimated from a version of the canonical gravity model that considers $d_{i,j}$ instead of $\log(d_{i,j})$, as explanatory variable of the logarithm of traffic flows. The estimated correlation between $\widetilde{A}_i$ and A_i is 0.91 and 0.90 for the national and the international versions. The result holds for values of the decay parameters θ, φ in a range between 0 and 1.5: the estimated correlations are always above 0.9. This indicates that given a spatial structure (**x, W**) the behaviour of the model is not sensitive to a particular functional form.

The results of the exercises presented in this section show that the behaviour modelled by the classic model is robust with respect to variations in the decay parameter θ, different definitions of vector **x**, and the functional specification of the model.

7 MEXICO'S ACCESSIBILITY GEOGRAPHY

This section presents a general description of the spatial structure of accessibility estimated under the domestic, international, and regional versions of the model. Table 5 presents a summary of the variables used and Figures 3, 4, and 5, their geographic distribution. All the figures presented in this section are estimated using the parameters of the canonical version of the model.

7.1 Domestic Accessibility

The regions in the central and western-central areas of Mexico (macroregions IV and V in Table 5) present the highest domestic accessibility levels in the country. This pattern is explained by the fact that the two most important regions are located in these areas (Mexico Basin and R147): these regions present a high population, the road network is dense, and it is composed of high quality links. Another area that presents an accessibility level above the national average is the eastern part of the country (macroregion VII in Table 5). This area is geographically close to the highly populated regions of central Mexico.

The southern regions (macroregion VI in Table 5) present an accessibility level close to the national average. Two important motorways link Mexico City and the southern states of Guerrero and Oaxaca with a RT equivalent to A4 and A3, respectively.[18] The presence of these links ensures that the overall accessibility of the regions in the area is not below the national average. However, there are some important exceptions. For example, the states of Chiapas, in the southernmost part of the country, the coast of Oaxaca, and the mountains of Guerrero, present some of the lowest accessibility levels.

The northeast and north central regions also present an accessibility level similar to the national average. These regions are characterised by higher population dispersion and lower road densities. In general, central Mexico has more and larger settlements than these areas. However, some important cities, such as the metropolitan area of Monterrey and the border city of Juarez (the third and fourth most important cities in the country by population), are located in this area. Both the northeast and the north central regions are well connected to central Mexico by high quality roads with a RT equivalent to A4.

The regions with the lowest accessibility are located in the northwestern part of Mexico and the Yucatan Peninsula (macroregions II and VIII in Table 5). The low accessibility levels of the California and the Yucatan Peninsula are mainly driven by the peculiarities of Mexican geography. Besides, the regions are linked to the rest of the country by limited and low quality roads. The effect of a poor road endowment is illustrated by the northwestern regions located on the east margin of the Gulf of California. These regions are practically isolated from the rest of the country by the mountain of Sierra Madre. In fact, there are no direct west-east roads to connect this area with the industrial cities of the northeast. This explains that some of the regions with the lowest accessibility levels are located in this zone.

[18] The classification of roads by *Road Type* (RT) follows the *Statistical Handbook of Transport* published by the Mexican Institute of Transport (Mexican Institute of Transport, IMT 2000). A4: 4-lane paved road, 22 meters width; A2: 2-lane paved road, eleven meters width; B2: 2-lane paved road, nine meters width; C: 2-lane paved road, seven meters width.

7.2 International Accessibility[19]

In general, the geographical distribution of international accessibility can be explained by the proximity of a region to the USA. However, there are important peculiarities that can be attributed to infrastructure endowments. The regions with the highest international accessibility are located in the northeast part of Mexico (macroregion I in Table 5), an industrial area close to the state of Texas and the most important gateway for the Mideast. Accessibility is evenly distributed in this area, and it does not drop with distance in comparison to the other border regions. The northwest of Mexico has a lower accessibility in comparison to the rest of the border areas. This is explained by the relatively longer distance between this area and the Great Lakes and the Mideast.

The north-central regions also present an international accessibility above the national average. This area has a unique location: it is in the intersection of the most important international trade corridor. Despite its unique location, the regions in this area depend predominantly on one single road corridor and the secondary network is not well developed. The further a region is from the international border the highest is its dependence on poor national roads and as a result, the lower is its accessibility. In fact, accessibility drops at a higher rate in this area in comparison to the northeast, illustrating the poor quality of regional roads.

In the central area of the country, international accessibility is evenly distributed. However, there are important and uneven variations in the south. For example, the state of Guerrero presents accessibility levels that are comparable to regions located 200 km closer to the border. In the State of Oaxaca, which is on the same latitude as Guerrero, accessibility drops significantly. These differences can be solely attributed to the structure of the road network

[19] This definition does not consider the market potential of other trade partners such as Canada and Central America. Strictly speaking, it is a partial measure of international accessibility. Nevertheless, it is worth recalling that the USA is the most important trade partner that Mexico has. This is reflected by the relative importance of the freight flows between the two countries, in comparison to the rest of international freight flows. For example, the total value of land freight between Mexico and Canada represent less than 2.5 percent of the value of the USA-Mexico freight. Therefore, this measure is a good approximation of Mexico's actual international market potential.

and the RT of the roads serving as international links. Finally, the regions with the lowest international accessibility are located in the Yucatan Peninsula.

7.3 Regional Accessibility

Regional accessibility measures the market potential within a particular region, as a function of the spatial structure of the urban settlements, the domestic structure of the regional roads, and the RT. The index is slightly correlated with total population and population density, with coefficients of correlation of 0.45 and 0.33, respectively. Population dispersion also plays a significant role: the number of settlements in the region has a correlation of 0.34 with the index. These properties make the model take larger values in high density regions, such as those with large metropolitan urban areas (Mexico City, Guadalajara and Monterrey). However, low density regions with high population dispersion also present high values for this variable (South Guanajuato and Central Oaxaca).

8 FINAL REMARKS

This paper introduced an empirical metric for accessibility in Mexico, based on a classic potential model. This variable depends on the spatial distribution of socio-economic activities in Mexico and on the provision of transport infrastructure. The first component captures the effect of the land-use system in Mexico, which in the model is taken as given. Transport infrastructure is characterised in terms of the structure of the transport network and its RT. The accessibility model can be indirectly interpreted as a proxy for welfare -given that it is not expressed in monetary units- under the theoretical results reviewed in the Section 2. Moreover, it can be used as a proxy to the value of road infrastructure, under the assumption that producers are able to translate variations in accessibility into value. The results of this analysis show that the behaviour of the model is robust with respect to alternative specifications, the choice of the parameters within a reasonable space, and the variables that measure the activity level of each location.

Acknowledgements and disclaimer

The authors are grateful to two reviewers for helpful suggestions on an earlier version of this paper. 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

Anderson, J. (1979) "A theoretical foundation for the gravity equation" The American Economic Review (69:1) pp106-116.

Anderson, J. and E. van Wincooop (2003) "Gravity with gravitas: A solution to the border puzzle" American Economic Review (93:1) pp. 170-192.

Banister, D. and Berechman, J. (2000) "Transport investment and economic development" University College London Press, London, UK.

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

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

Ben-Akiva, M. and Lerman, S.R. (1979) "Disaggregate travel and mobility-choice models and measures of accessibility" In Hensher, D.A. and Sotpher, P.R. (eds.) "Behavioural travel modelling". Croom-Helms, London, UK.

Bergstrand, J. (1985) "The gravity equation in international trade: Some microeconomic foundations and empirical evidence" The Review of Economics and Statistics (67:3) pp. 474-481.

Berry, B. and Marble, D. (1968) "Spatial analysis: A reader in statistical geography" Prentice-Hall, NJ, USA.

Bigman, D. and Deichmann, U. (2000) "Spatial indicators of access and fairness for the location of public facilities" In Bigman, D. and Fofack, H. (eds.) "Geographic targeting for poverty alleviation: methodology and application" World Bank, DC, USA.

BTS (2008) "Toll Facilities in the US Bridges-Roads-Tunnels-Ferries" Publication from the Office of Highway Policy Information, DC, USA.

Crotte, A., Noland, R.B., and Graham, D.J. (2008) "Estimation of road traffic demand elasticities for Mexico City" Mimeo. Centre for Transport Studies, Imperial College, London, UK.

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.

Disdier, A.C. and K. Head (2006) "The puzzling persistence of the distance effect on bilateral trade" WP No. 186, Centro Studi Luca d'Aglian, University of Milan, Italy.

Domencich T. A. and McFadden D. L. (1975) "Urban travel demand: A behavioural analysis" North-Holland, Amsterdam, Netherlands.

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

Ewing, R. (1993) "Transportation service standards– as if people matter" Transportation Research Record (14) pp.10-17.

Fotheringham, S. (1981) "Spatial structure and distance-decay parameters" Annals of the Association of American Geographers (71:3) pp. 425-436.

Fujita, M., Krugman, P. and Mori, T. (1999) "On the evolution of hierarchical urban systems" European Economic Review (43) pp. 209-251.

Fürst, F. and Schürmann, C. (2000) "The SASI Model: demonstration examples Deliverable D15 EU SASI" Report for the European Commission, prepared by Institut für Raumplanung, Universität Dortmund, Dortmund, Germany.

Geurs, K. and van Wee, B. (2004) "Accessibility evaluation of land-use and transport strategies: review and research directions" Journal of Transport Geography (12) pp.127-140.

Hägerstrand, T. (1970) "Space-time and human conditions," In Karlqvist A., Lundqvrist, L. and Snickards, F. (eds.) "Dynamic allocation of urban space" Saxon House, Lexington, KY, USA.

Hansen, W. (1959) "How accessibility shapes land use" Journal of the American Institute of Planners (25) pp.73-76.

IMT (2000) "Manual Estadístico de Transporte" Internet publication from the Instituto Mexicano del Transporte. http://www.imt.mx/manual/index.php

INEGI (1999) "Censo Económico 1999: Origen y destino de los pasajeros y la carga transportada." Publication from the Instituto Nacional de Estadística, Geografía e Informática, Aguascalientes, México.

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

Keeble, D. and Walkers, S. (1988) "Peripherals regions in a community of twelve member states" Report for the Office for Official Publications of the EC, Luxembourg, Luxembourg.

Lopez-Cordova, E. (2002) "NAFTA and Mexico's Manufacturing Productivity: An Empirical Investigation Using Micro-Level Data" Mimeo. Inter-American Development Bank, DC, USA.

Martinez, F. (1995) "Access: The transport-land use economic link" Transport Research B (29:6) pp. 457-470.

McCallum, J. (1995) "National borders matters: Canada-US regional trade patters" American Economic Review (85:3) pp. 615-623.

Miller (2000) "Within transportation networks: Basic theory and computational procedures" Geographical Analysis (31) pp. 187-212.

Miller, H. (1999) "Measuring space-time accessibility benefits within transport networks: Basic theory and computational procedures" Geographical Analysis (31) pp. 187-212.

Montenegro, C.E. and I. Soloaga (2006) "NAFTA's trade effects: New evidence with a gravity model" Estudios de Economía (33:1) pp. 45-63.

NATS (2006) "North American merchandise trade" Internet publication from the North American Transportation Statistics http://nats.sct.gob.mx.

Preston, J. and Raje, F. (2007) "Accessibility, mobility, and transport-related social exclusion" Journal of Transport Geography (15) pp.151-160.

Rietveld, P. (1989) "Infrastructure and regional development: A survey of multiregional economic models" The annals of Regional Science (23) pp. 255-274.

Santos Silva, J.M.C. and Tenreyro, S. (2006) "The log of gravity" The Review of Economic and Statistics (88:4) pp. 641-658.

Schürman, C. and Tallat, A. (2000) "Towards an European peripherally index" Report for the European Commission, prepared by Institut für Raumplanung, Universität Dortmund, Dortmund, Germany.

Soloaga, I. Wilson, J.S. and A. Mejia (1996) "Moving forward faster: Trade facilitation reform and Mexican competitiveness" WP No. 3953, World Bank Policy Research, World Bank, DC, USA.

Thill, J.C. Horowitz, J.L. (1997) "Travel time constrains in destination choice sets" Geographical Analysis (29) pp. 108-123.

Weibull, J. (1976) "An axiomatic approach to the measurement of accessibility" Regional Science and Urban Economics (6) pp. 357-379.

Weibull, J. W. (1980) "On the numerical measurement of accessibility," Environment and Planning A (12) pp. 53-67.

Williams H. C. (1976) "Travel demand models, duality relations, and user benefit analysis" Journal of Regional Science (16) pp. 147-166.

Williams H. C. (1977) "On the formation of travel demand models and economic evaluation measures of user benefit" Environment and Planning A (9) pp. 285-344.

Wolf, H.C. (2000) "International home bias in trade" The Reviews of Economics and Statistics (82:4) pp. 555-563.

Ypma, B. (2000) "International comparison of the role of accessibility in transport policy" B&A Group, The Hague, Netherlands.

Table 1 Cost Structure of Freight Industry in Mexico

	Total Cost	
	MEX$ '000	%
Wages	9,878.90	26
Fuel	15,863.51	42
Tolls	3,095.29	8
Other Costs	8,866.49	24
Total Costs	37,704.19	100

Source: National Economic Census 2004

Table 2 Selected Socio-economic Indicators in Mexico and the USA

	USA	Mexico	Ratio
Demographic (individuals)			
Population	281,421,906	99,879,361	2.82
No farm Employment	163,958,700	28,157,810	5.82
Manufacturing Employment	16,782,221	4,138,520	4.06
Production			
No Farm Personal Income			
Total (USAD'000)	8,672,369,281	284,926,140	30.44
Average (USAD)	52,893.62	10,118.90	30.44
Wage and Salaries			
Total (USAD'000)	4,938,031,880	72,013,586	68.57
Average (USAD)	30,117.53	2,557.50	68.57

Source: USA Bureau of Economic Analysis; National Economic Census 2004, USA Population Census 2000, INEGI Population Census 2000, Mexico.

Table 3 Metropolitan Areas

Code	Metropolitan Area	Code	Metropolitan Area
ACA	Acapulco	MTY	Monterrey
AGU	Aguascalientes	MLM	Morelia
CPE	Campeche	NOG	Nogales
CUN	Cancun	NLD	Nuevo Laredo
CJS	Cd. Juarez	OAX	Oaxaca
CEL	Celaya	ORI	Orizaba
CHE	Chetumal	PCH	Pachuca
CUU	Chihuahua	PNG	Piedras Negras
CPG	Chilpancingo	PZC	Poza Rica
ACU	Ciudad Acuña	PBC	Puebla
CME	Ciudad del Carmen	PVR	Puerto Vallarta
CEN	Ciudad Obregon	QRO	Queretaro
CVL	Ciudad Valles	REY	Reynosa
CVM	Ciudad Victoria	SMC	Salamanca
COA	Coatzacoalcos	SLW	Saltillo
CLQ	Colima	SCC	San Cristobal de las Casas
CUA	Cuautla	SJR	San Juan del Rio
CUE	Cuernavaca	SLP	San Luis Potosi
CUL	Culiacan	SRC	San Luis Rio Colora
DGO	Durango	TAM	Tampico
ENS	Ensenada	TCL	Tapachula
GDL	Guadalajara	THC	Tehuacan
HMO	Hermosillo	TPQ	Tepic
IGL	Iguala	TIJ	Tijuana
IRA	Irapuato	TLA	Tlaxcala
PAZ	La Paz	TOL	Toluca
BJX	Leon	TRC	Torreon
LMC	Los Mochis	TGZ	Tuxtla Gutierrez
ZLO	Manzanillo	URU	Uruapan
MAM	Matamoros	VER	Veracruz
MZT	Mazatlan	VSA	Villahermosa
MER	Merida	JAL	Xalapa
MXL	Mexicali	ZCL	Zacatecas
MEX	Mexico City	ZHG	Zamora de Hidalgo
LOV	Monclova		

Based on INEGI (2000b)

Table 4 Robust OLS Estimation for the Decay Parameter θ
Interstate Gravity Model in Mexico: Estimated Data
(Standard Errors in Parenthesis)

	Canonical				No Dummies			
	$\gamma_1 \neq \gamma_2$		$\gamma_1 = \gamma_2$		$\gamma_1 \neq \gamma_2$		$\gamma_1 = \gamma_2$	
	Coefficient[/9]	R^2[/10]	Coefficient[/9]	R^2[/10]	Coefficient	R^2	Coefficient	R^2
Complete Dataset								
GDP [/1]	-0.73	0.81	-0.73	0.81	-0.86	0.61	-0.85	0.60
	(0.1138)		*(0.1138)*		*(0.1018)*		*(0.1023)*	
GDP II[/2]	-0.73	0.81	-0.73	0.81	-0.79	0.68	-0.79	0.68
	(0.1138)		*(0.1138)*		*(0.0974)*		*(0.0984)*	
GDP III[/3]	-0.73	0.81	-0.73	0.81	-0.94	0.59	-0.93	0.58
	(0.1138)		*(0.1138)*		*(0.1042)*		*(0.1050)*	
Non Farm GDP [/4]	-0.73	0.81	-0.73	0.81	-0.90	0.63	-0.90	0.62
	(0.1138)		*(0.1138)*		*(0.0995)*		*(0.1003)*	
Population	-0.73	0.81	-0.73	0.81	-0.70	0.45	-0.70	0.44
	(0.1138)		*(0.1138)*		*(0.1184)*		*(0.1186)*	
Workforce [/5]	-0.73	0.81	-0.73	0.81	-0.75	0.50	-0.74	0.49
	(0.1138)		*(0.1138)*		*(0.1125)*		*(0.1129)*	
Workforce II [/6]	-0.73	0.81	-0.73	0.81	-0.86	0.21	-0.86	0.21
	(0.1138)		*(0.1138)*		*(0.1417)*		*(0.1415)*	
Workforce III [/7]	-0.73	0.81	-0.73	0.81	-0.80	0.53	-0.79	0.52
	(0.1138)		*(0.1138)*		*(0.1087)*		*(0.1090)*	
Non Farm Workforce[/8]	-0.73	0.81	-0.73	0.81	-0.70	0.44	-0.70	0.43
	(0.1138)		*(0.1138)*		*(0.1193)*		*(0.1198)*	
N	496		496		496		496	

Source: Own calculation
/1 GDP includes production of all economic sectors in the state
/2 GDP II includes production of the secondary sector in the state
/3 GDP II includes production of the tertiary sector in the state
/4 Non Farm GDP includes production of the secondary and tertiary sector in the state
/5 Workforce includes total employment in the state
/6 Workforce includes employment in the secondary sector in the state
/7 Workforce includes employment in the tertiary sector in the state
/8 Non Farm Workforce includes employment in the secondary and tertiary sector in the state
/9 The coefficient and standard error are the same for all the cases up to the fifth decimal
/10 The statistics are the same for all the cases up to the fifth decimal

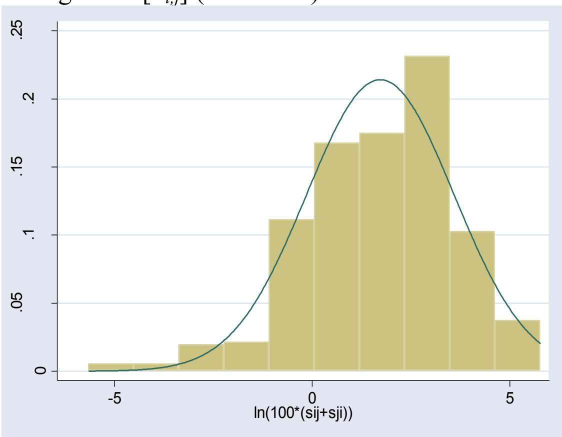
Table 5 Domestic, International, and Regional Accessibility
Canonical Accessibility Model
Normalisation: Average=100

		Domestic	International	Regional [/1]
I	North East			
	Population	91.11	125.95	42.99
	Production	91.63	124.93	n.a.
II	North West			
	Population	62.56	108.31	64.62
	Production	58.05	108.97	n.a.
III	North Centre			
	Population	84.66	116.85	46.64
	Production	82.10	116.43	n.a.
IV	Centre			
	Population	170.15	96.44	142.15
	Production	173.93	96.48	n.a.
V	Centre-West			
	Population	118.32	98.05	131.51
	Production	111.34	98.14	n.a.
VI	South			
	Population	99.48	86.02	105.12
	Production	98.83	86.25	n.a.
VII	East			
	Population	108.74	92.01	194.51
	Production	115.70	92.09	n.a.
VIII	Yucatan			
	Population	64.98	76.37	72.47
	Production	68.41	76.71	n.a.

Normalisation: Average=100
Source: Own calculation
[/1] Production based index not available due to lack of data on local production

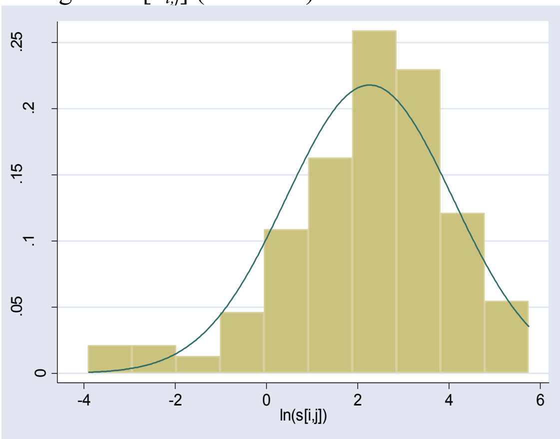
Figure 1 Distribution of Freight Flows

Figure 1A
Histogram $\ln[T_{i,j}]$ (Estimated)



Source: Own elaboration based on the Origin and Destination of Passengers and Freight Survey in Mexico

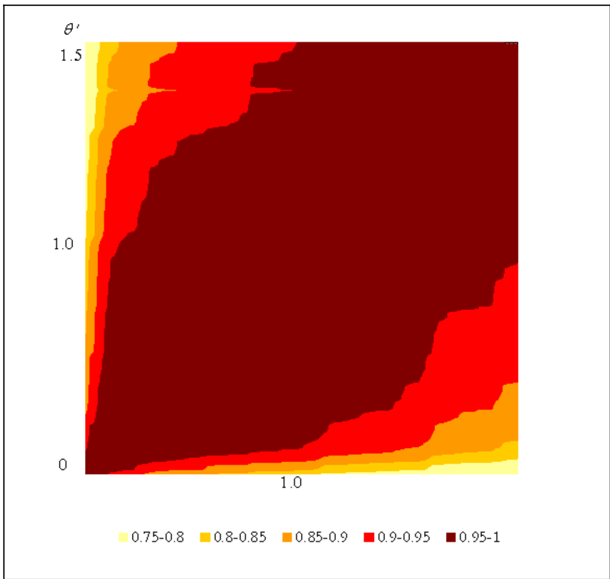
Figure 1B
Histogram $\ln[T_{i,j}]$ (Observed)



Source: Own elaboration based on the Origin and Destination of Passengers and Freight Survey in Mexico

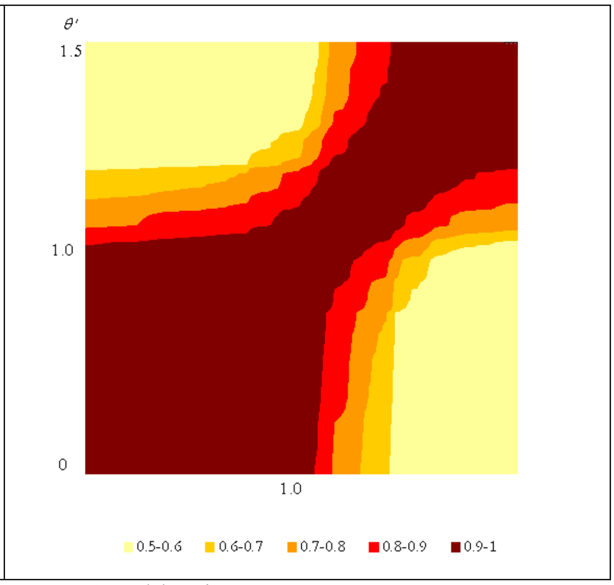
Figures 2 Correlation Surfaces

Figure 2A
$\Psi(\mathbf{x}, \mathbf{x}', \mathbf{W}, \theta, \theta)$ for National Accessibility



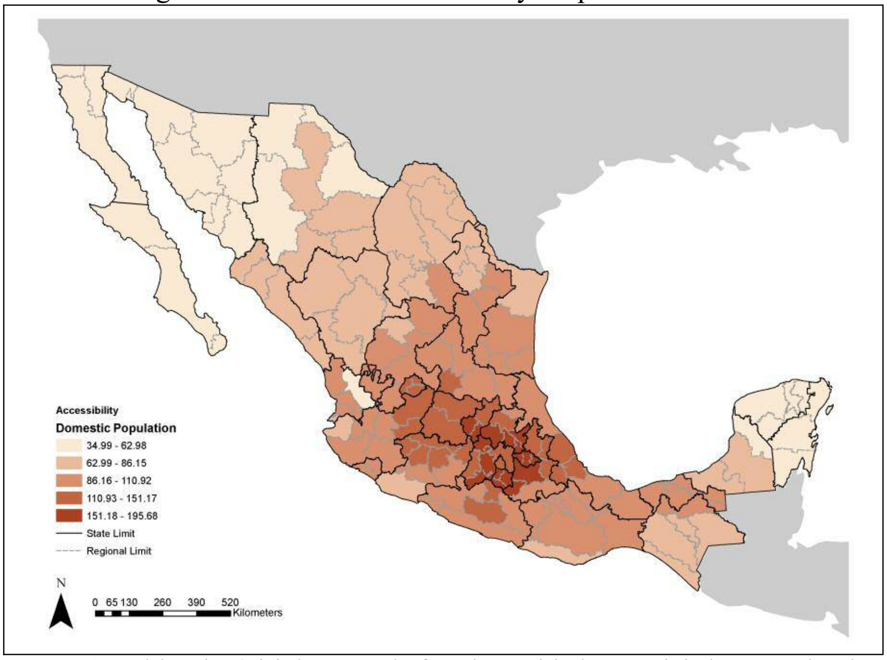
Source: Own elaboration

Figure 2B
$\Psi(\mathbf{x}, \mathbf{x}', \mathbf{W}, \theta, \theta)$ for International Accessibility



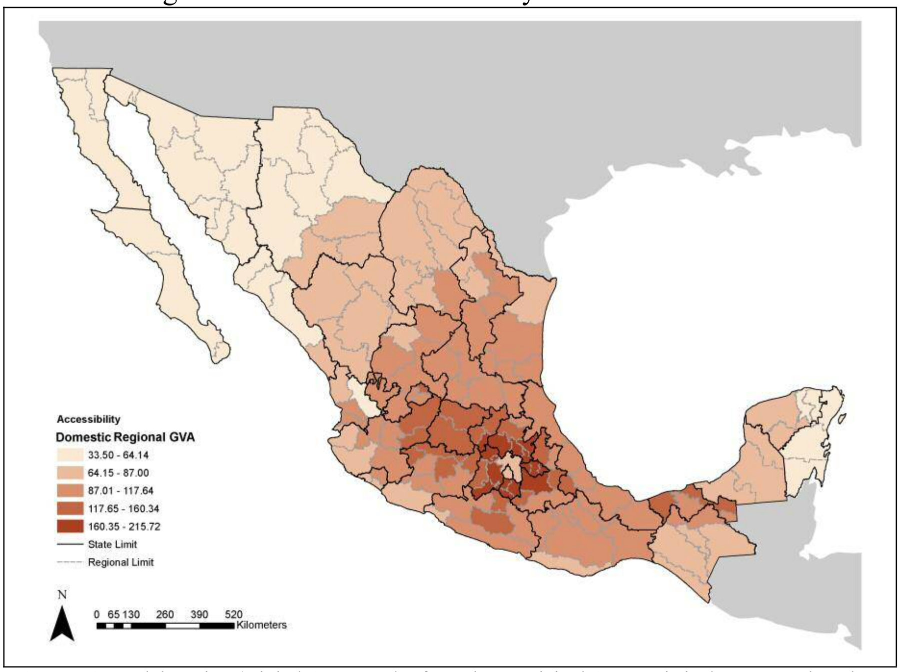
Source: Own elaboration

Figure 3A Domestic Accessibility: Population Based



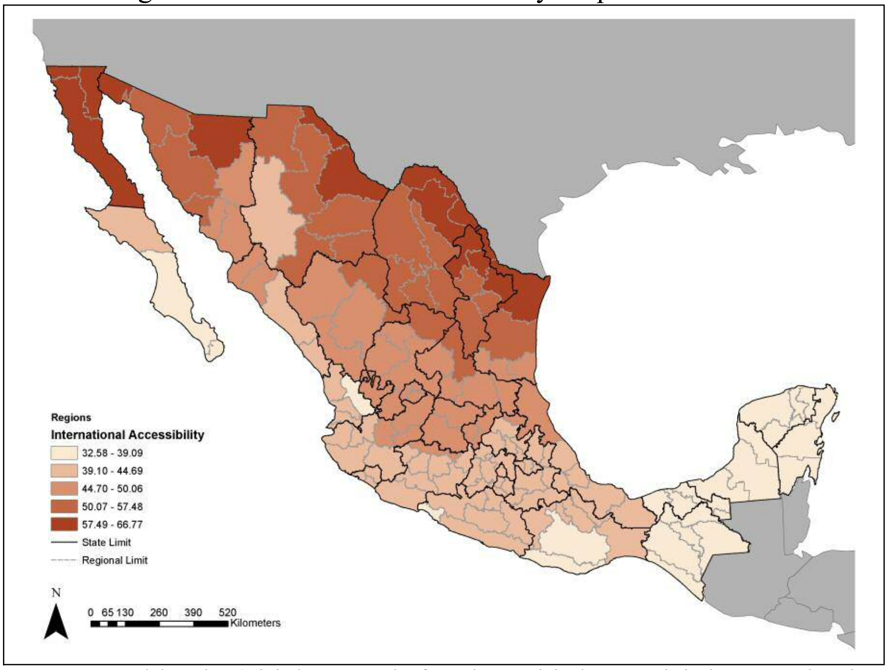
Source: Own elaboration (Digital Cartography from the Municipal Geo-statistical Framework and the Topographic Digital Dataset, INEGI)

Figure 3B Domestic Accessibility: Production Based



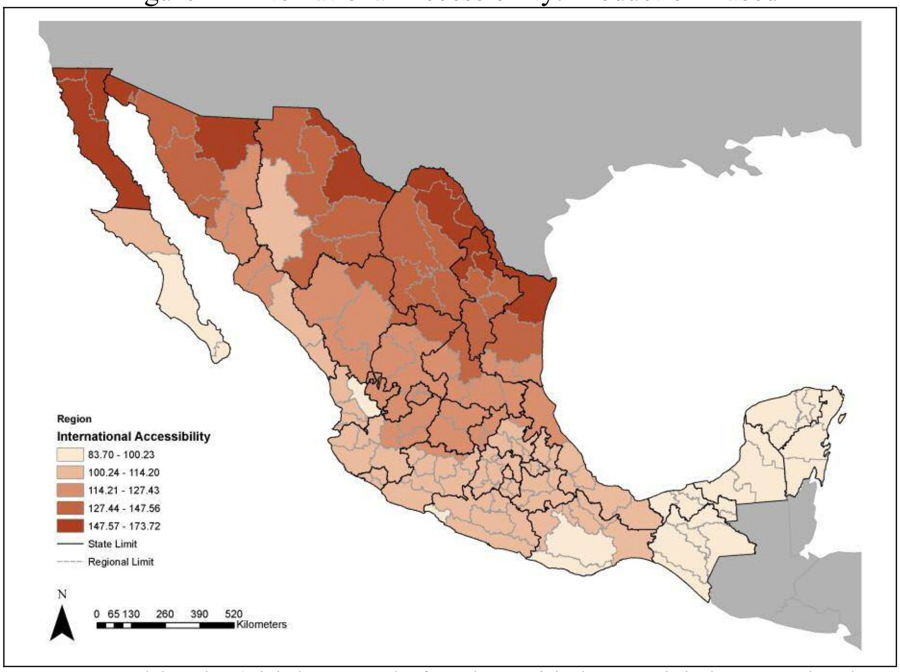
Source: Own elaboration (Digital Cartography from the Municipal Geo-statistical Framework and the Topographic Digital Dataset, INEGI)

Figure 4A International Accessibility: Population Based



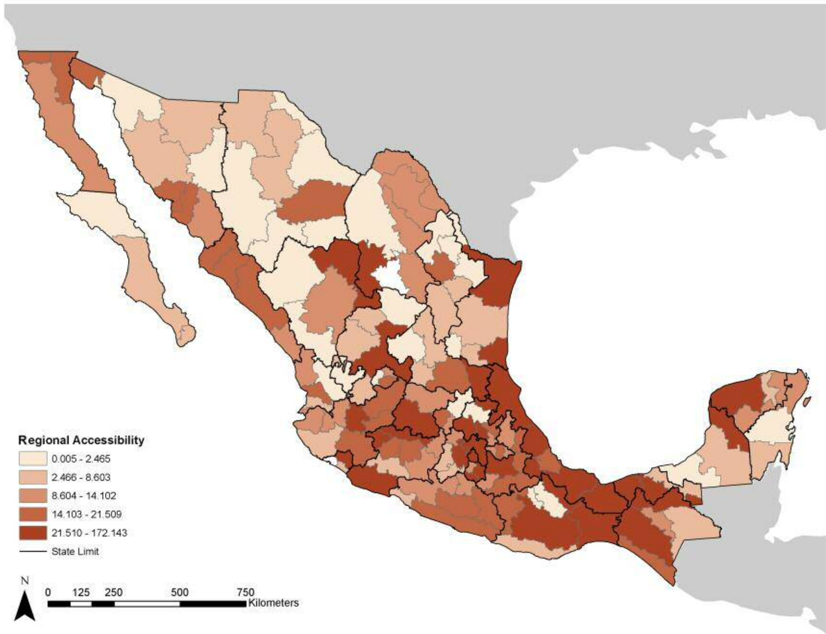
Source: Own elaboration (Digital Cartography from the Municipal Geo-statistical Framework and the Topographic Digital Dataset, INEGI)

Figure 4B International Accessibility: Production Based



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

Figure 5 Regional Accessibility: Population Based



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