Roads and the Geography of Economic Activities in Mexico *

Brian Blankespoor† Théophile Bougna‡
Rafael Garduno-Rivera§ Harris Selod ¶

October 21, 2017

Abstract

This paper estimates the impacts of road improvements on local employment structure and specialization in Mexico over the 1985-2016 period. Using geo-referenced panel data, it measures access to domestic markets from each locality as a weighted sum of surrounding populations (market access) or incomes (market potential), with weights inversely related to travel time or travel cost. Instrumenting for road placement endogeneity and addressing the recursion problem in regressions that involve access to markets, the analysis finds significant and positive causal effects of improved accessibility on employment and specialization. Heterogeneous effects are found across sectors and regions.

Keywords: Mexico, Roads, Market Access, Market Potential, Specialization, Industrial Localization.
JEL classification: R12; R30; C23.

*We thank Luis Rodrigo Gonzalez Vizuet, Shiyan Zhang, and Lucie Letrouit for excellent research assistance. We are indebted to Mike Bess for the many inputs he provided regarding the history of roads in Mexico, and to Uwe Deichmann, Maria Marta Ferreyra, Mark Roberts and Matthew Turner, who provided useful comments on the preliminary versions of this paper. We also thank participants to the authors' workshop for the World Bank study on 'Cities, Productivity and Growth in Latin America and the Caribbean'. Funding from the World Bank and from DFID under the World Bank Strategic Research Program (SRP) '*Transport for Sustainable and Inclusive Growth*' is gratefully acknowledged.

†Development Data Group, The World Bank, E-mail: bblankespoor@worldbank.org
‡Development Research Group, The World Bank, E-mail: tbougna@worldbank.org
§Centro de Investigacion y Docencia Economicas, Mexico. E-mail: rafael.garduno@cide.edu
¶*Corresponding author*: Development Research Group, The World Bank. E-mail: hselod@worldbank.org

1 Introduction

There are many benefits associated with investments in transport infrastructure which can stimulate growth through trade, structural transformation, agglomeration and productivity (see Redding and Turner, 2015, and Berg el al., 2017 for surveys). The impacts can be large: Calderón, Moral-Benito and Servén (2009) estimate that a 10 percent improvement in the stock of infrastructure (including transport infrastructure) increases per capita GDP by 0.7 to 1 percent. In Latin America, in particular, the potential to stimulate growth through transport investments could be especially significant given the large infrastructure gap in most Latin American countries (see Fay and Morrison, 2007; Gonzalez, Guasch and Serebrisky, 2007). Perrotti and Sánchez (2011) estimate that 'closing the infrastructure gap' in Latin American countries would require investments in the amount of up to 6.2 percent of annual GDP for the period 2012- 2020. In theory, benefits could then materialize in the form of a reduction in transportation costs (or travel time), more intense competition, and deeper economic integration. Locations where access to markets is improved would become more attractive for firms, potentially triggering a process of concentration and specialization of local economic activity whereby innovation, the sharing of input markets, and improved matching on the labor market would increase productivity, and in turn lead to local economic development and an increase in living standards.

In this paper, we focus on the first steps in this potential chain of events, in the case of Mexico, a country which experienced relatively large investments in roads over the past decades and from which lessons could be derived for the rest of the continent. In Mexico, road transport accounts for 57% of freight transport and remains a vital component of the economy. The Mexican road network links the country from north to south and between its two oceanic coastlines; and some of the most important road connections link the capital city with border crossings to the US. We investigate the extent to which these road improvements increased access to markets and, in turn, affected the location of economic activities.

Our paper contributes to a growing literature that looks at the effect of transport infrastructure on local economic development outcomes (Donaldson and Hornbeck, 2015; Duranton and Turner, 2012; Michaels, 2008; Baum-Snow, 2007; Baum-Snow et al., 2017; Jedwab and Storeygard, 2016; Alder, 2016; Alder, Robert and Tewari, 2017; Straub and Bird, 2015, Ghani et al., 2015; Faber, 2014; Rothenberg, 2015, Fyre, 2015; Galasso and Oettl, 2017). In view of this literature, our paper looks at the effects of roads on the geographical concentration of economic activities and the specialization of localities.

There are two challenges to estimate the causal impact of roads on local economic outcomes: the non-random placement of roads, and the recursion problem inherent in regression of economic outcomes on market access measures. To address these two problems, we adopt three separate identification strategies. First, we use the so-called 'doughnut' strategy, which allow to instrument access to markets excluding zones that are more likely to be prone to an endogeneity bias due to non-random road placement. Second, in order to avoid the recursion problem, we substitute market access indicators with measures of access to infrastructure that do not involve population or income. These measures include the number of roads intersecting a circle of a given radius around the centroid of each locality, or the number of kilometers of roads within a given distance to that centroid. Third, we use these access to infrastructure variables as instruments for market access indicators.

Our paper makes three principal contributions. First, we construct a new panel of geo-referenced roads data in Mexico over three decades (for the period 1986-2014). Second, we also take advantage of the panel nature of our data to estimate time-varying measures of market access for more than 2,000 localities (whereas previous studies, either just used a few dates or had a sample with a limited number of observations). Third, to the best of our knowledge, our paper is the first to study the effect of market access on specialization.

Our key findings are that transportation improvements over the last three decades increased access to markets in Mexico—which we measure in concurrent ways—and that these improvements in accessibility had a positive impact on local employment and economic specialization. The effects are large and heterogeneous across sectors. A 10 percent increase in market access results in a 1.6 to 2.1 percent increase in employment and a large increase in locality specialization. The effect of our market potential measure is slightly larger. A 10 percent increase results in a 2.9 to 6.5 percent increase in employment, and a 13.0 percent increase in output specialization. We also find heterogeneous effects across sectors with employment in commerce and services benefiting more than manufacturing from road improvements.

The rest of the paper is structured as follows. Section 2 presents the Mexican context. Section 3 describes the data used, explains the construction of our market access and market potential measures, and provides information on the trends in industrial concentration and specialization in Mexico. Section 4 presents our empirical methodology and discusses identification strategies. Section 5 discusses our empirical results regarding the effects of market access and market potential on employment and specialization, and Section 6 concludes.

2 The Mexican context: Lessons from the literature

2.1 The concentration of industries and specialization of localities

Economic activity in Mexico is highly concentrated in the Mexico City Metropolitan Area, which in 2010, contributed to a quarter of the national Gross Value Added even though it covers less than 0.3 percent of the national territory. In the wake of Michael Porter's work (Porter, 1998), there has been increasing policy interest in Mexico to facilitate the industrial concentration and specialization of localities and the development of industries, so as to ultimately achieve higher economic growth. However, economic studies that shed light on these issues are scarce and only a few case studies focus on industrial concentration and specialization.

A first set of such studies measures the level of concentration of various industries and its increase over time (Unger, 2003) and tries to identify the determinants of industrial clustering, concluding that, in Mexico, labor force skill play a very important role to attract clusters, much more than wage differentials (Unger and Chico, 2004). Complementing these observations, several authors identify the locations that have become more specialized over time, focusing on different spatial scales. Pérez and Palacio (2009) find that during the 1994-2004 period, specialization has increased. Focusing on Mexican cities, Kim and Zangerling (2016) find that, between 1990 and 2010, specialization only increased in the Mexico City Metropolitan Area, but did not in other cities which remained very much diversified.

A second set of case studies points at the potential benefits of clustering in Mexico. Monge (2012) suggests that clustering of the tequila industry in the state of Jalisco (in Western Mexico) reduced transaction costs and that entrepreneurs in that sector benefitted from labor specialization. Dávila (2008), who studied the economic performance and the commercial integration with Texas of industrial clusters in northeastern Mexico[1] during the 1993-2003 period, suggests that economic clusters played a role to foster productive innovations and stimulate bilateral trade.

2.2 Road investments: A historical perspective

In Mexico, large investments in roads started during the Spanish Colony (1521-1810), but these roads were mainly focused on connecting natural resources (especially silver and gold) with the port of Veracruz, to ship them to Spain.

During the presidency of Porfirio Días (between 1884 and 1911), while railways flourished throughout Mexico, the road network, which essentially dated back to the colonial period, received little financial support from the government and consequently deteriorated (Bess, 2016b). The Mexican Revolution, between 1910 and 1920, did not improve the situation. In 1918, municipal surveyors in Mexico City evaluated that the conflict had damaged around 4,000 km of roads serving the federal capital (Bess, 2016b). In the 1920s and 1930s, during the presidency of Alvaro Obregon and Plutarco Elias Calles, national and state political leaders engaged in the reconstruction and enlargement of the road network, mobilizing, to this aim, the private sector, public school teachers and rural communities all over the country (Bess, 2016a). Thousands of kilometers of highways were constructed.

The 1940s and 1950s were characterized by a great hope in the promises of industrialization and a generalized drive toward economic modernization. Road construction was perceived as necessary to allow market growth and improve the accessibility of regions (Bess, 2014). This

[1] i.e. Chihuahua, Coahuila, Nuevo Leon and Tamaulipas

conception led to the building of hundreds of kilometers of new roads in Mexico by state road-building agencies, mobilizing large public spending and private domestic investments as well as foreign ones. The U.S. invested millions of dollars of direct investments in Mexican transportation industry and infrastructure.

During President Miguel Aleman's term (1946-1952), the first freeway (from Mexico City to Acapulco) was opened and became a model for the construction of future freeways. During this period, road building played a key role in the modernization of the Mexican economy and the development of major commercial industries (Bess, 2014).

In the 1960s, roads were built to respond to the needs of private firms and also to serve the national and state governments' objective to build strategic relationships with rural communities. Under President Adolfo Lopez Mateos (in office between 1958 and 1964), an unprecedented amount of 300 million pesos was raised in bonds for the building of new highways (Bess, 2017). In addition, a government owned company was founded to build more than a thousand kilometers of toll roads in the center of Mexico.

The presidency of Carlos Salinas de Gortari (1988-1994) constituted a significant milestone in the history of road construction in Mexico with the launch of a very ambitious program, the *Programa Nacional de Solidaridad*, which led to the construction of 5,800 kilometers of privately financed highways at a cost of $15 billion (Foote, 1997). At the same time, cutting its own spending on road infrastructure, the government privatized toll road operations, a lucrative business for road-building firms. This allowed Mexico to build, in just six years, 'what had taken two decades to achieve in western European countries' (Foote, 1997). However, the extremely high tolls ended up deterring trucks from using these new roads. In response, the federal government announced, in January 1997, the mobilization of $3.3 billion over 30 years to restructure the highway network. This outlay was added to a $ 1.7 billion toll-road rescue plan to help the state-owned Mexican banks that had financed road building on non-market terms (Foote 1997).

After 2000, the opposition party, the *Partido de Acción Nacional*, eventually gained power. Road building policies were nevertheless pursued in continuity with past policies, which considered the construction of roads as a symbol of modernization. President Vicente Fox (in office between 2000 and 2006) mobilized hundreds of millions of pesos for road building through a program for basic infrastructure, and his successor, President Felipe Calderón (in office between 2006 and 2012), constructed and renovated more than 23,000 km of roads as part of a program to address rural poverty. Eventually, in April 29, 2014, the federal government launched the National Infrastructure Program[2] 2014-2018 (NIP), projecting a substantial increase in investment compared to the last twenty years (Pérez-Cervantes and Sandoval-Hernández, 2015). The most ambitious part of this new program focuses on the south of the country, while no large project is planned in the north and while the center of Mexico, more populated and richer,

[2]Programa Nacional de Infraestructura.

receives smaller projects.

2.3 Roads and economic activities

To our knowledge, the impact of roads on the location of economic activities in Mexico has not been studied yet. A set of studies, however, have investigated the link between infrastructure and productivity. For instance, Becerril-Torres et al. (2010) study the effect of total infrastructure (measured with an index that includes roads, ports, airports, and telecommunications) on the convergence across states in technical efficiency. They find that their infrastructure index is associated with greater regional productivity, but mainly during the Import Substitution Industrialization (ISI) period (1970-1985), before Mexico entered the North American Free Trade Agreement (NAFTA) in 1994. Similarly, Brock and German-Soto (2013) find that lower levels of infrastructure investments had a lower effect on regional productivity during the NAFTA period, and conclude that continued investment in transportation will be necessary to boost industrial sector growth. Focusing specifically on the effect of road infrastructure on productivity growth in the manufacturing sector in Mexico, Duran-Fernandez and Santos (2014a and 2014b) conclude that road infrastructure has positive effects on productivity and on the average product of labor.

3 Data and descriptives statistics

3.1 Firms and employment

We use panel data for employment and firm locations from two different sources published by the Mexican National Institute of Statistics (INEGI).[3] The first source is the Economic Census for the years 1986, 1989, 1994, 1999, 2004, 2009, and 2014, which provides employment figures and breakdown by broad sector of activity, as well as total income at the level of municipalities and for formal businesses. The second source of data is the Directory of Economic Units (DENUE)[4] for the years 2004, 2009 and 2014. It is an exhaustive dataset which contains detailed micro-geographic information on all formal establishments, including the 6-digit NAICS classification, the number of employees of the establishment, turnover, the municipality identifier, and exact geographic coordinates. For the year 2014, the database contains information on approximately 3,000,000 establishments throughout the country. Note that although the DENUE database has finer geographic identifiers than the Economic Census (actual coordinates vs. municipality) and a finer industry identifier (6-digit NAISC vs. broad sectors of activity), it covers less years than the Economic Census. In the analysis, we use either one, as appropriate.

[3]Instituto Nacional de Estadísticas y Geografía.
[4]Directorio Estadístico Nacional de Unidades Económicas.

Table 1 presents the distribution of formal jobs in Mexico between 1999 and 2014 according to Economic Censuses. As can be seen, the bulk of jobs are in services, commerce and manufacturing. Although employment has increased in all sectors over the period, the greatest increases have occurred for commerce and service jobs, which relative shares have gone up while the relative shares of all the other sectors have gone down. The decrease in the relative share of manufacturing employment from 30 to 24 percent is particularly noticeable. In 2014, more people are actually employed in the commerce sector than in manufacturing, showing that Mexico is following a trend of tertiarization as developed economies have in previous decades.

Table 1: Sectoral distribution of formal jobs in Mexico (1999 – 2014)

Sector	1999	share	2009	share	2014	share
Agriculture	174,127	1.27	180,083	0.90	188,566	0.87
Mining	113,189	0.83	14,2325	0.71	166,548	0.77
Manufacturing	4,1754,00	30.47	4,661,062	23.17	5,073,432	23.51
Commerce	3,792,466	27.68	6134758	30.50	6,389,648	29.61
Services	5,257,100	38.37	8,762,918	43.56	9,537,235	44.20
Other	190,033	1.39	235,688	1.17	220,929	1.02
Total	13,702,315	100.00	20,116,834	100.00	21,576,358	100.00

Source: Economic Censuces (INEGI) 1999, 2009 and 2014.

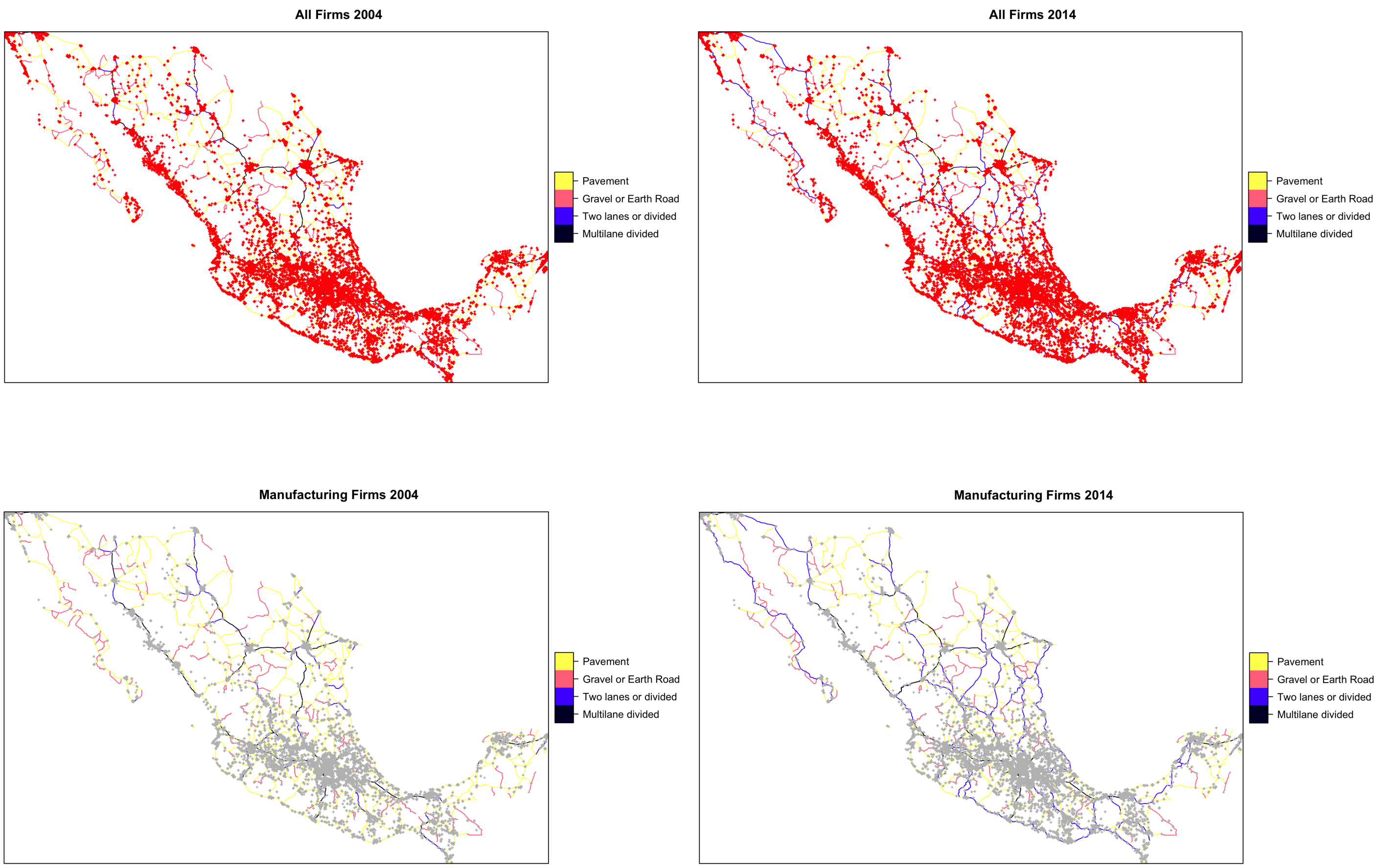
Figure 1 plots the location of establishments in the DENUE database in 2004 and 2014 for all firms (top panel) and for manufacturing firms (bottom panel) respectively.[5] It reveals a high concentration of economic activity in the center of the country and in the surroundings of Mexico City, with a greater concentration of manufacturing firms than for the other sectors, but no visually noticeable change over the past decade.

3.2 Roads

In this paper, we focus on the later waves of road construction and improvement that occurred since the mid-eighties. To have a consistent measure of road extent and road types over time, we constructed a new geo-referenced database for the period 1985-2016. This was done by importing historical road type and road extent information from the American Automobile Association (AAA) paper maps into the 2014 road geometry published by DeLorme. Because the DeLorme road geometry is network enabled and topologically correct, it allows for clean travel time computations that we need to measure access to markets (see below). Using this fixed geometry, we then imported information on road type and extent for the reference years

[5]On Figure 1, establishment locations are overlayed on roads. See the following subsection for details regarding the roads data construction.

Figure 1: The spatial distribution of firms in Mexico (2004 and 2014)



Sources: DeLorme, AAA and DENUE (INEGI). *Note*: Universe of all formal firm in 2004 (top left panel) and 2014 (top right panel). Manufacturing in 2004 (bottom left panel) and 2014 (bottom right panel).

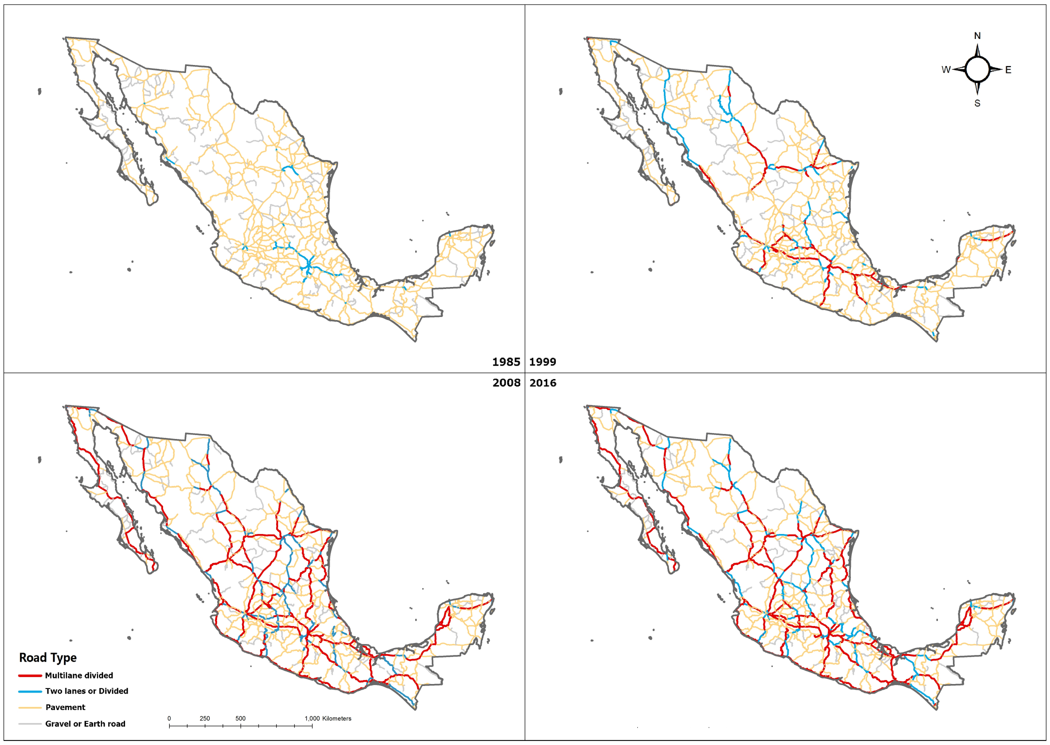
or the nearest available year before which we have firm data in the DENUE and in the Economic Censuses (1986, 1994, 1999, 2004, 2009 and 2014), leaving us with a road panel for the years 1985, 1993, 1999, 2004, 2008 and 2016.[6] We grouped the AAA functional classes into four categories, which we label 'Multilane divided', 'Two lanes or Divided', 'Pavement', and 'Gravel and Earth Road'.

Figure 2 represents the evolution of the road network for the years 1985, 1999, 2008, and 2016. As can be seen, major road improvements took place in the middle of the studied period, between 1999 and 2008. This is confirmed by table 13 in Appendix A, which provides road length by road category for all six years during the 1985-2016 period.[7]

[6]We use the roads information for 2016 as the AAA map for recent years before or in 2015 were not available to us.

[7]Observe that official statistics provided by INEGI for Mexico provide significantly greater total road lengths

Figure 2: The road network in Mexico (1985 – 2016)



Source: DeLorme (2014) and authors' calculations. Note: These maps represent a cross-sectional road geometry derived from DeLorme (2014) and updated by the authors using road category information from AAA maps (1985, 1999, 2008, 2016).

3.3 Geographic unit of analysis

Because our focus is on accessibility and the location of economic activities, the natural geographic unit of analysis is the urban or metropolitan area. However, because the data we use is produced at at a finer geographic scale—the municipality level—it is necessary to merge some municipalities into metropolitan areas. As INEGI provides data for 2,377 municipalities[8] and identifies 316 metropolitan areas—either as isolated municipalities or as groups of two municipalities—we are able to reconstruct a sample of 2,094 localities along a mixed classification that consists of the 316 metropolitan areas identified by INEGI and the remaining 1,778 isolated municipalities that are not classified by INEGI as metropolitan areas.[9] In what follows, we use the term locality to refer to either an isolated municipality or a metropolitan area in our

than what can be inferred from the AAA. This is because INEGI data accounts for all types of roads, including minor segments, whereas the AAA maps focus on the main roads, which are likely to be more relevant for trading.

[8]Since the 1989 census, new municipalities were created by splitting some of the old municipalities. In order to have a definition of localities that is stable across time, we merged these new municipalities back to their 1988 boundaries using the list provided by INEGI (2006).

[9]We consider here the 1990 definition of metropolitan areas.

sample. As a robustness check, all the regressions that we run on our mixed sample of 2,094 municipalities and metropolitan areas are also run on the sample of 2,377 municipalities.

3.4 Measures of accessibility

Market access

Each locality is characterized by its accessibility to markets and we resort to several accessibility measures. The first measure of accessibility for a locality i, which we refer to as market access (MA), is given by the following formula:

$$MA_{i,t} = \sum_{j \neq i} P_{j,t} \tau_{ij,t}^{-\theta} \tag{1}$$

where $P_{j,t}$ is the population of locality j at time t (which proxies for the size of the local market in j), $\tau_{ij,t}$ is the time required to travel between locality i and j given the state of the road network at time t, and θ is a measure of trade elasticity.

From formula (1), it is easy to see that the market access indicator is the discounted sum of the population of all the localities j that surround locality i, where the discount factor is inversely related to travel time. Travel times $\tau_{ij,t}$ are calculated on the reconstructed country-wide road network assuming that speed is a function of road type.[10] As for the trade elasticity parameter, in the absence of a specific study for Mexico, we use the same value suggested by Donaldson (2016) in the case of India (θ = 3.8).[11] Market access, which reflects the size of domestic markets accessible from location i, is frequently used in the literature in the absence of information on local incomes.

Market potential

The second measure of accessibility for a locality i, which we refer to as market potential (MP), is given by the following formula:

$$MP_{i,t} = \sum_{j \neq i} \frac{Y_{j,t}}{TC_{ij,t}^{\sigma-1}} \tag{2}$$

[10]We use the following speed assumptions: Multilane divided: 90 km/h; Two lanes or Divided: 80 km/h; Pavement: 70km/h; Gravel or Earth Road: 40 km/h; and unknown category (i.e., not shown on AAA paper maps but present in the 2014 DeLorme geometry): 5 km/h. These speed assumptions are consistent with the travel times published by INEGI for their own road geometry.

[11]In the regressions presented in the next section, we also perform robustness checks by using alternative values of the market access indicator constructed with other values for θ, the maximum value encountered in the literature being θ = 8.2 (see Pérez Cervantes and Sandoval Hernández, 2015)

where $Y_{j,t}$ is locality j's total income (in real terms) at time t, $TC_{ij,t}$ is a transport cost function between localities i and j, and σ is an elasticity term. Since the market potential formula has already been calibrated for Mexico, we use the same iceberg transport cost specification as the one estimated by Pérez Cervantes and Sandoval Hernández (2015), with $TC_{ij,t} = e^{(.0557+.0024\tau_{ij,t})}$ for $j \neq i$ and σ = 9. Similarly to formula (1), formula (4) measures the potential demand for goods traded from location i but in terms of income.[12]

Observe that in both formulas (1) and (2), we exclude 'own locality' to reduce endogeneity concerns.[13]

Table 2 below details the mean of the market access and market potential indicators at the locality level over the study period. As one can see, there have been large increases in both market access and market potential over the studied period.

Table 2: Mean and median market access and market potential of localities (1986 – 2014)

	Market access			Market potential		
Year	Mean	Median	SD	Mean	Median	SD
1986	24.41	0.00	389.89	–	–	–
1994	23.11	0.00	376.21	16,713.55	1,444.03	25,204.14
1999	24.90	0.00	393.58	445,492.00	55,543.83	617,044.60
2004	48.57	0.01	1,000.78	753,410.80	111,578.70	1,010,617.00
2009	53.09	0.01	1,093.79	1,271,577.00	220,128.80	1,680,713.00
2014	56.19	0.01	1,152.08	1,364,670.00	257,035.20	1,814,512.00

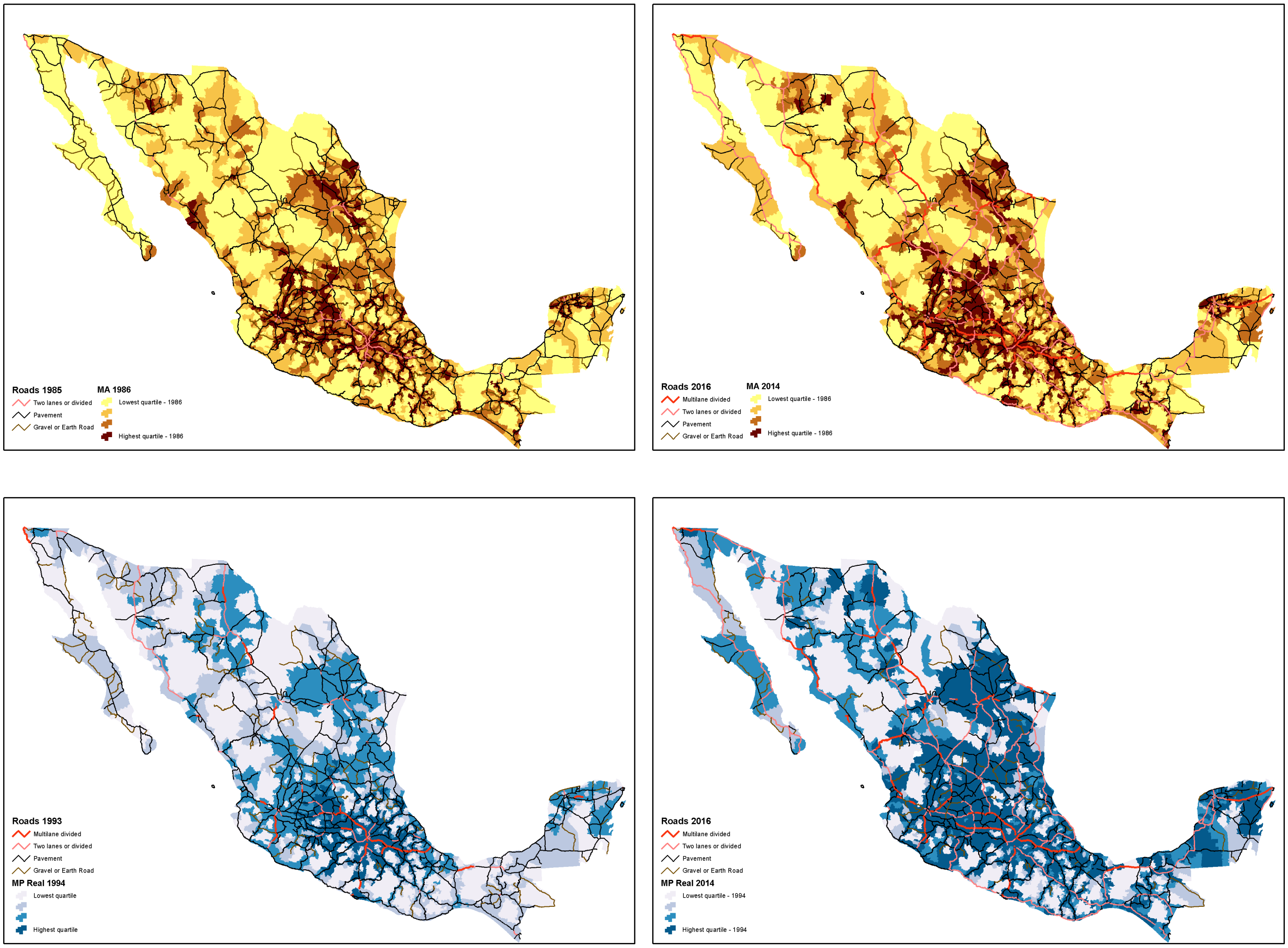
Sources: DeLorme, AAA and Economic Censuses (INEGI). *Notes:* Mean and median market access and market potential values are calculated over the sample of localities in Mexico.

On Figure 3, we represent market access and market potential at the locality level (overlayed on the road network) in 1986 and 2016. For comparison purposes, on both maps, the four categories correspond to the 1986 municipality quartiles of market access or market potential respectively. Although both maps show improvements along our two indicators, the contrast is starker with the market potential measure. This is understandably due to increases in real per capita income in addition to population growth.

[12]We use the real gross production provided by the economic census at the municipality level to proxy for local income.

[13]Also note that to increase accuracy of our measures, we actually calculate metropolitan area market access and metropolitan area market potential by first calculating municipality level formulas (1) and (2) excluding all localities j that belong to the same metropolitan area as municipality i. We then compute a metropolitan area weighted average of these municipality indexes using municipality area as the weight.

Figure 3: Market access and market potential in Mexico



Sources: DeLorme, AAA and Economic Censuses (INEGI). *Note*: Market access in 1986 (top left panel) and 2014 (top right panel). Market potential in 1994 (bottom left panel) and 2014 (bottom right panel).

Counts of road intersections and efficient road length

In addition to market access and market potential measures, we also construct local indicators of road availability. We do this in two ways as in Baum-Snow et al. (2017). The first measure is the number of roads intersecting a circle of a given radius centered on a locality's centroid. Another measure is the weighted length of roads within the same circle, where the length of each road type is weighted by the corresponding speed, but it excludes the roads in the locality itself.[14] Table 12 in Appendix A shows the mean and median average efficient

[14]The reason for this exclusion parallels the 'own locality' exclusion in the market access and market potential formulas and is made to reduce endogenity issues (see the econometrics section below).

kilometers of roads within a circle of 200km radius across localities for the 1986-2014 period. It can be seen that accessibility according to this variable has increased on average by 19.1 percent.

Access to external markets

Finally, to account for access to external markets, we also construct the minimum travel time and minimum travel cost (using the Pérez Cervantes and Sandoval Hernández, 2015) formula) to one among six major Mexican ports (documented by INEGI) and to one among forty-four entry ports to the U.S (documented by the U.S Department of Transportation, see US DOT, 2014). Table 12 in the Appendix A provides the mean and median minimum travel time and minimum travel cost to these ports and U.S. border entry ports. Over the 1986-2014 period, there is a slight decrease in the average minimum travel time (8 percent) and in the average minimum travel cost (3 percent) to a U.S. border entry port. The same is true for the average minimum travel time (6.1 percent) and travel cost (1 percent) to a major port.

3.5 Trends in industrial concentration and specialization

3.5.1 Trends in industrial concentration: 2004 – 2014

We examine how industrial concentration measures evolved between 2004 and 2014. Although the literature provides several measures of industrial concentration, we report the Ellison and Glaeser concentration index (see Ellison and Glaeser, 1997, and the Appendix B for details) calculated using plant-level data at both the 4-digit and 6-digit NAICS industrial classification. We calculate the Ellison and Glaeser index (henceforth EG index) for 2004, 2009, and 2014 at the localities level for all industries and for manufacturing only.

Table 3 reports our results, which leads to the three following comments. First, industries have become more geographically concentrated over the 2004-2014 period. This can be seen in the observed increase in the mean value of the EG index from 0.381 in 2004 to 0.430 in 2014 (a 10.3 percent increase). Second, manufacturing industries are on average more concentrated than the overall industries in Mexico. The mean value of the manufacturing EG index is on average 12 to 15 percent higher than the mean value of the EG index calculated for all industries. Third, applying commonly agreed upon thresholds in the literature, it can be seen that about 98-100 percent of industries (for overall industries and for manufacturing) can be said to be concentrated. The fraction of concentrated manufacturing industries (98 percent) is greater than the one reported for Canada (75 percent) in Behrens and Bougna (2015) and is roughly similar to the one reported for the U.S. (97 percent), France (95 percent), and the U.K. (94 percent) in Ellison and Glaeser (1997), Maurel and Sédillot (1999), and Duranton and Overman (2005) respectively. Fourth, the average level of concentration increases as one moves from the 4 to the 6-digit industry classification (concentration is relatively more intense among more specific industry segments). This result is consistent with Rosenthal and Strange (2003) for the U.S. case and with Behrens and Bougna (2015) in the case of Canada.

Table 3: Mean and median EG indices at the locality level, NAICS 4- and 6-digit industries (2004 – 2014)

	2004		2009		2014	
	Manufacturing	Overall	Manufacturing	Overall	Manufacturing	Overall
	NAICS 6-digit Industries					
Mean	0.381	0.256	0.388	0.26	0.43	0.282
Median	0.284	0.146	0.297	0.156	0.319	0.154
Minimum	-0.436	-0.739	0.045	-0.168	0.031	0.005
Maximum	1.001	1.006	1	1.006	1.353	1.013
Share < 0	1.23	0.95	–	0.27	–	–
Share $\in (0, 0.05]$	1.63	20.14	2.1	20	1.61	21.85
Share > 0.05	97.14	78.91	97.9	79.73	98.39	78.15
	NAICS 4-digit Industries					
Mean	0.344	0.222	0.348	0.216	0.359	0.236
Median	0.264	0.116	0.273	0.12	0.285	0.118
Minimum	-0.085	-0.225	0.033	0.004	0.02	0.005
Maximum	1	1	0.992	0.997	0.999	1
Share < 0	1.16	0.38	–	–	–	–
Share $\in (0, 0.05]$	2.33	25.95	3.49	24.23	3.49	25.84
Share > 0.05	96.51	73.66	96.51	75.77	96.51	74.16

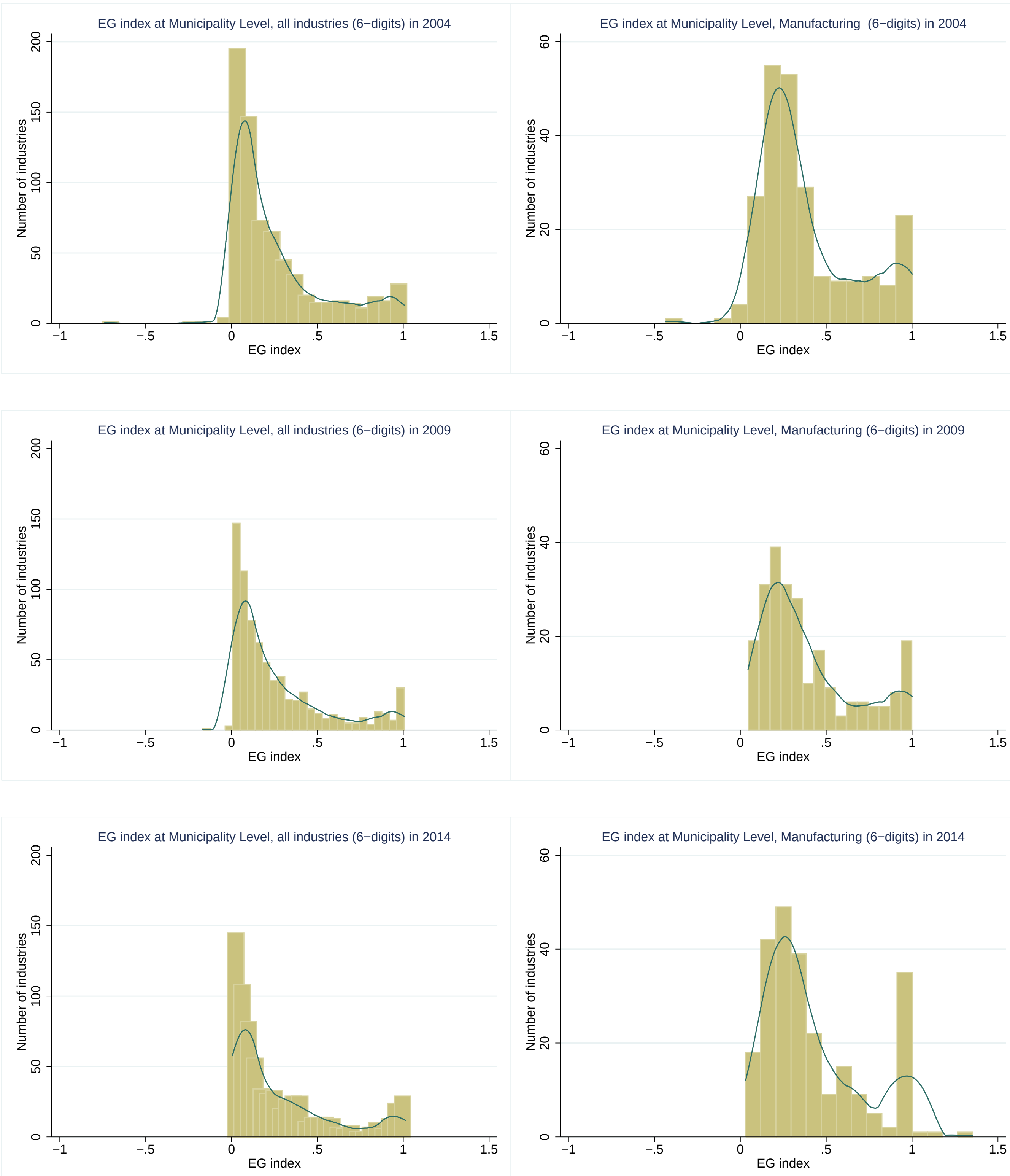
Notes: Mean and median values for 248 (resp. 245 in 2004, and 238 in 2009) 4-digit and 6-digit NAICS classification of industries. Share< 0 means 'not clustered'. Share$\in (0, 0.05]$ means 'weakly clustered'. Share> 0.05 means 'strongly clustered'. See Ellison and Glaeser (1997) for details.

Figure 4 plots the distributions of the EG indices for the 6-digit manufacturing industries in 2004, 2009, and 2014. As can be seen, these distributions are skewed towards 0.05, which shows that many industries are highly agglomerated, whereas only few of them are weakly agglomerated (the EG index is positive but smaller than 0.05).

3.5.2 Trends in Mexico Specialization : 2004 – 2014

We now examine how specialization has evolved between 2004 and 2014. We compute the Krugman's specialization index at the 4-digit NAICS level (see Appendix B for details) and consider as highly specialized localities for which the specialization index is higher than 0.75 whereas localities for which the specialization index is below 0.35 can be considered as weakly specialized. Our results, shown in Table 4, can be summarized as follows. First, localities are

Figure 4: Distribution of the EG Index at the locality level (NAICS 6-digit) in 2004 (top panel), 2009 (middle panel), and 2014 (bottom panel). All industries (left panel) and manufacturing (right panel).



less specialized in terms of employment than in terms of output. The average output specialization index is on average 10 to 12 percent higher than the average employment specialization index. Second, there is an increasing trend toward specialization over the period, both in terms of output and employment specialization.

Table 4: Mean and median specialization indices in Mexico : 2004 – 2014

	2004		2009		2014	
	Employment	Output	Employment	Output	Employment	Output
Mean	0.117	0.126	0.155	0.174	0.199	0.209
Median	0.014	0.030	0.021	0.063	0.046	0.087
Minimum	0.000	0.000	0.000	0.000	0.000	0.000
Maximum	1.972	1.978	1.963	1.983	1.984	1.973
ksi < 0.35	91.750	90.72	88.90	88.20	85.20	84.73
ksi $\in (0.35, 0.75]$	2.200	3.28	2.71	4.15	4.19	6.05
ksi $>$ 0.75	6.040	6.00	8.40	7.65	10.61	9.22

Notes: Mean and median values of the Krugman Specialization Index (KSI) for the 316 metropolitan area and 1,832 standalone municipalities in Mexico. KSI < 0.35 means 'not specialized'. KSI $\in (0.35, 0.75]$ means 'weakly specialized'. KSI > 0.75 means 'highly specialized'.

Figure 5 maps the employment and output specialization indices. These maps show the increasing trends in Mexico towards both industrial concentration and locality specialization. To identify the factors that drive these changes in the employment and specialization of localities, we will resort to a multivariate analysis.

4 Econometric specification and identification issues

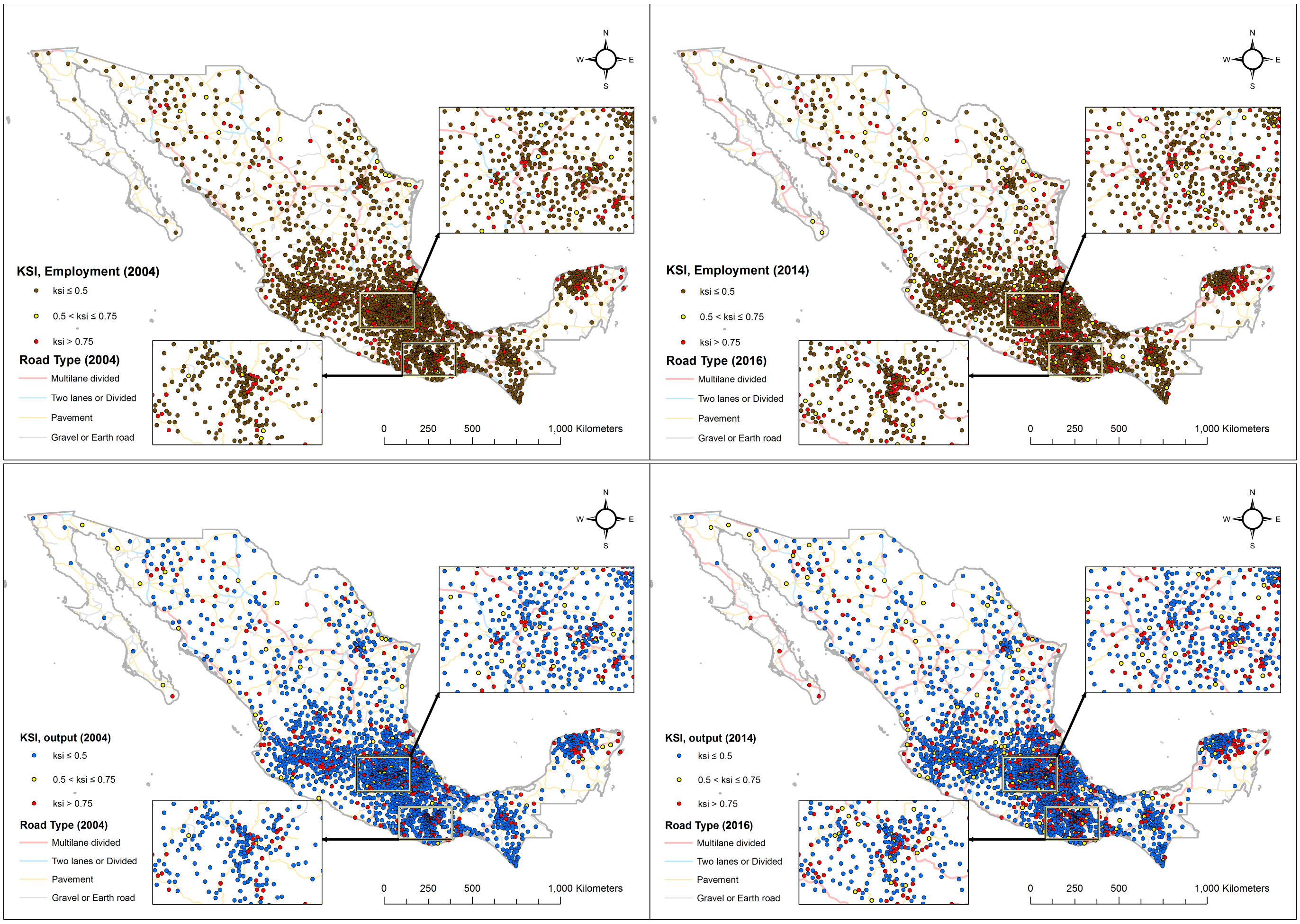
4.1 Econometric specification

We take advantage of the panel nature of our data to estimate the effects of road improvements on the localization of economic activities, and the specialization of localities in Mexico. Our main empirical specification is as the following 'market access' model:

$$Y_{m,t} = \beta_M M_{m,t} + \beta_C^{int} M_{m,t} I_{m,t} + \beta_C C_{m,t} + \alpha_t + \mu_m + \epsilon_{m,t} \quad (3)$$

where Ym, t is the dependent variable (employment, specialization index, and localization index) for locality m, $M_{m,t}$ is the market access or market potential of the same locality, $C_{m,t}$ is a vector of time-varying locality characteristics (education, population, oil-reserves, and pre/post NAFTA period), $I_{m,t}$ is an interaction term (education, population, oil-reserves, pre/post NAFTA period, and capital city dummy), α_t is a time dummy and μ_m is the location fixed effect, which

Figure 5: Maps of employment and output locality specialization in Mexico (2004 – 2014)



Sources: DENUE (INEGI). *Note*: Employment specialization in 2004 (top left) and 2014 (top right). Output specialization in 2004 (bottom left) and 2014 (bottom right).

absorbs time-invariant location characteristics. $\epsilon_{m,t}$ is the error term. The coefficients of interest are β_M and β_C^{int}, which account for the effect of accessibility and of its interaction with controls. To avoid endogeneity with income and population growth, we create a dummy variable for the above and below the initial (1986) median education and population level when interacting market access and market potential with education and population characteristics.

Specifically, we regress the log value of employment and/or of the specialization index of locality m and year t on the log of the market access (MA) or of the market potential (MP), so that our coefficients estimates can be interpreted as elasticities. Standard errors are clustered at the locality level to adjust for heteroskedasticity and within-locality correlation over time. The regression sample is a balanced panel of 2,094 localities with employment data for the years 1986, 1994, 1999, 2004, 2009, and 2014. The specialization and concentration indices are calculated at the establishment and locality levels. Since the latter micro-geographic data only span the 2004-2014 period, we will therefore restrict our analysis on specialization and concentration to this time period.

We estimate three specifications based on equation (3). These specifications differ by the inclusion of the interaction of market access or market potential with various local characteristics such as the education rate (dummy equal to 1 if the locality has above median average education), population size, the pre and post NAFTA period, a capital city dummy, and whether oil is produced in the locality. We also add two sets of locality controls (median education dummy education and population size).

We also estimate an 'infrastructure model' where market access is proxied by our measure of efficient roads within a fixed radius surrounding a locality's centroid. The latter model makes it possible to contrast the impact of access to domestic and international markets (measured by the minimum time or travel cost to a port or port of entry into the U.S.).[15]

Table 5 presents the summary statistics for the key variables of our analysis and a decomposition of the between and within component of the standard deviation.[16]

4.2 Identification issues

The panel structure of our data allows us to control for location-specific time-invariant factors and general macroeconomic trends. However, we need to accounts for the the three following problems: (i) omitted variables bias, (ii) the non-random road placement, and (iii) the structural recursion problem of our market access indicators.

Omitted variables bias. In our framework, the problem of omitted variables is mitigated by the panel structure of our data. We include locality fixed effects in the panel estimation, which absorbs all time-invariant local characteristics such as initial wealth. Therefore, the estimated effects of market access and market potential cannot be attributed to locality differences. This is an advantage over a cross-sectional framework, where unobservable location characteristics would likely be correlated both with population or income and transport infrastructure. Despite the panel nature of our data, and the different controls included, our model is still subject to problem of simultaneity bias or reverse causality. There are two main empirical challenges with our model specification: the non-random placement of roads and a recursion problem inhering with regression on market access or market potential measures. We discuss below how we address these issues.

Non-random road placement and the construction of the 'doughnut' IV. The expansion of the road network surrounding a locality can be endogenous and this may create a correlation between increases in local market access or local market potential and our left-hand side

[15] It is not desirable to introduce our measures of access to external markets in the market access specification because of their high correlation with our market access or market potential indicators.

[16] The latter decomposition confirms the presence of time variation in our data (although most of the variation remains cross-sectional).

Table 5: Key variables and summary statistics

Variables	Observations	Mean	Standard Deviation Overall	Standard Deviation Between	Standard Deviation Within
Total employment	12,564	6,812.303	90,112.890	85,792.800	27,619.900
Manufacturing employment	12,564	1,998.672	23,597.920	23,277.810	3,901.393
Commerce employment	12,564	2,149.258	26,322.070	25,274.430	7,369.452
Service employment	12,564	2,229.871	37,222.180	33,232.070	16,779.740
Mining employment	12,564	62.447	1,149.786	642.266	953.764
Market access	12,564	22.496	399.887	396.782	50.361
Market potential	10,470	365,276.400	573,321.300	430,809.700	378,379.700
Employment KSI	5,286	0.237	0.445	0.359	0.266
Output KSI	5,286	0.305	0.495	0.397	0.303
Education	12,408	0.046	0.049	0.034	0.035
Total population	12,564	47,149.370	445,092.100	442,553.500	48,283.380
Efficiency roads within 200 km	11394	12.440	0.419	0.411	0.081
Minimum travel time to port	12,544	6.885	0.551	0.544	0.090
Minimum travel time to border	12,544	6.300	0.543	0.531	0.113
Minimum travel cost to port	12,544	0.795	0.136	0.134	0.024
Minimum travel cost to border	12,544	0.648	0.121	0.119	0.020

Note: the above descriptive statistics are calculated on a sample of 12,564 observations covering 2,094 localities over 6 different dates. The standard deviation is decomposed into between and within components, which measure the cross sectional and the time series variation, respectively.

variable, for instance employment. This may occur if road construction occurs in localities that would otherwise have experienced relative increases or decreases in average employment (our review or road investments in Mexico suggest that roads may have been built to increase accessibility in lagging regions during the studied period). To deal with that problem, we adopt the so-called 'doughnut' strategy as proposed by Jedwab and Storeygard (2017). The basic idea behind this approach is to instrument the market access or market potential with similar measures excluding surrounding areas within a fixed radius (in addition to 'own locality' exclusion). Variations in this instrument are more likely to be exogenous. Following the identification strategy of Jedwab and Storeygard (2016) also implemented by Blankespoor et al. (2017), we construct the counterfactual measures of market access and market potential excluding all localities j located within a 25, 50 or 75 km radius of locality i, hence the reference to a 'doughnut'.[17]

[17]Because of the mixed nature of our sample, for metropolitan areas, we first compute these counterfactual measures at the municipality, excluding both municipalities within a same metropolitan area and municipalities within our distance threshold. We then define the metropolitan area 'doughnut' as the weighted average of these municipality indices, using municipality area as weights.

The structural endogeneity of market access and market potential. This is also called the structural recursion problem. If considering on the left-hand side a variable which is correlated with other variables that enter the calculation of the accessibility regressor in formulas (1) or (2) – for instance if regressing total employment on market access or market potential in Equation (3) – a recursion problem is bound to emerge. This is the case even when excluding 'own locality' from the measure of accessibility. To see this, notice that locality i's market access is a function of locality j's population, which in turn is correlated with locality j's employment, which is a function of locality j's market access, and thus of locality i's population. Locality i's employment is then structurally correlated with its own market access in the absence of any causal effect of market access on employment. Following Baum-Snow et al. (2017), we address this problem by calculating either counts of road intersections or a 'doughnut' of efficient road lengths in a given radius (in a similar way to the market access and market potential 'doughnuts').[18] We then use the road count or efficient road variable instead of the market access or market potential indicator, or use it to instrument market access or market potential. Sources of variations in 'accessibility' are then only due to variations in roads, thus avoiding the recursion problem.

The validity of the instrument (for market access and market potential) is tested by means of the Kleibergen-Paap statistics which is robust in the presence of heteroskedasticity. After performing this test, our result excludes the hypothesis of weak instruments in both cases as it exceeds the rule of thumb suggested by Staiger and Stock (1997) that the statistic must be larger than 10. As for the exogeneity of the instruments, in both cases we rely on the Hansen-J statistics which, in both cases, strongly accepts the exogeneity hypothesis of the instruments.[19] Finally, it is worthwhile noticing that, besides being efficient, our estimation results are also consistent with respect to either heteroskedasticity or autocorrelation because of the Stock-Yogo weak ID specification employed for the estimation. In all our IV regressions, the first stage F-statistic is large, suggesting that the instrument is strong (Stock and Yogo, 2005). The coefficients on market access and market potential are positive and statistically significant in all second stage regressions.

To differentiate the role of access to domestic and external markets, we also estimate the infrastructure model as previously discussed. In this model, domestic market access is proxied by efficient road length[20] and access to external markets by the minimum travel cost or travel time to a major port and to the U.S. border. One key advantage of this approach is that the infrastructure measure (efficient road length) does not have a structural dependence on population (like the market access) or on income (like the market potential). However, this infrastructure measure may still suffer from non-random placement. To deal with this

[18] In this 'doughnut', the exclusion of surrounding areas addresses endogeneity of road placement.

[19] See Baum et al. (2017) for a detailed explanation of test implementation in STATA and for references.

[20] Efficient road length is calculated for different radii: 50, 75, 100, 150, 200, and 300 km.

problem, we use as an instrument the 'doughnut' road efficient length. We also report the results from cross-section regressions where the efficient road length of the corresponding year is instrumented with the 1949 efficient road length.[21]

5 Estimation Results

Our 1986 – 2014 data allows us to exploit changes in the transportation network over time to estimate the effect of market access and market potential on localities employment and specialization in Mexico.

5.1 The effect of market access and market potential on employment

Table 6 reports estimates from the regression of employment on market access, in which our market access measure is instrumented in three ways: using the doughnut market access, the road count within a 10 km circle, and the efficient road length within a 100km radius. The first column (MA only) reports our basic estimates without any controls. As can be seen, the coefficient on the market access is positive and significant. There is a positive and significant causal effect (IV) of market access on employment. First stage coefficients are positive and significant, which means that each market access measure is predicted by the appropriate instrument.[22] Results indicate that between 1986 and 2014, a 10 percent increase in market access results in a 1.6 to 2.1 percent increase in employment. To avoid potential endogeneity of our control variables, we use as interaction terms population and education variables (above mean dummies) set at their initial (1986) levels.

More urbanized locations (localities which had a population above the 1986 median population) appear to benefit more from an increase in market access, while areas with a more educated population benefit less from an increase in market potential. The estimates also provide evidence that the positive effect of improved domestic market access on employment is partially attenuated after the NAFTA agreement (the net effect is still positive). One key explanation of this result that after the NAFTA agreement, what is important for localities is to improve their access to international market, especially the northern localities given their proximity to the U.S. market. Therefore, our result can explain the substitution effect of domestic versus international market access. This effect cannot be capture by our our market access and market potential indicators, which by construction only estimates domestic market access and does not take into account potential market size effects across national boundaries. Finally, an

[21] 1949 is the earliest date for which we could recover information on road extent and road types from paper maps. This historic measure is likely to be exogenous to economic outcomes observed more than 40 years later.

[22] These first stage results are not reported in the paper but are available upon request.

improved market access has a beneficial and statistically significant large effects on oil-based localities.

Table 6: The effect of market access on employment

	Total employment				
Variables	MA only	OLS	IV (Doughnut)	IV (Road count)	IV (Efficient roads)
Market Access (MA)	0.176***	0.149***	0.208***	0.163***	0.170***
	(0.0153)	(0.0111)	(0.0144)	(0.0202)	(0.0219)
MA x Metropolitan dummy		0.0717***	0.0528***	0.0629***	0.0574***
		(0.0117)	(0.0146)	(0.0154)	(0.0164)
MA x Education86		-0.0468***	-0.0486***	-0.0474***	-0.0452***
		(0.00334)	(0.00403)	(0.00329)	(0.00339)
MA x NAFTA dummy		-0.114***	-0.119***	-0.115***	-0.117***
		(0.00233)	(0.00251)	(0.00213)	(0.00219)
MA x Capital city		-0.0225	-0.0294	-0.0228	-0.0237
		(0.0226)	(0.0390)	(0.0250)	(0.0251)
MA x Oil dummy		0.839**	0.756*	0.832**	0.809**
		(0.385)	(0.403)	(0.399)	(0.401)
Above median education		0.155***	0.124**	0.155***	0.146***
		(0.0581)	(0.0550)	(0.0522)	(0.0508)
Above median population		0.851***	0.736***	0.805***	0.779***
		(0.0803)	(0.0823)	(0.0918)	(0.0959)
Constant	6.836***	5.790***			
	(0.0844)	(0.0755)			
Observations	11379	11379	9778	11251	10293
Adj. R-squared	0.052	0.423	0.398	0.424	0.440

Note: * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level. The doughnut IV is calculated by excluding all localities within a 25km circle. The road count is the number of roads intersecting a circle of 10km radius, and efficient road length is calculated within a 100 km radius.

Table 7 reports estimates from the regression of employment on market potential, using OLS and IV. Overall, we find a positive and significant causal effect of market potential on employment. The first column (MP only) reports our basic estimates without any controls and the estimated coefficient are positive and statistically significant. In words, increasing market potential within localities are associated with an increase in their employment level, and the effects are large. A ten percent increase in market potential is associated with a 2.9 to 6.5 percent increase in total employment. More urbanized locations appear to benefit less from an increase in market potential, whereas areas with more educated population benefit more from an increase in market potential. These results are at odds with what we found in the market access regression.

5.2 Heterogeneous effects of market access and market potential

Market access and market potential have heterogeneous effects on employment in different sectors and economic regions. We did a breakdown of the total employment and report the

Table 7: The effect of market potential on employment

Variables	Total employment MP only	OLS	IV (Doughnut)	IV (Road count)	IV (Efficient roads)
Market Potential (MP)	0.192***	0.351***	0.322***	0.292**	0.645***
	(0.00324)	(0.0401)	(0.0484)	(0.148)	(0.0342)
MP x Metropolitan dummy		-0.0180***	-0.00188	-0.00954	-0.0680***
		(0.00652)	(0.00822)	(0.0232)	(0.0144)
MP x Education94		0.0168***	0.0188***	0.0195***	0.000723
		(0.00207)	(0.00286)	(0.00705)	(0.00270)
MP x NAFTA dummy		-0.0591***	-0.0557***	-0.0424	-0.141***
		(0.0127)	(0.0140)	(0.0423)	(0.00928)
MP x Capital city		-0.00822*	-0.0113**	-0.0108**	-0.0109*
		(0.00430)	(0.00450)	(0.00427)	(0.00646)
MP x Oil dummy		0.0227	0.0168	0.0248	0.0135
		(0.0157)	(0.0190)	(0.0172)	(0.0229)
Above median education		0.187***	0.151***	0.185***	0.203***
		(0.0417)	(0.0404)	(0.0374)	(0.0359)
Above median population		0.406***	0.219**	0.302	0.973***
		(0.0872)	(0.101)	(0.267)	(0.171)
Constant	3.749***	2.134***			
	(0.0382)	(0.334)			
Observations	10229	10229	8168	10098	9248
Adj. R-squared	0.463	0.503	0.511	0.503	0.471

Note: * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level. The doughnut IV is calculated by excluding all localities within a 25km circle. The road count is the number of roads within a 10km radius.

estimate results for the three main sectors namely services, manufacturing, and commerce. The results from our IV estimations are reported in Table 8. The first three columns report the estimated coefficients for market potential and the last three columns for the market access. Compared to other sectors, it is the service sector that benefits more from an increase in market access or market potential.[23] A 10 percent increase in market potential is associated with 1.2 percent, 3.0 percent, and 4.1 percent increases in employment in the manufacturing, commerce, and services sectors respectively. Similarly, a 10 percent increase in market access is associated with a 1.6, 1.8, and 2.5 percent increase in manufacturing, commerce, and services employment respectively.

Results also show that more urbanized or more educated localities benefit more from an increase in market potential, and less from an increase in market access.[24] Finally, localities seem to benefit less from an increase in domestic market access or domestic market potential over the NAFTA period. As mentioned earlier, the net effect is positive, and the result can simply explain the substitution effect of domestic versus international market access.

Table 15 and Table 16 in Appendix C shows our results (when we control for endogeneity issues), both for the effect of market access and market potential on total employment. As

[23] The services sector account for about 60 percent of Mexico GDP

[24] This result is consistent with the use of a different human capital proxy, locality literacy.

Table 8: Sectoral heterogeneous effects of market access and market Potential

Variables	Market Potential: Employment			Market Access: Employment		
	Manufacturing	Commerce	Services	Manufacturing	Commerce	Services
Market Potential or Market Access	0.119*	0.303***	0.413***	0.158***	0.176***	0.246***
	(0.0623)	(0.0418)	(0.0807)	(0.0184)	(0.0125)	(0.0807)
MP or MA x Metropolitan dummy	0.0332**	0.00657	-0.0148	0.0363*	0.0711***	-0.0148
	(0.0139)	(0.00681)	(0.0161)	(0.0193)	(0.0133)	(0.0161)
MP or MA x Education86	0.0266***	0.0195***	0.0217***	-0.0399***	-0.0407***	0.0217***
	(0.00451)	(0.00241)	(0.00518)	(0.00590)	(0.00342)	(0.00518)
MP or MA x NAFTA dummy	-0.0146	-0.0596***	-0.0744***	-0.0850***	-0.101***	-0.0744***
	(0.0173)	(0.0122)	(0.0221)	(0.00356)	(0.00206)	(0.0221)
MP or MA x Capital city	-0.00819*	-0.0117***	-0.00924	-0.0470**	-0.0121	-0.00924
	(0.00427)	(0.00379)	(0.00568)	(0.0228)	(0.0363)	(0.00568)
MP or MA x Oil dummy	-0.0200	0.0163	-0.000259	-0.391	0.683***	-0.000259
	(0.0237)	(0.0253)	(0.0276)	(0.485)	(0.205)	(0.0276)
Above median education	0.197***	0.125***	0.152***	0.234***	0.0795*	0.152***
	(0.0648)	(0.0295)	(0.0492)	(0.0714)	(0.0441)	(0.0492)
Above median population	-0.299*	0.166*	0.438**	0.513***	0.829***	0.438**
	(0.171)	(0.0863)	(0.192)	(0.102)	(0.0779)	(0.192)
Observations	7839	8144	7726	9165	9747	7726
Adj. R-squared	0.168	0.534	0.412	0.162	0.403	0.412

Note: * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level. The reported results are the doughnut IV which is calculated by excluding all localities within a 25km circle. Results are robust to others IV strategies.

can be seen, the North and the Capital states benefits more for the observed improvements in Mexico transportation network . Interestingly, there is no significant effect of an increase in the market potential in the North region, while the effect is important for the South region (which is a less-developed region). Despite the low level of infrastructure development in the South region, our result shed light on the potential to enhance productivity and economic opportunities through investment in transport infrastructure in lagging regions. This results is also in line with the evidence that the south region has a greater potential to generate positive spatial spillovers (see Deichmann et al., 2004, Baylis et al., 2012 and Alvarez et al., 2017).[25]

5.3 The effect of market access and market potential on specialization

Table 9 presents the respective coefficient estimates of the causal effect of market potential and market access on locality specialization. First, better access to market and better market potential have a positive and significant effect on specialization. The causal effects of the increase in market access are larger than that of the market potential. A 10 percent increase in market access had a 7 percent increase in output specialization and a 3.4 percent in employment

[25]Despite the observed increase in Mexico, the per capita income in the northern states is two or three times higher than in the southern states, and the disparities in other social and infrastructure indicators are even more dramatic.

specialization. While an increase in market potential has no significant causal effect on output and employment specialization. Second, a better market access and market potential in more urbanized locations leads to less specialization. However, this effect is not significant when we control for the potential endogeneity of road placement.

Table 9: The effect of market access and market potential on specialization

	Krugman Specialization index: Output				Krugman Specialization index: Employment			
	Market Potential		Market Access		Market Potential		Market Access	
	OLS	IV (Doghnuts)	OLS	IV (Doghnuts)	OLS	IV (Doghnuts)	OLS	IV (Doghnuts)
Market Potential or Market Access	2.913***	0.154	0.455***	0.704***	1.884***	0.984	0.321**	0.339**
	(0.494)	(1.863)	(0.176)	(0.231)	(0.436)	(0.815)	(0.128)	(0.160)
MP or MA x Metropolitan dummy	-2.245***	0.464	-0.127	-0.286	-1.433***	-0.590	-0.0410	-0.0833
	(0.507)	(1.772)	(0.146)	(0.187)	(0.441)	(0.820)	(0.0942)	(0.0995)
MP or MA x Education86	0.0694	0.172**	-0.259***	-0.291***	0.0752*	0.122**	-0.269***	-0.214**
	(0.0448)	(0.0774)	(0.0768)	(0.0919)	(0.0432)	(0.0528)	(0.0798)	(0.105)
MP or MA x Capital city	0.00864	0.00877	0.0458*	0.0837***	-0.00180	-0.00183	0.0955***	0.102**
	(0.00871)	(0.00988)	(0.0239)	(0.0304)	(0.0138)	(0.0135)	(0.0170)	(0.0439)
MP or MA x Oil dummy	-0.423	-0.577	1.009	0.950	-0.243	-0.233	1.785	1.756
	(0.494)	(0.470)	(2.109)	(1.945)	(0.458)	(0.538)	(1.968)	(2.478)
Above median education	1.313***	1.264***	1.214***	1.286***	0.492	0.414	0.231	0.508
	(0.353)	(0.431)	(0.390)	(0.418)	(0.403)	(0.424)	(0.423)	(0.390)
Above median population	29.22***	-4.382	1.503	0.994	18.32***	7.713	0.780*	0.476
	(6.422)	(22.33)	(1.017)	(0.966)	(5.581)	(10.35)	(0.472)	(0.447)
Constant	-42.77***		-4.412***		-29.06***		-4.017***	
	(6.165)		(0.985)		(5.454)		(0.504)	
Observations	4599	3628	4303	3628	4491	4149	4203	3556
Adj. R-squared	0.0649	0.0498	0.0234	0.0233	0.0385	0.0381	0.0200	0.0106

Note: * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level. The reported results are the doughnut IV which is calculated by excluding all localities within a 25km circle. Results are robust to others IV strategies.

5.4 The effect of access to infrastructure on employment and specialization

We now turn to the estimation of the impact of change in roads improvement using a measures of access to infrastructure that do not involve population or income. this allow us to avoid the structural endogeneity of market access and market potential. Following Baum-Snow et al. (2017) and Emran and Hou (2013), we regress our dependent variable – employment and specialization – on measures of access to domestic and international markets. For access to external markets, we calculate two complementary measures: the minimum travel time and the minimum travel cost to a major port and to the U.S. border.[26] The domestic market access is proxied by efficient road length within a circle of 200 km radius outside the locality itself. To correct for the non-random placement of roads, we also instrument our domestic market access applying the doughnut approach. For this we calculate efficient road length within a 300 km

[26]Minimum travel cost is calculated using the same calibrated trade cost function as in the market potential formula.

radius while removing roads with a smaller circle of radius of 50, 100 or 200 kilometers. We also run a cross-section regression where our infrastructure metric is instrumented by the 1949 efficient road length in order to address the possible non-random placement of roads.

As in the previous section, first stage estimated coefficients are positive and significant, which means that our efficient road length measure is predicted by the appropriate instrument. Table 10 reports the estimates in the regression of employment on infrastructure, where efficient road length is instrumented using the doughnut approach. Better access to domestic and international markets (lower travel time or lower travel cost) has a positive and statistically significant effect on employment. These effects are large: a 10 percent increase in efficient road length leads to a 7.6 percent increase in total employment. They are also heterogeneous across sector: employment in the commerce sector (8.1 percent) benefits more than in the services sector (7.0 percent) or in manufacturing (5.0 percent).

Our results are consistent with Emran and Hou (2013) who used a rural sub-sample of Chinese households to document the importance of domestic and international market access on per-capita income of rural households in China. However, these results are at odds with Baum-Snow et al. (2017). As they argued, the counter-intuitive result with what would have been expected in an environment of free mobility are attributed to the migration restrictions under the Hukou system. However, given that they are looking at a long difference in a cross-sectional framework, we can test whether a panel structure framework is more adequate to assess the importance of transport improvement on local economic outcomes. [27]

Looking at the effects of road infrastructure on specialization, Table 11 shows that better access to domestic and international markets has a positive and statistically significant causal effect on specialization. A 10 percent increase in efficient road length leads to a 13.8 percent increase in locality output specialization, and a 8 percent increase in locality employment specialization. Results also indicate that, when we control for potential endogeneity problems, access to international markets has no significant effect on output and employment specialization.

6 Conclusion

Transport investments have the potential to stimulate growth through trade, structural transformation, agglomeration, and productivity. This paper uses extensive micro-geographic data and geo-referenced digitized maps of the transport network in Mexico over three decades to provide empirical evidence of the causal effect of improved access to markets following road improvements on local employment and specialization. In addition to shedding light on the

[27] To test this hypothesis, we estimate cross-section regressions instrumenting efficient road length with its 1949 level. As can be seen from Table 14 in Appendix C, the coefficients on efficient road length are negative and statistically significant using both OLS and IV, and across years.

Table 10: The effect of road infrastructure on employment

Variables	Total employment OLS	Total employment IV	Manufacturing OLS	Manufacturing IV	Services OLS	Services IV	Commerce OLS	Commerce IV
Road Efficiency Units	5.494***	7.607***	3.559***	5.007***	5.008***	6.979***	5.817***	8.092***
	(0.0939)	(0.109)	(0.144)	(0.166)	(0.0791)	(0.0872)	(0.109)	(0.125)
Min travel cost to border	-3.296***	2.084***	-3.831***	-0.00484	-3.352***	1.660***	-4.327***	1.296***
	(0.431)	(0.442)	(0.670)	(0.681)	(0.363)	(0.348)	(0.467)	(0.482)
Min travel cost to port	1.472***	-0.485	2.143***	0.665	1.675***	-0.145	1.542***	-0.343
	(0.461)	(0.414)	(0.726)	(0.660)	(0.380)	(0.342)	(0.519)	(0.475)
Above median education	0.224***	0.250***	0.339***	0.357***	0.159***	0.183***	0.129***	0.158***
	(0.0460)	(0.0419)	(0.0682)	(0.0631)	(0.0363)	(0.0327)	(0.0475)	(0.0441)
Above median population	0.427***	0.384***	0.315***	0.293***	0.415***	0.375***	0.419***	0.376***
	(0.0448)	(0.0412)	(0.0782)	(0.0667)	(0.0419)	(0.0371)	(0.0575)	(0.0516)
Constant	-60.99***		-38.45***		-55.82***		-65.70***	
	(1.337)		(1.975)		(1.135)		(1.597)	
Observations	11201	11201	10620	10615	11185	11185	10584	10576
Adj. R-squared	0.549	0.507	0.210	0.195	0.607	0.559	0.497	0.459

Note: * denotes significance at the 10% level, ** denotes significant at the 5% level, and *** denotes significance at the 1% level. The doughnut IV is calculated by excluding all localities within a circle of 200km radius.

Table 11: The effect of road infrastructure on specialization

Variables	Output Specialization OLS	Output Specialization IV	Employment Specialization OLS	Employment Specialization IV
Road Efficiency Units within 200 km	6.297***	13.75***	2.576**	8.024***
	(1.177)	(2.796)	(1.181)	(2.587)
Minimum travel cost to border	-5.287	4.458	-7.207**	-0.0656
	(3.564)	(4.696)	(3.647)	(4.852)
Minimum travel cost to port	5.453	0.712	10.54***	6.858*
	(3.583)	(3.923)	(3.641)	(4.016)
Above median education	1.124***	1.179***	0.332	0.366
	(0.358)	(0.368)	(0.406)	(0.420)
Above median population	1.183	1.061	0.512	0.456
	(0.747)	(0.712)	(0.599)	(0.590)
Constant	-82.91***		-37.48**	
	(15.30)		(15.59)	
Observations	4556	4367	4478	4274
Adj. R-squared	0.019	0.009	0.007	0.001

Note: * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level. The doughnut IV is calculated by excluding all localities within a circle of 200km radius.

effects of transport infrastructure (roads) on local economic development, this paper also provides evidence of the respective roles of domestic and external market access on the geography of economic activities.

To estimate the causal effect of roads on the geographical concentration of economic activity and on the specialization of localities, this paper addresses two important endogeneity issues, namely the potentially non-random placement of roads and the structural endogeneity of market access and market potential (recursion problem). For this, we adopt three separate identification strategies. We use the so-called 'doughnut' IV strategy to address non-random road placement of roads, and we address the structural endogeneity of market access and market potential by using a raw measures of access to infrastructure that do not involve population or income, either as substitutes for market access indicators or as instruments.

Finally, our empirical exercise sheds light on the causal effect of road improvements on the increasing trends in specialization and geographical concentration of employment in Mexico and confirms that transport investments can have large effects on local economic activity through improved accessibility. The quantitative effects are large. A 10 percent increase in market access results in a 1.6 to 2.1 percent increase in total employment, a 7 percent increase in output specialization, and a 3.4 percent increase in employment specialization. In a country where the share of the manufacturing sector has been declining, we find heterogeneity in sectoral effects, with employment in commerce and services benefiting more from road improvements than employment in manufacturing. We also provide evidence of the potential benefits of the investments in transport infrastructures in lagging regions.

Further investigation, however, will be needed to understand the mechanisms at play. For example, we do not know whether improved connectivity (through new or better roads) plays a role through improved accessibility to output markets, to intermediate inputs, to the pool of workers, or simply through reduction in travel times. We leave these important issues for further research.

References

[1] Abrams, A., Lozano-Gracia, N., and C. Oram. 2016. Deblluring Nightime Lights. *Unpublished manuscript*.

[2] Addison, D. and B. Stewart. 2015. Nighttime Lights Revisited: the Use of Nighttime Lights Data as a Proxy for Economic Variables. *World Bank Policy Research Working Paper 7496*.

[3] Alder, S. 2017. Chinese Roads in India: The Effects of Transport Infrastructure on Economic Development. *Unpublished manuscript*.

[4] Alder, S., Roberts, M. and M. Teewari. 2017. The Effect of Transport Infrastructure on India's Urban and Rural Development. *Unpublished manuscript*.

[5] Alvarez, A., Garduno-Rivera, R., and Nunez, H. M. 2017. Mexico's North-South divide: The regional distribution of state inefficiency 1988 – 2008. *Papers in Regional Science*.

[6] Anas, A. and K. Xiong. 2003. Intercity Trade and the Industrial Diversification of Cities, *Journal of Urban Economics*, 54, 258-276.,

[7] Asher, S. and P. Novosad. 2016. Market Access and Structural Transformation: Evidence from Rural Roads in India. *Unpublished manuscript.*

[8] Becerril-Torres, O., Àlvarez-Ayuso, I. and L. Del moral-Barrera. 2010. Do Infrastructures Influence the Convergence of Efficiency in México? *Journal of Policy Modeling*, 32(1), 120-137.

[9] Baum-Snow, N., 2007. Did Highways Cause Suburbanization? *The Quarterly Journal of Economics*, 122(2), pp.775-805.

[10] Baum-Snow, N., 2017. Urban Transport Expansions, Employment Decentralization, and the Spatial Scope of Agglomeration Economies. *Unpublished paper.*

[11] Baum-Snow, N., Henderson, J.V., Turner, M.A., Zhang, Q., and L. Brandt. 2017. Roads, Railroads, and Decentralization of Chinese Cities. *Review of Economics and Statistics* 99(3), 435-448.

[12] Baylis KR, Garduno-Rivera R, Piras G. 2012. The distributional effects of NAFTA in Mexico: Evidence from a panel of municipalities. *Regional Science and Urban Economics*. 42: 286 – 302.

[13] Bazant, J. 1977. A Concise History of Mexico from Hidalgo to Cardenas 1805-1940. *New York: Cambridge University Press.*

[14] Becerril-Torres, O. , Álvarez-Ayuso, I. and L. del Moral-Barrera. 2010. Do Infrastructures Influence the Convergence of Efficiency in México?, *Journal of Policy Modeling*, 32(1), 120-137.

[15] Behrens, K., Bougna, T. and M.W. Brown. 2015. The World is not yet Flat: Transportation Costs Matter!' *CEPR Discussion Papers, DP 10356.*

[16] Behrens, K. and T. Bougna. 2015. An Anatomy of the Geographical Concentration of Canadian Manufacturing Industries, *Regional Science and Urban Economics*, 51, 47-69.

[17] Berg, C., Deichmann, U., Liu, Y. and H. Selod. 2017. Transport policies and development, *Journal of Development Studies*, 53(4), 465-480, 2017.

[18] Bess, M. 2014. Routes of Conflict : Building Roads and Shaping the Nation of Mexico, 1941-1952, *The Journal of Transport History*, 35(1), 78-96.

[19] Bess, M. 2016a. 'Neither Motorists nor Pedestrians Obey the Rules : Transit Law, Public Safety, and the Policing of Northern Mexico's Roads, 1920s-1950s." *The Journal of Transport History*, 37(2), 155-174.

[20] Bess, M. 2016b. Revolutionary Paths : Motor Roads, Economic Development, and National Sovereignty in 1920s and 1930s Mexico. Mexican Studies / Estudios Mexicanos, 32(1), 56-82.

[21] Bess, M. 2017. Routes of Compromise: Building Roads and Shaping the Nation of Mexico, 1917- 1952. *University of Nebraska Press*, 246 pages.

[22] Bird, J. and S. Straub. 2014. The Brasilia Experiment: Road Access and the Spatial Pattern of Long-Term Local Development in Brazil. *World Bank Policy Research Paper*, WPS 6964.

[23] Blankespoor, B., Mesplé-Somps, S., Selod, S. and G. Spielvogel. 2017. Roads and Structural Transformation in Mali. *Unpublished paper.*

[24] Bosker, M., Deichman, U. and M. Roberts. 2015. Hukou and Highways: The Impact of China's Spatial Development Policies on Urbanization and Regional Inequality. *World Bank Policy Research Working Paper* WPS7350.

[25] Bougna, Théophile. 2016. The Determinants of Localization: A Conditional Distance-Based Approach. *Unpublished paper.*

[26] Brock, G. and B. German-Soto. 2013. Regional Industrial Growth in Mexico: Do Human Capital and Infrastructure Matter? *Journal of Policy Modeling*, 35(2), 228-242.

[27] Brulhart, M. and F. Sbergami. 2009. Agglomeration and Growth: Cross-Country Evidence, *Journal of Urban Economics*, 65, 48-63.

[28] Calderón, C., Moral-Benito, E., and L. Servén. 2015. Is Infrastructure Capital Productive? A Dynamic Heterogeneous Approach, *Journal of Applied Econometrics*, 30(2), 177-198.

[29] Calderón, C. and L. Servén. 2010. Infrastructure and Economic Development in Sub-Saharan Africa, *Journal of African Economies*, 19, AERC Supplement 1, i13-i87.

[30] Ciccone, A. 2002. Agglomeration Effects in Europe, *European Economic Review*, 46(2): 213-227.

[31] Dávila A. 2008. Los clústeres industriales del noreste de Mèxico (1993-2003). Perspectivas de desarrollo en el marco de una mayor integración económica con Texas?, *Región y Sociedad*, 20(41), 57-88.

[32] Deichmann, U., Marianne, Fay, Jun, Koo, Somik V. Lall. 2004. "Economic structure, productivity, and infrastructure quality in Southern Mexico", *The Annals of Regional Science*, 38:361?385.

[33] Donaldson, D. 2016. Railroads of the Raj: Estimating the Impact of Transportation Infrastructure, *American Economic Review*, forthcoming.

[34] Donaldson, D. and R. Hornbeck. 2016. Railroads and American Economic Growth: A Market Access Approach, *Quarterly Journal of Economics*, 131(2), 799-858.

[35] Duran-Fernandez, R., and G. Santos. 2014a. Road Infrastructure Spillovers on the Manufacturing Sector in Mexico, *Research in Transportation Economics*, 46, 17-29.

[36] Duran-Fernandez, R., and G. Santos, G. 2014b. Regional Convergence, Road Infrastructure, and Industrial Diversity in Mexico, *Research in Transportation Economics*, 46, 103-110.

[37] Duranton, G. and H. Overman. 2008. Exploring the Detailed Location Patterns of U.K. Manufacturing Industries Using Micro-Geographic Data,*Journal of Regional Science*, 48(1), 213-243.

[38] Duranton, G. and H. Overman. 2005. Testing for Localization Using Micro-Geographic Data, *Review of Economic Studies*, 72(4), 1077-1106.

[39] Duranton, G. and M. Turner. 2012. Urban Growth and Transportation, *Review of Economic Studies*, 79(4), 1407-1440.

[40] Duranton, G., Morrow, P. and M. Turner. 2014. Roads and Trade: Evidence from U.S. Cities, *Review of Economic Studies*, 81(2), 681-724.

[41] Ellison, G. and E. Glaeser. 1997. Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach, *Journal of Political Economy*, 105(5), 889-927.

[42] Ellison, G., Glaeser, E. and W. Kerr. 2010. What Causes Industry Agglomeration? Evidence from coagglomeration patterns, *American Economic Review*, 100(3), 1195-1213.

[43] Emran, M.S. and Z. Hou. 2013. Access to Markets and Rural Poverty: Evidence from Household Consumption in China, *Review of Economics and Statistics* 95(2), 682-697.

[44] Faber, B. 2014 Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System, *Review of Economic Studies*, 81(3), 1046-1070.

[45] Faggio, G., Silva, O. and W. Strange. 2014. Heterogeneous Agglomeration. SERC *Discussion Papers*, SERC DP0152. Spatial Economics Research Center, London School of Economics and Political Science, London, U.K.

[46] Fay, M., and M. Morrison. 2007. "Infrastructure in Latin America and the Caribbean. Recent Developments and key Challenges. Directions in Development. Infrastructure." The World Bank. Washington, D.C.

[47] Foote, W. 1997. Mexico's Troubled Toll Roads. *ICWA Letters*, February 3.

[48] Frye, D. 2015. Transportation Networks and the Geographic Concentration of Industries. *Unpublished manuscript.*

[49] Galasso, A. A., and A. Oettl. 2017. Roads and Innovation. *Review of Economics and Statistics* 99(3), 417-434.

[50] Ghani, E., Goswami, A., and W. Kerr. 2015. Highway to Success: The impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing", *The Economic Journal*, 126(591), 317-357.

[51] Gibbons, S., Lyytikainen, T., Overman, H. and R. Sanchis-Guarner. 2016. New Road Infrastructure: The Effects on Firms. CEPR *Discussion Paper* DP11239.

[52] Guasch, J.L., M. Fay. 2003. "Economic Activity, Agglomerations and Logistics in the Mexican Southern States". In *Mexico: Southern States Development Strategy*. The World Bank. Washington, D.C.

[53] Jedwab, R. and A. Storeygard. 2016. The Heterogeneous Effects of Transportation Investments: Evidence from sub-Saharan Africa 1960-2010. *Unpublished Manuscript.*

[54] Kim, Y. and B. Zangerling. 2016. Mexico Urbanization Review. World Bank Publications.

[55] Monge, M. 2012 Analisis de la cadena productiva de Tequila: El caso de Jalisco. *License dissertation. Departamento de Economía. UAM-Azcapotzalco.* 69 pages.

[56] Mun, S. 1997. Transport Network and System of Cities, *Journal of Urban Economics*, 42(3), 205-221.

[57] Rothenberg, A. 2013. Transport Infrastructure and Firm Location Choice in Equilibrium: Evidence from Indonesia's highways. *Unpublished manuscript.*

[58] Pérez, S. ands M. Palacio. 2009. Desarrollo regional y concentración industrial: Impacto en el Empleo (1994-2004). In *Observatorio de la Economía Latinoamericana*, Número 117.

[59] Pérez Cervantes, F. and A. Sandoval Hernández. 2015. Estimating the Short-Run Effect on Market-Access of the Construction of Better Transportation Infrastructure in Mexico. Banco de México *Working Papers*, N.2015-15.

[60] Perrotti, Daniel and Ricardo J. Sànchez. 2011. La brecha de infraestructura en América Latina y el Caribe. *Recursos naturales e Infraestructura* Series No. 153, United Nations publication, Santiago, Chile, July.

[61] Porter, M. 1998. Clusters and the New Economics of Competition, *Harvard Business Review*, 76(6), 77-90.

[62] Redding, S. and M. Turner. 2015. Transportation Costs and the Spatial Organization of Economic Activity. *Handbook or Regional and Urban Economics* 5B, Chapter 20, Elsevier.

[63] Rosenthal, S. and W. Strange. 2003. Geography, industrial Organization, and Agglomeration, *Review of Economics and Statistics* 85(2), 377–393.

[64] Staiger, D. and J. Stock. Instrumental Variables Regression with Weak Instruments. *Econometrica*, 65(3):557-586, 1997.

[65] Roberts, M., Deichmann, U., Fingleton, B. and and T. Shi. 2012. Evaluating China's road to prosperity: A new economic geography approach, *Regional Science and Urban Economics*, 42(4), 580-594.

[66] Rothenberg, A. 2013. Transport Infrastructure and Firm Location Choice in Equilibrium: Evidence from Indonesia's Highways. *Unpublished manuscript.*

[67] Schaffer, M. 2010. xtivreg2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models. http://ideas.repec.org/c/boc/bocode/s456501.html

[68] Storeygard, A. 2016. Farther Down on the Road: Transport Costs, Trade and Urban Growth In Sub-Saharan Africa, *Review of Economic Studies*, 83(3), 1263-1295.

[69] Unger, K. and R. Chico. 2004. La Industria Automotriz en tres Regiones de México. Un análisis de clùsteres, *El Trimestre Económico*, 71(284), 909-941.

[70] Unger, K. 2003. Los Clústeres Industriales en Mèxico: Especializaciones Regionales y la Política industrial, CIDE, División de Economía. *Documento de trabajo* número 278.

[71] Volpe Martincus, C. and J. Blyde. 2013. Shaky Roads and Trembling Exports: Assessing the Trade Effects of Domestic Infrastructure Using a Natural Experiment, *Journal of International Economics*, 90(1), 148-161.

[72] Volpe Martincus, C., Carballo, J. and A. Cusolito. 2014. Routes, Exports, and Employment in Developing Countries: Following the Trace of the Inca Roads. *Unpublished manuscript.*

Appendix A: Descriptives Statistics

Summary Statistics of roads, travel time and travel cost data

Table 12: Mean and median minimum travel cost & time, and efficient road length (1986 – 2014)

Statistics	Time to border	Cost to border	Time to port	Cost to port	Road efficiency
			1986		
Mean	1,150.14	2.28	663.12	1.94	244,535.80
Median	1,143.72	2.25	551.22	1.86	247,934.80
SD	449.24	0.34	398.62	0.30	83,307.17
			1994		
Mean	1,122.96	2.26	649.33	1.94	252,067.40
Median	1,110.02	2.23	542.41	1.86	255,266.10
SD	443.76	0.33	387.05	0.29	86,187.57
			1999		
Mean	1,111.30	2.25	637.15	1.93	265,420.80
Median	1,107.47	2.23	531.42	1.85	276,233.40
SD	427.93	0.32	374.74	0.28	89,278.03
			2004		
Mean	1,092.18	2.23	625.16	1.92	283,430.80
Median	1,080.79	2.21	518.52	1.84	289,366.60
SD	423.49	0.31	364.70	0.27	99,398.15
			2009		
Mean	1,007.39	2.17	583.87	1.89	297,425.70
Median	997.12	2.15	480.98	1.82	303,141.90
SD	388.01	0.28	363.45	0.26	99,650.99
			2014		
Mean	1,059.40	2.21	622.91	1.92	291,227.70
Median	1,053.18	2.19	519.05	1.84	296,489.10
SD	407.02	0.30	369.76	0.27	97,588.48

Note: Mean and median minimum travel time (in minutes) and travel cost (iceberg costs) to US border entry point and ports are calculated over the sample of 2,094 localities in Mexico. Efficient kilometers of roads reported are within a 200 km radius.

Table 13: Road network improvements in Mexico (1985 – 2016, in km)

Road type	1985	1993	1999	2004	2008	2016
Multilane divided	0	2,299	4,885	6,239	16,233	16,036
Two lanes or Divided	1,888	3,221	4,013	4,706	6,315	6,591
Pavement	45,933	43,426	42,410	42,843	32,688	31,629
Gravel or Earth road	11,506	12,054	10,865	10,458	9,964	9,247
Total	59,328	60,999	62,173	64,246	65,200	63,504

Note: Authors' calculations based on AAA map information for the corresponding years.

Appendix B: Concentration and specialization Indices

Krugman index

The Krugman Specialization index (KSI) is a widely-used specialization measure. It measures the standard error of industry shares, by computing the share of employment which would have to be relocated to achieve an industry structure equivalent to the average structure of the reference group. The Index can take values in between zero (identical territorial structures) and two (totally different structures). In our case, M = 2,094 localities, i.e., the number of metropolitan area and standalone municipalities in Mexico (consistent across years) and we will measure the locality specialization in terms of output and employment.

$$KSI_m = \sum_{j=1} |S_{m,i} - \overline{S}| \tag{4}$$

Where $S_{m,i}$ is the output or employment share of industry i in locality m, and is the average share of industry i in the total output or employment across all localities in Mexico. Interpretation: If the relative specialization measure is zero, then the economic structure of a locality is identical to the economic structure of the overall economy. The higher the index, the more the economic structure of the locality deviates from the overall economy (reference group) and the more that locality is specialized.

Ellison and Glaeser geographic concentration index

The Ellison-Glaeser index (Ellison and Glaeser, 1997) defines concentration as agglomeration above and beyond what we would observe if plants simply chose locations randomly (as opposed to a uniform spatial distribution). This measure provides an unbiased estimate of agglomerative forces independently of their source. It can be interpreted as the probability that a firm choosing its location follows the prior firm rather than locating randomly. The Ellison-Glaeser index is given by the following formula:

$$\gamma_i \equiv \frac{G_i - (1 - \sum_r x_r^2) H_i}{(1 - \sum_r x_r^2)(1 - H_i)}, \tag{5}$$

where:

- $G_i \equiv \sum_r (s_{ri} - x_r)^2$ is the spatial Gini coefficient of industry i;
- x_r is the share of total employment in each locality r;
- s_{ri} is the share of employment of locality r in industry i;
- $H_i \equiv \sum_j z_{ji}^2$ is the Herfindahl index of the plant size distribution of industry i;
- z_{ji} represent the employment share of a particular firm j in industry i.

Interpretation: Following Ellison and Glaeser (1997), an industry is strongly concentrated if $\gamma_i > 0.05$, weakly concentrated if $\gamma_i \in (0, 0.05]$, and not concentrated if $\gamma_i < 0$.

Appendix C: Additional Results

Table 14: Cross-section estimates of the effect of infrastructure on employment.

	1986		1994		1999		2004		2009		2014	
Variables	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Efficient road length	-0.516***	-0.658***	-0.441***	-0.574***	-0.489***	-0.752***	-0.472***	-0.665***	-0.496***	-0.642***	-0.505***	-0.718***
	(0.101)	(0.109)	(0.0998)	(0.109)	(0.0938)	(0.110)	(0.0882)	(0.0992)	(0.0961)	(0.108)	(0.0945)	(0.107)
Min travel cost to border	-1.138***	-1.208***	-1.177***	-1.254***	-1.387***	-1.464***	-1.319***	-1.397***	-0.673**	-0.686***	-0.663***	-0.695***
	(0.249)	(0.248)	(0.242)	(0.240)	(0.243)	(0.242)	(0.241)	(0.239)	(0.267)	(0.265)	(0.252)	(0.250)
Min travel cost to Port	-0.0702	-0.360	0.448	0.189	0.196	-0.273	0.253	-0.122	-0.144	-0.479	-0.106	-0.521*
	(0.300)	(0.305)	(0.282)	(0.287)	(0.297)	(0.299)	(0.294)	(0.296)	(0.303)	(0.305)	(0.299)	(0.298)
Above median education	1.784***	1.761***	1.882***	1.857***	1.845***	1.803***	1.796***	1.771***	1.815***	1.794***	1.786***	1.750***
	(0.0748)	(0.0750)	(0.0745)	(0.0747)	(0.0712)	(0.0718)	(0.0711)	(0.0714)	(0.0722)	(0.0727)	(0.0736)	(0.0742)
Above median population	1.133***	1.136***	1.053***	1.060***	1.039***	1.047***	1.157***	1.156***	1.109***	1.114***	1.020***	1.033***
	(0.0709)	(0.0710)	(0.0701)	(0.0701)	(0.0677)	(0.0680)	(0.0677)	(0.0679)	(0.0691)	(0.0696)	(0.0694)	(0.0697)
Observations	1853	1850	1868	1865	1870	1867	1870	1867	1870	1867	1870	1867
Adj. R-squared	0.520	0.517	0.537	0.535	0.551	0.547	0.565	0.562	0.554	0.552	0.540	0.536

Note: * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level.
The doughnut IV is calculated by excluding all localities within a circle of 200 km radius.

Table 15: Spatial heterogeneous effects of market access

Variables	Total employment				
	Border	Capital	Center	North	South
Market Access (MA)	0.154***	0.328	0.188***	0.336***	0.221***
	(0.0348)	(0.274)	(0.0248)	(0.106)	(0.0216)
MA x Metropolitan dummy	0.0756	-0.0251	0.0613**	-0.118	0.0949***
	(0.0473)	(0.123)	(0.0242)	(0.0999)	(0.0247)
MA x Education	-0.0499***	-0.150	-0.0428***	-0.0173	-0.0495***
	(0.0130)	(0.212)	(0.00744)	(0.0151)	(0.00508)
MA x Capital city	-0.0799***	-0.186***	-0.137***	-0.107***	-0.123***
	(0.00637)	(0.0232)	(0.00423)	(0.00587)	(0.00413)
MA x Oil dummy	-0.0124		0.111	0.0246	-0.0785*
	(0.0319)		(0.0896)	(0.0234)	(0.0453)
Above median education					0.671*
					(0.406)
Above median population	-0.141	0.249	0.0165	0.250	0.225**
	(0.120)	(0.774)	(0.0775)	(0.181)	(0.106)
Constant	0.862**	0.608	0.746***	-0.471	1.006***
	(0.341)	(0.449)	(0.125)	(0.417)	(0.128)
Observations	1117	206	3749	896	3798
Adj. R-squared	0.311	0.381	0.420	0.395	0.422

Note: * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level. The reported results are the doughnut IV which is calculated by excluding all localities within a 25km circle. Results are robust to others IV strategies.

Table 16: Spatial heterogeneous effects of market potential

Variables	Total employment				
	Border	Capital	Center	North	South
Market Potential (MP)	-0.0257	0.690***	0.381***	0.0595	0.424***
	(0.0951)	(0.0757)	(0.0791)	(0.118)	(0.0893)
MP x Metropolitan dummy	0.0284	0.0338	0.0454***	0.0737**	-0.0223
	(0.0294)	(0.0400)	(0.0144)	(0.0339)	(0.0154)
MP x Literacy	0.0218**		0.0141**	0.0284**	0.0189***
	(0.00920)		(0.00584)	(0.0133)	(0.00398)
MP x Capital city	0.0339	-0.166***	-0.0846***	-0.00115	-0.0791***
	(0.0263)	(0.0219)	(0.0206)	(0.0279)	(0.0273)
MP x Oil dummy	0.00265		-0.00494	-0.0132**	-0.00822
	(0.00984)		(0.00426)	(0.00523)	(0.00866)
Above median education					-0.00480
					(0.0207)
Above median population	0.146	0.142**	0.0980*	-0.0134	0.157**
	(0.104)	(0.0674)	(0.0560)	(0.144)	(0.0689)
Constant	-0.0609	-0.722**	-0.373**	-0.748*	0.470**
	(0.352)	(0.368)	(0.174)	(0.396)	(0.195)
Observations	933	172	3119	744	3191
Adj. R-squared	0.214	0.750	0.634	0.489	0.534

Note: * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level. The reported results are the doughnut IV which is calculated by excluding all localities within a 25km circle. Results are robust to others IV strategies.