



**UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO
FACULTAD DE ESTUDIOS SUPERIORES
ACATLÁN**

**EL TRANSPORTE Y SU INFLUENCIA EN LA PRODUCTIVIDAD Y EL
COMERCIO INTERREGIONAL, 2009-2014.**

T E S I S

**QUE PARA OBTENER EL TÍTULO DE
LICENCIADO EN ECONOMÍA**

P R E S E N T A

MARIA VICTORIA MUÑOZ TAPIA

ASESOR: DR. ROLDÁN ANDRÉS ROSALES

SANTA CRUZ ACATLÁN, NAUCALPAN, ESTADO DE MÉXICO, NOVIEMBRE 2016



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

No exagero si digo, que desde que tengo memoria he esperado este momento en el que me encuentro. Si ahora me preguntaran, ¿Con qué soñabas cuándo tenías 10 años? Yo contestaría inmediatamente, con ir a la Universidad y no cualquiera, la UNAM obviamente. Y si, para esa edad es algo en lo que nadie piensa aún, te llaman la atención otros miles de cosas. Sin embargo yo sabía que eso era lo que quería, no tenía ni idea de lo que tendría que hacer, pero de lo que si estaba segura era que no quería nada más que eso.

La motivación ahí estaba conmigo a lo largo de todo el tiempo. Claro que, alguien tenía que aumentar esa motivación y convertirla en una bola de nieve que cada vez se hiciera más grande y fuerte. Mis padres, ellos son los que se llevan todo el crédito de mi constancia y perseverancia por hacer crecer mi motivo cada día. Ellos estaban ahí 24x24 para lo que se me ofreciera, la realidad es esa, que todo lo invertían en mí, en mí futuro. Y cómo es que yo iba a tener el descaro de echar a la basura su confianza, jamás. El compromiso y la responsabilidad, por lo tanto, fueron las otras dos semillas que crecieron dentro de mí gracias a ellos. Nunca tendré la vida necesaria ni las palabras correctas como para agradecerles que la persona que ven ahora, en lo que me he convertido es por ellos y no por nadie más.

Debo confesar que este camino no ha sido nada fácil, que la meta siempre estaba en mi mente pero que a veces se nublaba con las dificultades que se me presentaban. Me llevó momentos complejos en los que dudaba de lo que podía llegar a lograr o me daba miedo pensar, qué iba a ser de mí si yo no tenía lo que se necesita y más para superar esas dificultades. Afortunadamente debo agradecerle a Dios por haber puesto en mi camino personas que me ayudaron a salir a delante con su apoyo y sus consejos. Maestros y amigos que vieron en mí algo, no podría afirmar que fue, pero la confianza que depositaron en mí ha sido enorme.

Espero no olvidar a nadie; Gracias tutor Dr. Roldán porque sin su asesoría en este paso final no me encontraría escribiendo esto. Gracias a mis hermanas por soportar tanto tiempo sin que les pusiera más atención por estar estudiando. Gracias a mi abuelita por entenderme que si no la iba a visitar no era porque no quisiera sino porque no tenía tiempo para hacer todas las tareas que tenía. Gracias a mi familia que resta, por no molestarme en las fiestas familiares mientras estaba con mi amiga la computadora. Gracias a Rosario, Alex y Jaime porque de ustedes aprendí todo lo que necesité al final de este proyecto, me terminaron de formar como profesionista, los considero sin duda, mi familia. Y demás, que me mostraron que la vida no es fácil si la estamos analizando en todo momento, por centrarme en la realidad, gracias por creer en mí y por mostrarme que también yo debo creer en la persona que soy, por haberme dado ese equilibrio que necesitaba en todo el camino.

Sin duda en mi memoria se queda la mejor etapa de mi vida, los cinco años que pasé en mi *alma mater*. Descubrí tanto de mí, aprendí temas infinitos, conocí la mejor parte de mi historia. Agradezco la calidad de personas que me acompañó durante este tiempo fuera y dentro de la universidad, gente que más allá de compartirme su conocimiento me transmitió su experiencia. Invaluable es todo lo que me llevo y espero haber dejado algo bueno por lo menos, en cada uno de ellos. Aquí estoy, logrando por fin lo que más soñé, lo que algún día hicieron todos ustedes por mí déjenme decirles que valió ABSOLUTAMENTE la pena. Emoción es lo que siento y con motivaciones infinitas seguiré adelante con lo que viene. Las bases ya están dentro de mí y las personas correctas a mi lado también, para conseguir muchas más metas de estas.

¡Gracias Universidad, Gracias Acatlán, por tanto y tanto!

VICTORIA MUÑIZ

¡Por mi raza hablará el espíritu!

C O N T E N I D O

| | |
|--|----|
| INTRODUCCIÓN | 5 |
| CAPITULO 1 DISCUSIÓN TEÓRICA REGIONAL Y DEL TRANSPORTE | 7 |
| 1.1 La Región y el Espacio en la Actualidad | 7 |
| 1.1.1 Crecimiento Regional | 11 |
| 1.1.2 Competitividad y Productividad Regional | 16 |
| 1.2 Costos Del Transporte | 23 |
| CAPITULO 2 ANÁLISIS DESCRIPTIVO – ESPACIAL | 29 |
| 2.1 Discusión Teórica Del Análisis Descriptivo | 29 |
| 2.2 Análisis Exploratorio de los Datos | 31 |
| 2.2.1 Especificación De Las Variables | 32 |
| 2.2.2 <i>Exploratory Data Analysis (EDA) – Exploratory Spatial Data Analysis (ESDA)</i> | 36 |
| 2.2.2.1 Análisis Exploratorio de los Datos | 3 |
| | 6 |
| 2.2.2.2 Análisis Exploratorio Espacial de los Datos | 38 |
| 2.2.3 Dependencia Espacial | 44 |
| CAPITULO 3 EVIDENCIA EMPÍRICA SOBRE EL EFECTO DEL TRANSPORTE Y EL COMERCIO EN LAS REGIONES | 54 |
| 3.1 Importancia Del Análisis Econométrico | 54 |
| 3.1.1 Modelo Lineal Clásico | 57 |
| 3.1.2 Lag y Error | 59 |
| 3.2 Datos De Panel | 65 |
| 3.2.1 Modelo de Efectos Fijos: Within | 68 |
| 3.3 La Matriz De Distancias y El Modelo Durbin | 70 |
| 3.4 Evidencia Empírica Del Modelo | 72 |
| 3.5 Errores del Modelo | 76 |
| CONCLUSIONES | 78 |
| BIBLIOGRAFÍA | 81 |

INTRODUCCIÓN

La infraestructura es una de las principales herramientas y al mismo tiempo la principal restricción para el crecimiento económico regional de un país. Funge como vía de cohesión territorial y económica dado que une al territorio para hacerlo accesible desde cualquier punto, permitiendo al comercio conectarse con el mercado interno y externo. Las carencias de la infraestructura en el transporte traen consigo costos que abren una brecha entre el precio recibido por el productor y el precio pagado por el consumidor final.

Argumenta Pérez (2006) que existe un gran incentivo en la utilización de nuevas y mejores obras de infraestructura. Estas nuevas obras implican una reducción de costos en las actividades de integración de los mercados de insumos (tales como materias primas y mano de obra) localizados en poblaciones distintas (Gannon y Liu, 1997). Se incentiva el comercio regional (e internacional) cuando los costos de transporte disminuyen para los agentes económicos que intervienen en las actividades de integración.

Al estudiar las variables infraestructura y comercio es necesario situar un marco de análisis económico espacial, poniendo especial énfasis en las implicaciones regionales. La movilidad de bienes (productos o mercancías) a través del transporte es por definición un hecho geográfico dada su indiscutible expresión espacial; de aquí que su dimensión geográfica resulte fundamental en los métodos de investigación y en la formulación del modelo econométrico.

Autores regionales como; Richardson (1986), Polése (1998), Fujita (2000), Asuad (2001) entre otros, destacan la importancia del área geográfica en la teoría regional. Todos ellos abordan el tema del transporte y sus factores como variable de estudio que explica la importancia de la localización y el espacio.

El replanteamiento de la ideología con la adición del sistema de comercio a nivel regional se presenta a raíz de la carencia de información analítica que permita explicar y construir modelos que tengan en cuenta a la estructura espacial.

El enfoque de esta investigación utiliza el recurso infraestructura del transporte de la economía de México para conocer la influencia que tiene en la productividad manufacturera¹ y a su vez en el comercio regional de las 32 entidades federativas. A través de los recientes desarrollos tanto estadísticos como de software se pretende conocer la dimensión espacial de los procesos infraestructura del transporte – comercio y productividad regional.

El trabajo está estructurado en tres capítulos, en el primero de ellos se aborda la discusión teórica regional y del transporte relacionado con la productividad y competitividad². En el segundo capítulo, se hace un análisis descriptivo-espacial de las variables recolectadas en los periodos observados. En el último capítulo se exponen los resultados obtenidos de la evidencia empírica sobre el efecto del transporte y el comercio en las regiones. Finalmente, se presentan las conclusiones.

¹La productividad es una variable que muestra que tan eficiente es una región para utilizar sus recursos. Menciona Iregui, et al., (2006) que la productividad de la industria manufacturera ha sido objeto constante de estudio. El parámetro de productividad manufacturera en el análisis comprende datos del sector manufacturero a nivel entidad federativa.

² La competitividad en términos generales es la capacidad de competir. En esta investigación se desarrolla a profundidad el concepto con base en el contexto y propósito del análisis.

La competitividad se relaciona directamente con la productividad manufacturera. La dotación en infraestructura del transporte para la conexión entre regiones favorece la productividad manufacturera específicamente en el comercio de mercancías lo que hace a una región más competitiva con respecto a las demás.

CAPITULO 1 DISCUSIÓN TEÓRICA REGIONAL Y DEL TRANSPORTE

1.1 La Región y el Espacio en la Actualidad

Hoy en día, en el ámbito nacional, es necesario e importante analizar el comportamiento de las economías regionales e impulsar su crecimiento y desarrollo a fin de combatir problemas específicos. El estudio de la región como manera de introducirse a la economía regional. Estamos sumergidos en una economía altamente desigual con regiones o estados que poseen características distintas que están muy lejos de ser homogéneas, es por ello, la importancia de orientar los estudios a los marcos regionales que se adapten al contexto de cada entidad.

La interpretación del contexto regional consiste en una concepción integral entre economía y espacio, al incorporar al análisis económico su dimensión espacial, la atención se enfocará a la estructuración regional y territorial de la economía. Entonces, no es aventurado señalar que el comportamiento de la economía depende sustancialmente de sus componentes espaciales y territoriales (Asuad, 2001). Fujita *et al.*, (2000) menciona que todas las actividades económicas se localizan en el espacio. Sin embargo, señala que el espacio no fue parte importante sino hasta el siglo XX.

El interés histórico de la economía regional lo explica Cuadrado (2006), quien considera que esta ciencia empieza con los temas de la localización que se reaviva en la primera mitad del siglo XX, al tiempo que se producen otras aproximaciones más ambiciosas. Entre los autores más destacados hay que citar necesariamente los trabajos de Hotelling, Engländer, algo más tarde a Lösch y Hoover. Cuando realmente se producen cambios importantes que marcan el inicio del despliegue de la Economía Regional es a partir de los cincuenta.

Por ejemplo, a finales de esa época, Arrow y Debreu (1954) realizaron ya algunas incursiones de interés en lo que hoy se califica como Economía Regional, al proponer la doble integración del espacio y del tiempo en la teoría económica. A lo largo de las décadas de los sesenta y los setenta, economistas de cierto renombre interesados por los aspectos „espaciales“ de los procesos económicos y su incorporación a la teoría económica, acabó ampliándose muy notablemente.

Economistas como Isard, quien desde principios de los cincuenta venía realizando relevantes aportaciones al análisis regional, propugnando la necesidad y justificación de una „Ciencia Regional“, Beckmann y Greenhut con sus contribuciones a la teoría de la localización o con los estudios sobre crecimiento y disparidades regionales de Friedmann, Hansen, Perroux, Aydalot, Richardson, Von Böventer y Anselin además de algunos autores latinoamericanos que hicieron excelentes aportaciones a la teoría del desarrollo y el sub-desarrollo económico, como Furtado y Prebisch.

Todos ellos encontraron en el campo regional un excelente „laboratorio“ de análisis porque veían en los datos unas estadísticas más equilibradas y comparables según lo que ofrecía cada región. Ahora se puede afirmar que la economía regional ha tenido una rápida evolución tanto teórica, como empírica, particularmente a partir de los noventa (Quintana y Andrés, 2014). Ésta se había desarrollado en la periferia de la teoría económica dominante a partir de sus dos enfoques principales; el microeconómico constituido por la teoría de la localización industrial y el macroeconómico que se integraba por la teoría del crecimiento y desarrollo (Capello, 2011).

El espacio en general es una realidad abstracta, más global, sin detenerse en las particularidades. Donde lo que más interesa son las relaciones económicas entre los seres humanos. Las características del espacio influyen en el comportamiento del ser humano y las acciones de éste transforman el espacio.

Económicamente se analizan tres tipos de tratamiento del espacio; el espacio como distancia, como superficie y como lugar (Polése, 1998). Para Richardson (1986), en el espacio se encuentra una división de la clasificación de regiones, pero es más importante lo que ocurre dentro de las regiones.

Asuad (2001) habla del concepto de espacio, considera que podría ser un concepto sencillo, pero solamente es a primera vista, ya que aparte de ser un término utilizado en otras disciplinas como la filosofía éste ha evolucionado en el tiempo. Losch (1944), define al espacio económico como un campo de fuerzas económicas entre unidades económicas, dentro de las cuales actúan polos económicos como centros de atracción y repulsión.

Para Perroux (1949), el espacio económico rebasa las fronteras nacionales, generalmente no coinciden con el espacio territorial ni con el país en su conjunto, como las relaciones económicas que establecen las empresas, independientemente del espacio geoeconómico y del político. Mientras que para Boudeville (1970), el espacio económico es tridimensional, se caracteriza por la aplicación de las variables económicas “sobre” o “en” un espacio geográfico determinado a través de las transformaciones matemáticas, las cuales describen un proceso económico y pueden ser consideradas por su homogeneidad, polarización o finalidad, similar a la visión de Perroux (1949).

En años más recientes, Asuad (2013) explica al espacio como social, el cual se integra por un conjunto de sub espacios; físico y biológico. Constituidos cada uno por el tipo de actividades humanas; económicas, socioculturales y políticas, éstas interactúan con el espacio natural, integrado sinérgicamente por los sub espacios. El desarrollo económico y social desigual propicia barreras y restricciones económicas, sociales, políticas y espaciales, que tienen que ser removidas para impulsar el desarrollo económico y social del espacio y del territorio que lo integra en su conjunto y así lograr el desarrollo con cohesión social. Por lo tanto, al espacio en la toma de decisiones no se puede separar del comportamiento económico.

La región se distingue del espacio por su homogeneidad interna y por la continuidad que presenta, mientras que el espacio es heterogéneo y no continuo. Delimitar regiones dependerá de los objetivos de la investigación, definir regiones con precisión es entrar en dilemas económicos administrativos e históricos y hasta políticos. Para Richardson (1986), la región es presentada como bloques de países contiguos dentro del sistema económico internacional, igual que para Polése (1998), las regiones se integran unas a otras y agrega que las fronteras nacionales son cada vez menos eficaces; es decir, que en un sentido económico no deberían de existir y así las personas, las mercancías, el capital y las ideas pueden circular libremente en la región, entonces los desplazamientos de bienes y servicios de una región a otra no deberían tener trabas.

Asuad (2001) piensa que no existe regla absoluta para definir los límites de una región y, que la región se define como un área contigua y homogénea, única en localización y características, por lo que se concibe como una unidad espacial concreta y definida. La relatividad de la región se refiere a su carácter cambiante y dinámica, como producto de las fuerzas que interactúan y también como procesos en el tiempo, es la división del espacio nacional en número limitado de regiones adyacentes y que está en función de una dimensión temporal como espacial (Richardson, 1986).

Asuad (2013) cree en la dinámica de la economía regional para mostrar los cambios por los procesos de las actividades económicas y menciona que la configuración económica de las regiones y de su territorio constituido por sistemas de ciudades y redes de transporte que se vinculan entre sí, que son consecuencia directa de la concentración y es una de las características fundamentales de la economía espacial. Por lo tanto, en este análisis económico regional se realiza la apertura del territorio, la interdependencia y la movilidad de los factores de producción.

No se niega como enfatiza Asuad (2001) que existen restricciones considerables debido a que la teoría económica regional ha sido relegada por mucho tiempo y se carece de investigaciones y estudios aplicados. A esto contribuye la falta de información y de conocimientos aplicados que propongan explicaciones regionales. Por esta razón es trascendental la revisión, el desarrollo técnico y metodológico de esta teoría para mostrar pruebas empíricas de la evolución regional.

1.1.1 Crecimiento Regional

El estudio del crecimiento económico regional es prácticamente nuevo en México. La pregunta que se hace dentro de muchas cuestiones que existen es, ¿Por qué el crecimiento de la economía mexicana, a pesar de sus existentes ventajas territoriales con los demás países se ha visto estancado por años? Actualmente la nueva literatura económica reconoce que el territorio y sus alcances no pueden quedar al margen de cualquier análisis porque no solo dificulta o facilita los movimientos de los factores de producción y bienes producidos, sino que todos los agentes económicos toman decisiones con base en el espacio-región.

Las teorías del desarrollo económico corresponden a interpretaciones más amplias e integrales sobre el proceso de desarrollo en el tiempo y espacio. En sus propuestas de explicación, generalmente se consideran varios factores determinantes de dicho proceso, por lo que se asume que es resultado de un proceso multifactorial complejo en el que se entrelazan diversos aspectos económicos, sociales, políticos y espaciales (Asuad, 2013). Por ejemplo, Duch (1998) expone la necesidad de integrar la dimensión espacial, ya que la teoría económica convencional crea modificaciones rigurosas y tradicionales al análisis competitivo.

Recordando que la teoría neoclásica (*mainstream* de la economía) del sector externo, fundamentada en las ventajas comparativas de los países, con el principio de la existencia de competencia perfecta y la ausencia de economías de

escala en la producción, explica que los países con una economía abierta al comercio internacional podrían maximizar el bienestar y el ingreso nacional gracias a una óptima asignación de los recursos disponibles, recursos en los que no están incluidos el espacio o la localización.

La justificación para adoptar medidas de liberalización en el comercio radicó en la relación positiva que podría existir entre libre comercio y crecimiento económico, que es tema importante para los países en desarrollo. En general, lo que el modelo tradicional (neoclásico) buscaba era convertir a las exportaciones en el motor del crecimiento y dinamizar el aparato productivo nacional, exponiéndolo a la competencia internacional. La existencia de una relación entre comercio y crecimiento no podía basarse en el modelo tradicional de comercio internacional. Ante esto nace con el desarrollo (a comienzos de los noventa), la necesidad de un cambio de enfoque considerando las nuevas teorías del crecimiento económico que incluyen al comercio y espacio.

Es claro que la consideración de la dimensión espacial es relativamente insuficiente e ignorada hasta este punto. Hasta que después con la Nueva Geografía Económica (NGE), una de las áreas más influyentes de la economía regional contemporánea que actualmente considera al espacio no sólo como una variable o elemento adicional o particular del análisis económico, sino como una dimensión fundamental del comportamiento económico, la dimensión espacial de la economía, la cual debe integrarse, a fin de desarrollar una teoría económica espacial, aspectos fundamentales que propuso inicialmente Isard, mismos que se han ido desarrollando en las teorías regionales y urbanas.

Anteriormente, el tema fue retomado por Krugman (1995), quien hace referencia a la ubicación de la actividad económica regional con un modelo de competencia monopolística para explicar las disparidades regionales del crecimiento. Es decir, La NGE ha proporcionado el replanteamiento de algunos problemas económicos en relación con la localización de las actividades productivas,

Los efectos de aglomeración, el desplazamiento de los factores de la producción y de los productos. Específicamente la teoría se ha centrado en los costos de transporte, el conocimiento tecnológico y su crecimiento, una vez ocurrida una aglomeración en la actividad económica. Los costes están directamente asociados al comercio y/o transporte.

Cuadrado (2012) reconoce que la NGE ha contribuido a vincular a la economía y a la geografía con el aspecto internacional, que puede servir para revisar varios aspectos en los que la Economía Regional está tratando de avanzar. A manera de ejemplo, Asuad (2001) argumenta que el comercio internacional ha dependido más de la economía de las regiones que de los países en su conjunto y, la evolución de los servicios que ha acompañado el desarrollo económico de los países se ha traducido en un crecimiento sin precedentes de los centros regionales.

Las tendencias de la integración económica dependen en gran medida de la concentración de la actividad económica en las regiones y de los principales centros urbanos que las conforman. De aquí que sea fundamental el análisis de las economías regionales al interior de los países, que debe precisar tanto su vinculación internacional como la expansión de sus industrias y servicios, identificando las relaciones y efectos que genera el desarrollo regional y económico del país.

En el caso específico de la expansión de las industrias en México, se conoce que su grado de participación obtenido en el mercado internacional en las últimas décadas ha sido el de irse “adaptando” a los requerimientos del mercado. Las exportaciones manufactureras han ayudado a disminuir el déficit del comercio del país: puede decirse que en el país se ha creado un patrón de especialización exportadora que influye en el grado de diversificación por destino y por producto. Sin embargo, pese a esta evolución en la productividad manufacturera, no ha resultado suficiente si el país no le apuesta al desarrollo del mercado interno.

De acuerdo con Ocegueda (2003), México ha consolidado el liderazgo de un grupo de industrias con alta capacidad exportadora que habría impactado favorablemente al crecimiento de la economía mexicana, mientras que las dificultades para articular estas industrias con los proveedores nacionales de insumos es un obstáculo debido a una red de infraestructura ineficiente. En suma, sugiere la necesidad de reconsiderar el papel de fortalecer el mercado interno (proveedores internos) a fin de que las industrias exportadoras no tengan que consumir bienes de proveedores externos.

Hay que destacar que el intercambio internacional con otros países y sus obstáculos han recibido ya suficiente atención en la literatura teórica y empírica, incluso el enfoque tradicional ya no es capaz de explicar los paradigmas del crecimiento económico como pasó en 1970, North (1970) se empezó a ver limitado en su análisis por el enfoque neoclásico para poder explicar el desempeño económico a través del tiempo y trató de construir un marco analítico más apropiado, en el que integraba supuestos con base en la dinámica de los agentes económicos.

La teoría de la Base Económica o Exportadora y la de Localización, aunque son teorías surgidas en contextos diferentes, convergen, mediante la desregulación, en un punto de interdependencia y complementariedad. Casas (2006) menciona que el éxito de una industria en producir bienes exportables sólo podía entenderse a partir de la teoría de la localización y de la teoría de la Base Económica.

Ambas teorías tanto la de la base exportadora como la de la localización discuten que las regiones alcanzan una etapa de especialización productiva, a fin de dotar al sector externo y, en condiciones de infraestructura del transporte más favorecedores, les permiten enviar bienes a otras a menores costos haciéndolas productivas. En la medida que las regiones crecen alrededor de los bienes principales y exportables se genera un proceso de agrupamiento y se localizan empresas, lo cual, mejora la competitividad de dichos bienes y propicia un mayor ingreso a la región (Casas, 2006).

Investigar la presencia de la “infraestructura del transporte” regional es fundamental, North (1977) manifiesta que se puede suponer que algún tipo de relación de equilibrio surgirá entre las regiones, a medida que los costos de transporte se hagan menos significativos. Las relaciones comerciales conllevan ubicaciones en distintos puntos, desde la empresa que se encarga de proveer las materias primas, a las que van a manufacturarlos, hasta los puntos donde se distribuyen y comercializan los productos finales. Como argumenta Polése (1998), cubrir distancias entre todos los puntos lleva esfuerzo, recursos y tiempo, lo que para las empresas implica costos.

Se señala entonces que los costos están en función de la distancia; el costo de producción se ve afectado por el costo de transporte y éstos son cubiertos por el consumidor. Por lo tanto, la infraestructura de transporte se convierte en la columna vertebral de las relaciones de comercio para las empresas manufactureras. Así que deberán coexistir dos modelos: el de desarrollo económico hacia fuera, mediante el sector exportador³ y, el de fortalecimiento del mercado interno que incluya a la infraestructura del transporte.

Por ahora se usará un marco referencial con autores⁴ como Weber (1909), quien delimitó el ámbito formal de los estudios de la localización industrial, Christaller (1966) quien desarrolló la Teoría del "Lugar Central", Lösch (1954) quien inició sus trabajos con la intención de producir una teoría sistemática de la localización tanto industrial como agrícola capaz de definir el equilibrio en el ámbito espacial, dada a conocer en su "Teoría de la localización y el equilibrio territorial".

³A partir de 1994, la economía mexicana ha tendido que especializarse hacia la industria maquiladora de exportación con el establecimiento del Tratado de Libre Comercio con América del Norte (TLCAN). La expectativa era que mediante este tratado se podría incrementar el crecimiento de las exportaciones mexicanas y, por tanto, un incremento al crecimiento de la economía. Sin embargo, las tasas de crecimiento anual de las exportaciones de México siguen cerca del mismo comportamiento, lo que hace cuestionar la eficiencia del sector manufacturero y de la infraestructura de transporte para distribuir bienes a Estados Unidos, Canadá e incluso hacia otros países.

⁴ Autores de la escuela neoclásica según, (Miguel, 2004)

Según Miguel (2004) ha sido la teoría de Lösch, la que mejor se adecúa a las premisas del análisis para explicar la localización espacial de las actividades. En tanto que para Weber el desarrollo del espacio geográfico sustentado en la localización industrial se orientará hacia los centros donde se concentra la fuerza de trabajo, las materias primas y/o los mercados de venta del producto, preferencialmente hacia los puntos del espacio en los cuales los costos de transporte resulten mínimos; para Lösch, la localización de la industria, y en general, de cualquier actividad, no solamente será la resultante de la minimización de los costos, sino también de la maximización de las ganancias, y ésta se logrará en las áreas geográficas donde la captación de ingresos sea máxima y/o los costos sean mínimos.

Sin dejar a un lado la teoría de Christaller (1933) ya que aborda la existencia de una red de transporte que vincula directamente áreas de productos y servicios formando lo que llama él una jerarquía espacial dentro de un territorio. Entre los supuestos de la teoría de la localización los costes de transporte son proporcionales a la distancia. Aunque Lösch presenta una versión del principio de jerarquía más formalizado y con mayor fundamentación económica. En este caso es posible sustentar a la econometría espacial a través del uso diversas teorías de crecimiento regional (TCR), por la heterogeneidad que posee el país, bajo el sustento de Vázquez (2002), la diversidad de la economía conduce a diferentes sendas de crecimiento para cada región, brindando diferentes etapas y niveles de desarrollo.

1.1.2 Competitividad y Productividad Regional

Analizar al espacio y región es esencial en la planeación del desarrollo y crecimiento regional, dándole un sentido geoeconómico al territorio, aclarando que la definición a estos dos conceptos depende de las ideas y visión de quién haga uso de ellos para explicar su análisis. Boudeville (1966), por ejemplo, concibió al espacio como la sangre y el suelo de la geopolítica, únicamente para el uso exclusivo en ese contexto que se encontraba.

Ahora los conceptos competitividad y productividad aparecen con una noción más operativa y su definición dependerá del contexto. Por ejemplo, Corona (2005) agrega al capital de inversión, las redes de transporte, la industria y las técnicas agrícolas como un nuevo escenario y antecede a la competitividad regional. Contrariamente Polése (1998) aborda la competitividad desde una visión regionalista, considera que depende tanto de factores tecnológicos y administrativo como de aspectos regionales y urbanos, la competencia es resultado de economías en el uso de factores y reducción de costos.

Al igual que la región y el espacio, la competitividad resulta también un concepto variado y diferente que dependerá de la corriente teórica a la que pertenezca cada autor y el contexto al que se enfoque. El nuestro será relacionado con la competitividad económica que se refiere a las condiciones que permiten que la producción de un bien o servicio compita en mercados. Black (1997) lo ve como la habilidad para competir en mercados por bienes o servicios basada en una combinación de precio y calidad, así las regiones pueden ser competitivas solo si los precios de sus productos son tan bajos como los de sus rivales en el mercado externo.

El término competitividad se ha ido enriqueciendo a partir de propuestas en el análisis, considerando que la competitividad de las empresas y, de las regiones se ve afectada por el contexto macroeconómico y las políticas económicas aplicadas tanto en el mercado doméstico como en el internacional, al afectar los costos de producción y distribución. En ese sentido, Asuad (2001) considera que la condición para competir internacionalmente (considerando los problemas regionales que se tienen), son las políticas de reforma estructural, que implican dentro de muchas cosas la reestructuración industrial y los procesos de regionalización como respuesta.

Los procesos de regionalización de los que habla Asuad (2001) se caracterizan por el conjunto de intervenciones orientadas a reorganizar las capacidades económicas y tecnológicas de la región. Según la definición por la Organización para la Cooperación y el Desarrollo Económicos (OCDE), competitividad es la capacidad de un país de tener ventaja (o desventaja) a la acción de vender sus productos en mercados internacionales.

El desempeño comercial es importante para la mejora de la competitividad de una nación, aparecen la empresa y el Estado (o gobierno) como elementos que contribuyen para mejorar la competitividad. Por lo tanto, el concepto se extiende a la empresa, un elemento fundamental en la contribución de la competitividad, el cual se integra de cuatro tipos según Asuad (2001): 1) infraestructura física-transporte y comunicaciones, 2) infraestructura humana – calidad y productividad de la fuerza de trabajo, 3) infraestructura institucional; económica, legal, financiera, instituciones de apoyo a las exportaciones y al desarrollo tecnológico: normas, estándares, pruebas, control de calidad, investigación y desarrollo y 4) el medio ambiente macroeconómico.

Porter en 1980 desarrolla su teoría de la competitividad nacional, con una muestra de 10 países representativos del comercio para ese año de diferentes partes del mundo, su objetivo era encontrar las causas de la productividad, pero también a formular políticas orientadas a promover la competitividad de una nación. De acuerdo a la teoría de Porter (1980), la prosperidad nacional no se hereda, sino que es creada por las oportunidades que brinda un país a sus empresas, ya que son las únicas de crear innovación.

La empresa puede crear ventajas frente a otras en diversas circunstancias, para mantener esa ventaja, necesita primero renovarse, porque si no lo hace, los competidores superaran a cualquier empresa que deje de mejorar e innovar.

Para Buendía (2013) una nación es competitiva si es capaz de utilizar eficientemente sus recursos, alcanzar altos niveles de productividad y proporcionar un buen nivel de vida a su población. Por lo tanto, las empresas tienen que lograr y mantener altos niveles de productividad mediante actos de innovación (Porter, 1980).

Krieger (1999) en discusiones sobre liberalización y competitividad, enfatiza en que las decisiones que toman los países deberían ser adoptadas con mesura. Dirige sus críticas a las propuestas de liberalización externa del Banco Mundial (BM) y Fondo Monetario Internacional (FMI) como condición suficiente para mejorar la competitividad internacional en países en desarrollo. De entre las recomendaciones de política económica que existen, aconseja que competitividad, competencia e innovación deberían ser entendidas como un complejo sistema de relaciones en el que para aumentar las oportunidades de las empresas locales y sostener su competitividad a través del tiempo no se debe perder el ritmo de la productividad en la industria (Correa, 2009).

En perspectiva de Rodrik (2006), la distancia respecto de la ortodoxia recomendada por el FMI y el BM que han conducido al éxito comercial de China. El argumento se ve reflejado actualmente en la economía mexicana que está siendo desplazada por la economía china, México pierde competitividad y China gana terreno en el mercado norteamericano gracias a sus bajos costos de producción. Según Correa (2009), considerar los problemas de competitividad internacional para México significa que las perspectivas de perder mercados externos en un modelo que considera las exportaciones como base de la dinámica económica, cuestiona no solo la posibilidad de lograr mejores resultados, sino la factibilidad del propio modelo e indicaría la necesidad de tomar decisiones para modificarlo.

En un trabajo referente a China, Rodrik (2006) sostiene que han sido las intervenciones de política económicas, entre ellas las que se mencionaron en el

párrafo anterior las que sirvieron de base a su sistema de comercio. Sin embargo, específicamente él menciona que son las políticas de atracción de Inversión Extranjera Directa (IED) y fomento a las exportaciones, las que han colocado al comercio chino en un nivel con alto valor agregado, mayor al que se esperaría de un país que llevaba tiempo presentado el nivel de ingreso como el de uno en vías de crecimiento.

De este modo, podemos preguntarnos, ¿Qué hace a las regiones competitivas y a las naciones? ¿Cómo influye el transporte internacional en la competitividad? Asuad (2001) explica que la relación de la competitividad con las regiones se establece porque éstas inciden significativamente en su desempeño, y se debe a las diferencias en la dotación de factores regionales y locales y las restricciones de su movilidad, facilitadas por la revolución de los transportes y las comunicaciones, que se reflejan en una disminución de los costos de transporte e información, facilitando la integración positiva. Es decir, no sería premeditado afirmar en este punto que la competitividad depende tanto de factores tecnológicos y administrativos como de aspectos regionales y urbanos.

De la misma forma, Porter (1999), Krugman (2000) y el Foro Económico Mundial (WEF) confirman que la prosperidad es creada por la productividad con la que una nación puede utilizar sus recursos (o factores) para producir bienes y servicios. La competitividad es el elemento principal de la ventaja competitiva nacional. A su vez, la ventaja competitiva nacional se define como el conjunto de instituciones, políticas y factores que determinan el nivel de productividad de un país (WEF, 2010). La capacidad de una nación para proporcionar un alto nivel de competitividad depende de la capacidad de sus empresas para lograr altos niveles de productividad, la forma en cómo ellas utilizan los recursos. La función de las empresas es seguir mejorando su productividad, su calidad de productos, la tecnología y/o aumentando la eficiencia de la producción (Porter, 1999).

Se necesita trabajo en conjunto entre el Estado y el sector privado, ya que la competitividad no va a surgir espontáneamente, sino que el país debe brindar oportunidades al sector privado para que este las desarrolle. El Estado lo puede hacer a partir de la inversión en educación, infraestructura, buenas instituciones, estabilidad macroeconómica, centros de investigación tecnológica, entre otros. Un trabajo en conjunto como un sistema que ayude a crear y desarrollar sectores productivos, innovadores que se vuelvan competitivos capaces de competir a nivel mundial (Buendía, 2013).

Por estas razones, si se da una apertura del mercado conllevará a un gran número de competidores que buscarán acaparar la atención del consumidor, lo que provocará que las empresas se vean forzadas a ser cada vez más competitivas. Pero, ofrecer el mejor producto no es sinónimo de competitividad, durante la comercialización atravesará por diversos procesos que podrán convertirse en posibles obstáculos o bien en factores de competitividad, que van desde el paso por aduanas, procesos de carga y descarga, servicios de transporte y otros servicios logísticos (Zamora y Pedraza, 2013).

Una red de infraestructura de transporte bien desarrollada es un prerrequisito para el acceso a las actividades económicas y servicios a nivel mundial. Modos efectivos de transporte, incluyendo calidad de los caminos, vías férreas, puertos y transporte aéreo permiten a los emprendedores hacer llegar sus bienes y servicios a los mercados en forma segura y a tiempo facilitando el movimiento de los trabajadores hacia mejores empleos (Foro Económico Mundial, 2011). En un trabajo de investigación que realiza Zamora y Pedraza (2013) analizan al transporte internacional como factor de competitividad en el comercio y plantean como hipótesis que la competitividad en materia de transporte internacional está determinada por las distancias y, sus costos, la calidad de los servicios de transporte, plazos de ejecución, las carreteras, puertos, líneas férreas y aeropuertos, el tráfico (toneladas), el tráfico (contenedores) y el tráfico

(mercancías en general) y la importancia del transporte internacional en la balanza comercial.

Como argumentan Smith *et al.*, (2008), la mejora de los sistemas de transporte es un elemento crítico en el éxito de los negocios, comunidades y la gente; lo que conlleva a la medición de la competitividad. En cuanto al problema de la falta de competitividad del transporte, destacan: los elevados costos, la infraestructura del transporte en carretera, vías férreas y puertos. La nueva competencia según Best (1990) se caracteriza por los siguientes cuatro aspectos: 1) Unidad empresarial que se distingue por la búsqueda de continuas mejoras en los métodos en los métodos, productos y procesos, en su organización desde la cima a la base; 2) la importancia de la cadena de producción; es decir, la competitividad depende de sus proveedores; 3) la importancia del sector institucional, facilidades de cooperación; y 4) estrategia industrial concentrada en la producción en lugar de la distribución, orientadas a formar mercados, estimular y llevar a cabo inversiones complementarias en sistemas de soporte y a propiciar la celebración de alianzas estratégicas.

1.2 El Transporte: Barrera Al Comercio Interregional

Cuando se habla de “las barreras al comercio” se entienden como las desventajas relativas que enfrentan las empresas para ofrecer sus productos en mercados alejados de su sede respecto de los oferentes locales (Viego, 2004). El punto importante es aclarar que la única traba que aquí se aborda es la infraestructura del transporte, que limita a las empresas ubicadas en regiones diferentes de México, complicando que se conecten a diversos mercados, incluso a sus proveedores y, no tiene nada que ver con las definiciones que se encuentran en el comercio internacional o disposiciones legales. Dentro de los obstáculos que las empresas encuentran al momento de querer penetrar nuevos mercados son los costos de transporte y el tiempo-distancia. Así, se tiene que priorizar la atención a las barreras de transporte a las que se enfrentan las empresas manufactureras.

1.2 Costos Del Transporte

Las relaciones comerciales conllevan ubicaciones en distintos puntos, desde la empresa que se encarga de las materias primas, que es la que provee a las que van a producir los bienes hasta los puntos donde se distribuyen y comercializan los productos finales (punto de origen – puntos de distribución y producción - punto de venta). Como argumenta Polése (1998), cubrir distancias entre todos los puntos conlleva esfuerzos, recursos y tiempo, lo que para las empresas implica costos, mismos que están en función de la distancia, el costo de producción se ve afectado por el costo de transporte y son cubiertos por el consumidor, por otro lado, los costos económicos que da la distancia no son solo directos; es decir, estos costos económicos que ocasionan un desplazamiento se mide también por el valor de tiempo utilizando.

Por su parte Olsson (2009) sostiene que el efecto agregado positivo de carreteras, mejores y más largas se basa en una mejora de la eficiencia de costes en el transporte de mercancías a mercados. Los costos de transporte más bajos se explican por los tiempos de viaje más cortos combinados con menor pérdida en el camino, el acceso directo al mercado. La disminución de los costos de transporte se hace efectiva cuando una empresa se encuentra en un lugar céntrico y acompañado de economías de escala.

Las relaciones de intercambio y las demás relaciones entre agentes económicos dan origen a un lugar céntrico, en la realidad, es complicado que todas las empresas puedan concentrarse en el mismo lugar. A lo que Polése (1998) llama “el juego de los costos de transporte⁵ y de las economías de escala”, entonces las

⁵ Los costos de transporte desde la visión de Polése siempre estarán ligados a las economías de escala, argumenta que, Las economías de escala están en el campo del transporte. Los costos unitarios no aumentan proporcionalmente la distancia recorrida. Un viaje corto representa los mismos costos que un viaje largo. La cura de costos de transporte tendera a disminuir. Pueden tomar también otra forma, los costos no aumentan de manera proporcional a la cantidad transportada, el transportista tendrá ganancias importantes si logra utilizar sus equipos a plena capacidad en ambos sentidos.

empresas que quedan fuera del centro se ven obligadas a maximizar las ventas a partir de rendimientos de escala ya que no pueden minimizar sus costos de transporte.

Polése (1998) argumenta que la ciudad nació en el centro de la región para reducir los costos de interacción espacial de sus habitantes y sacar provecho de las economías de escala en la producción de bienes y servicios. Al igual que Viego (2004) la zona central, por ejemplo, posee mayor capacidad de atracción de actividades debido a la importancia del mercado (propio o próximo). De este modo, una región central es aquella que posee un mercado propio grande y/o una ubicación favorable en términos de acceso a otras regiones.

1.2.2 Infraestructura Del Transporte

La infraestructura genera el soporte físico común de los distintos modos de transporte. Las infraestructuras son consideradas con frecuencia como bienes públicos, debido a; sus elevados costes fijos, como de mantenimiento y explotación, que hacen que en numerosas ocasiones no puedan ser asumidos de forma exclusiva por un ente privado a precios de mercado y por su importante dimensión económica y social. Otra posibilidad es que aun pudiendo ser de titularidad privada, requiere de una regulación por el Estado, para garantizar con ello los derechos del ciudadano y velar por la seguridad de la circulación en cualquier modo de transporte. Las infraestructuras de transporte juegan un papel considerable en el desarrollo económico y regional de la zona en la que se ubiquen (Cendrero, 2012).

La dotación de infraestructura afecta el crecimiento, en la medida que una mayor disponibilidad y calidad de estos servicios conlleva a una mayor productividad de los factores y costos de producción más bajos para los productores. La mayor rentabilidad incentiva la inversión y, por ende, aumenta el crecimiento potencial del producto (CEPAL, 2004). En el paradigma actual de la globalización en la que las economías se encuentran sumergidas, tienen que estar necesariamente en

constante evolución e irse adaptando a los requerimientos de desarrollo y crecimiento para su población. Como argumenta Asuad (2001) la expansión y revolución tecnológica de los servicios, transporte y telecomunicaciones e informática han reforzado el patrón de globalidad y regionalismo económico, permitiendo la integración entre la localidad y la región y la economía mundial, articuladas a través de redes de ciudades mundiales, a lo que Castells y Borja (1998) llaman redes globales de nodos urbanos/regionales.

Uno de los pilares en los que se tiene que enfocar el gobierno es en la infraestructura, como menciona Vázquez (2005), para que una economía pueda aumentar su productividad y competitividad, debe contar con un sistema de infraestructuras económicas y sociales. Cuando se habla de una infraestructura económica se refiere a la infraestructura que apoya directamente las actividades productivas, por ejemplo, transporte y comunicaciones. De acuerdo a Owen (1959), los volúmenes de transporte y los niveles de actividades económicas en una nación tienen una estrecha relación, gracias a que la mayoría de las actividades económicas involucran la presencia del uso del transporte para que las haga funcionales.

En la actualidad en México, si bien es cierto que la infraestructura pública muestra avances importantes en la creación de elementos y espacios que se requieren para llevar a cabo los procesos económicos, estos han sido insuficientes comparados con el crecimiento que han tenido las economías desarrolladas. México al igual que países de América Latina presenta características similares en inversión pública de infraestructura, las cuales siguen sin mostrar resultados de productividad y competitividad a nivel internacional. Las infraestructuras de transporte dinamizan la centralización de la actividad económica.

La interacción social conlleva a una interacción espacial, el espacio geográfico es una variable con la cual el hombre debe convivir, para una economía es fundamental las relaciones entre los seres humanos (Polése, 1998). Esta relación

necesita de una conexión, de caminos que faciliten las relaciones personales, pero aún más importante en el tema de análisis, de las relaciones comerciales. Asimismo, es importante exponer los factores que se interponen en una interacción efectiva.

Polése (1998) expone que las vías de comunicación de mejor calidad siempre serán los primeros que reclamen a los gobernantes los habitantes de las regiones. En su mayoría habitantes que se encuentran alejados de las grandes ciudades. Son indiscutible las consecuencias económicas por la utilización del transporte como determinante para la producción y el consumo; es decir, la actividad del transporte es fundamental para una economía que quiere crecer y desarrollarse.

1.2.2.1 Inversiones En Infraestructura

Si bien es cierto que se ha estado hablando de la centralidad y cómo ésta influye en los costos de transporte. La infraestructura es necesaria para el paso del transporte, y el agente principal para reducir la brecha que se tiene en materia de infraestructura es el Estado a través de políticas regionales y de desarrollo regional (Polése, 1998). Según Polése (1998) los países industrializados fomentan la instalación de las industrias en regiones desfavorecidas al otorgar primas regionales. Aunque su objetivo es la distribución geográfica de las industrias por su impacto en las decisiones de localización distan mucho de haber dado pruebas de ser un instrumento eficaz de intervención. Se tiene poco éxito de las primas que busca compensar los inconvenientes de una mala localización para las empresas y no son suficientes las subvenciones para equiparar su localización.

La región asume costos económicos impuestos por la distancia, incluso las primas no son efectivas reduciéndolos. Las primas regionales solo cumplen un papel marginal en las decisiones de localización de las empresas e influyen únicamente en los sectores que ya sienten cierta atracción por la región designada, su impacto es más limitado si la región se encuentra lejos de los mercados importantes.

Polése (1998) considera que se puede orientar la industria por otros medios, las inversiones en las infraestructuras públicas en regiones desfavorecidas. En lugar de subvencionar para que se instalen, se pretende dotar a las regiones de un capital público para que las empresas se sitúen en la zona. Las inversiones en infraestructura son un incentivo del cual dispone el Estado para influir en las tendencias de localización. En la economía mexicana invertir capital público en regiones desfavorecidas es un problema de elección para el gobierno ya que es complicado mantener en el tiempo inversiones en este tipo de regiones porque los resultados son muy lentos y escasos. La existente disparidad regional que se encuentra en las regiones mexicanas, el favorecer siempre a las regiones centrales ya que en ellas se obtienen mayores beneficios y seguir desprotegiendo a las regiones periféricas es uno de los problemas que asumen las economías en desarrollo.

Canning y Bennathan (1999) abordan el impacto de las redes de transporte en el crecimiento y llegan a la conclusión de que la tasa de retorno de la inversión en infraestructura de transporte implica efectos sobre el crecimiento, la utilización del capital público para invertirlo en infraestructura es rentable para cualquier economía. Como lo demuestran los estudios macroeconómicos de Aschauer (1989) sobre los efectos de la infraestructura, con datos de Estados Unidos. Estiman una producción nacional donde el PIB o el crecimiento no sólo dependen de la mano de obra, capital y tecnología, sino también de capital público.

De la misma forma, Canning y Bennathan (1999) encuentran para algunos países de renta media la inversión en infraestructura está en un nivel inferior al óptimo, dejando a los países con escasez de infraestructura graves. Lo cual apoya la idea general de que la mejora de los caminos conduce a una mayor producción y bienestar.

En América Latina, el principal componente de la inversión pública era la inversión en infraestructura. Hasta 1980 fue cuando la tendencia decreciente de ese componente fue más pronunciada. La disminución de la inversión pública refleja, en buena medida, la menor participación del gobierno en las economías de la mayor parte de los países latinoamericanos (CEPAL, 2004).

A pesar de ser conocidos los efectos virtuosos del desarrollo de la infraestructura sobre el crecimiento económico, recurrentemente existe la tendencia a reducir la inversión en el sector cada vez que se enfrentan situaciones de crisis. La ausencia de infraestructura adecuada, así como la provisión ineficiente de servicios de infraestructura, obstaculiza la implementación eficaz de políticas de desarrollo y la obtención de tasas de crecimiento que superen los promedios internacionales para alcanzar un mayor grado de especialización productiva (Tomassian *et al.*, 2010).

CAPITULO 2 ANÁLISIS DESCRIPTIVO – ESPACIAL

2.1 Discusión Teórica Del Análisis Descriptivo

A partir del trabajo de Tukey (1977) se han comenzado a afianzar y difundir las técnicas del análisis exploratorio de datos (EDA, *Exploratory Data Analysis*) como herramienta que permite realizar un estudio inicial de los datos de las variables en forma de aproximaciones gráficas, con la ventaja de realizar procedimientos estadísticos de mayor complejidad, como por ejemplo, el análisis de regresión simple (Buzai, 2007). Anselin (1988), impulsor de la econometría espacial considera al EDA como el campo formado por métodos y técnicas que, sustentados en una representación formal de la estructura de la dependencia y heterogeneidad espacial, provee el medio para llevar a cabo la adecuada especificación, estimación, prueba de hipótesis y predicción para modelos en la ciencia regional.

El análisis exploratorio de datos se considera un instrumento indispensable al momento de que se van a realizar las primeras aproximaciones al estudio de la estructura de la información espacial. Para Chasco (2003) es un grupo de técnicas que describen y visualizan las distribuciones espaciales y siguiendo a Anselin (1999) dicho enfoque es el procesamiento estadístico con la particularidad de las descripciones geográficas de los datos y la clase de técnicas son de tipo visual y numérico en el sentido de ser resumen de los datos (Tukey, 1977).

Es importante destacar que, en este enfoque el análisis de datos se basaba fundamentalmente en el cálculo de estadísticos, conduciendo a dos consecuencias: En primer lugar, se disminuía la importancia visual de la representación de las observaciones obteniendo exclusivamente datos y cálculos. En segundo lugar, se equiparaba el análisis con el modelo confirmatorio y hasta ahí se terminaba el estudio (Batanero *et al.*, 1991).

Anselin (1999) resume en tres formas el uso del enfoque espacial para los científicos de las ciencias sociales a saber: a) la base para integrar datos (*data integration*), en tanto la información, atributo o característica de escalas espaciales y dimensiones temporales particulares se pueden observar en otros tipos de escala y dimensiones (ESDA); b) la necesidad de chequear relaciones en diferentes direcciones, y c) observar dichos fenómenos cartográficamente en combinación de escalas micro y macro, convierte a los SIG (Sistema de Información Georreferenciada) en herramienta potente para soportar el análisis de investigación social.

Con la información adicional que sugiere la georreferenciación se pueden mostrar correlaciones espaciales y análisis de vecindad. Muñetón y Vanegas (2014) mencionan que lo atractivo del enfoque espacial es el acompañamiento visual georreferenciado de los datos. Mientras tanto, otra característica básica del espacio geográfico es la multidireccionalidad y la multidimensionalidad de las relaciones que se establecen. La necesidad de la existencia de un Análisis Exploratorio de Datos Espaciales constituyó una tendencia importante del desarrollo metodológico cuantitativo, potenciado por los nuevos desarrollos digitales en materia de representación (Fotheringham *et al.*, 2000).

El EDA es una colección de técnicas para resumir propiedades de los datos (estadística descriptiva), el ESDA (*Exploratory Spatial Data Analysis*) detecta patrones en los datos, identificando características interesantes o inusuales, detecta errores, distingue características accidentales de las que son relevantes y formula hipótesis a partir de los datos (Good, 1983).

El Análisis Exploratorio de los Datos se toma para examinar resultados de modelos, proporcionar evidencia sobre los supuestos del modelo que se cumplen y, determinar si existen efectos influyentes de los datos en el modelo estimado. No hace supuestos sobre la distribución de los datos y generalmente excluye las pruebas de hipótesis; no emplean teoría inferencial. Son técnicas principalmente visuales (tablas, gráficas, figuras) y/o cuantitativas como resúmenes de los datos.

Utiliza estadísticos “resistentes”, que no se afectan mucho por la presencia de valores extremos (Anselin, 1998).

Mientras que el ESDA sintetiza las propiedades espaciales de los datos, detecta patrones espaciales, formula hipótesis de la geografía de los datos, identifica casos y sub conjuntos de casos inusuales dados su localización en un mapa. Al igual que el EDA, son técnicas visuales y numéricamente resistentes. A diferencia del EDA, aquí el mapa asume un papel central ya que busca detectar propiedades de los datos a través de visualización cartográfica (Anselin, 1988).

En resumen, el Análisis Exploratorio de Datos es un modo de análisis de los datos que utiliza el resumen numérico y visual para explorar datos en busca de patrones no anticipados. Autores clásicos como Tukey (1980), Hartwig y Dearing (1979) lo catalogan como un “estado mental” ante el conocimiento, como un reconocimiento de que el ojo que mira al horizonte es el mejor instrumento que tenemos para observar, de manera completa lo no anticipado (Tukey, 1980). Es necesario aprender todo lo posible acerca de una variable o conjunto de variables antes de utilizar los datos para probar hipótesis o teorías acerca de las relaciones sociales (Hartwig y Dearing, 1979). Se tiene el fin de que el “razonamiento” exploratorio sea un esquema de análisis que engrandece las posibilidades del investigador para hallar nuevas respuestas a los problemas que se plantee (Horber y Ladiray, 1995).

2.2 Análisis Exploratorio de los Datos

Esta investigación presenta una técnica novedosa que supone la unión de dos tecnologías informáticas muy poderosas: los GIS (Sistemas de Información Geográfica) y los paquetes estadísticos propios del análisis de datos espaciales. Por lo tanto, al igual que Chasco *et al.* (2000) las técnicas adecuadas para el análisis estadístico-gráfico-cartográfico es a través de herramientas que llevan el nombre de análisis exploratorio de datos espaciales (AEDE) y se conciben como una disciplina dentro del más general análisis exploratorio de datos (AED) también

denominado “*data mining*”, que ha sido diseñado para el tratamiento específico de los datos espaciales o geográficos.

El AEDE combinan el análisis estadístico con el gráfico, haciendo posible el estudio de las distribuciones espaciales y sus valores atípicos, esquemas de asociación espacial, agrupamientos espaciales, entre otros. Se hace uso de los programas informáticos, GeoDa⁶, concebido como un producto autosuficiente que no requiere de un sistema específico de SIG (Sistema de Información Georreferenciada) tiene la ventaja de ser, hasta el momento, un producto de libre acceso en Internet. Al igual que QGIS (anteriormente llamado también Quantum GIS) que es un Sistema de Información Geográfica (SIG), una de las grandes fortalezas de este software es que trabaja en cualquiera de los sistemas operativos, además de su facilidad de interconexión con muchas bases de datos geoespaciales utilizando toda la potencia de análisis de este último en un entorno de trabajo más amigable.

2.2.1 Especificación De Las Variables

Se debe de tener en cuenta que, a pesar de utilizar las técnicas y herramientas econométricas espaciales más actualizadas, tanto en el capítulo 2 como en el capítulo 3, son insuficientes las variables recolectadas para poder desarrollar un modelo más robusto y que se pueda acercarse a la realidad. Decía Asuad (2001), el reto de la teoría económica regional es que existen restricciones considerables respecto a ella motivo, por el cual ha sido relegada por mucho tiempo y la falta de información y los escasos de datos regionales pertenecen a esas restricciones.

⁶ GEODA, software desarrollado por el Profesor Luc Anselin de la Universidad de Illinois.

Tabla No. 2.1 “ESPECIFICACIÓN DE VARIABLES”

| CODE VARIABLE | DESCRIPCIÓN | UNIDADES |
|------------------|--|----------------------------|
| FMM* | Flujo de Mercancías del Sector Manufacturero | Miles de pesos (B=2012) |
| LC | Longitud Carretera (Km/superficie territorial) | Kilómetros |
| LF | Longitud Ferroviaria (Km/Km(superficie territorial) | Kilómetros |
| UVC | Unidades Vehiculares de Carga | Unidades |
| VCF | Flujo de carga por ferrocarril | Miles de Toneladas |
| IPT | Inversión Pública en el Sector Transporte | Miles de pesos (B=2012) |
| EM | Establecimientos Manufactureros | Unidades |
| POT | Personal Ocupado Total | Unidades |
| FBI ⁱ | Flujo de Bienes que Abastecen los Establecimientos Manufactureros al Mercado Interno | Miles de pesos (B=2012) |
| FBE | Flujo de Bienes que Abastecen los Establecimientos Manufactureros al Mercado Externo | Miles de pesos (B=2012) |
| FBC | Flujo de Bienes Consumidos por los Establecimientos Manufactureros | Miles de pesos (B=2012) |
| A | Numero de Aeropuertos | Unidades |
| P | Numero de Puertos | Unidades |

Fuente: Programa de la industria manufacturera, maquiladora y de servicios de exportación INEGI
2008-2013. Elaboración propia.

Las variables analizan al sector manufacturero enfocado al sector que comercializa bienes maquilados para el exterior. Sin embargo, se hace énfasis en; la productividad del sector y en la interacción de comercio de bienes entre regiones antes de que salgan al exterior en los años 2008 y 2013. El análisis está basado en variables deflactadas, normalizadas y tratadas en logaritmos. La fuente de información proviene de los anuarios estadísticos por entidad federativa 2009, 2014 y 2015 de INEGI (Instituto Nacional de Estadística y Geografía) y del Programa de la Industria Manufacturera, Maquiladora y de servicios de Exportación (IMMEX).

Las variables endógenas son dos (con la justificación de obtener dos diferentes resultados o modelos. Al final del capítulo 3 se elegirá el modelo con la variable más significativo para el análisis); Productividad del sector manufacturero (PROD) y Flujo de Mercancías del Sector Manufacturero (FMM) tomada como variable proxy para mostrar el comercio regional o interno. Mientras que las variables exógenas son; Longitud Carretera en Km/superficie territorial (LC), Longitud Ferroviaria en Km/superficie territorial (LF), Unidades Vehiculares de Carga (UVC), Flujo de carga por ferrocarril en miles de toneladas (VCF), Inversión Pública en el Sector Transporte en miles de pesos (IPT), Establecimientos Manufactureros (EM), Flujo de Bienes que Abastecen los Establecimientos Manufactureros al Mercado Interno en miles de pesos (FBI), Flujo de Bienes que Abastecen los Establecimientos Manufactureros al Mercado Externo en miles de pesos (FBE),

Flujo de Bienes Consumidos por los Establecimientos Manufactureros en miles de pesos (FBC), Número de Aeropuertos (A) y Número de Puertos (P). Todas las variables monetarias fueron transformadas de dólares a pesos y deflactadas con el índice de precios del año 2012 para finalmente convertirlas a logaritmos⁷. También es notable mencionar que la longitud carretera y de vías férreas se normalizó con la superficie territorial de cada estado Km/Km² (LC y LF).

⁷ Como regla general, se trata de escoger una transformación que conduzca a una distribución simétrica, y más cercana a la distribución normal. De este modo, se pueden aplicar numerosas técnicas de inferencia estadística (Marin, J, PDF, Tema 4. Transformación de variables.)

De igual forma, es preciso señalar que se construye el índice de productividad usando la variable población ocupada total y valor agregado censal del sector manufacturero a dos dígitos (sector manufacturero total) y, será como ya se mencionó una de nuestras variables endógenas. La construcción del índice o coeficiente sigue la línea de Keeble *et al.*, (1986) que habla de los índices como la distinción entre una región y otra que puede realizarse con base en el índice de centralidad (mismo que él propuso):

$$\frac{Y_i}{Y_j} [1]$$

Donde i representa la región bajo análisis, j indica al resto de las regiones, Y_i e Y_j son variables que indican el tamaño de los mercados de cada aglomerado. Y_i es el valor agregado censal bruto del estado o entidad, Y_j es la población ocupada total del estado, Y_k es el valor agregado censal bruto nacional y por último es la población ocupada total nacional. Al nivel de desagregación del sector manufacturero a dos dígitos (ramo 31).

Este índice suele ser considerado como una proxy del tamaño de mercado disponible para una firma localizada en i , representado tanto por el propio mercado como el de autoridades cercanas. Esta relación permite evaluar el rendimiento de una unidad económica durante un periodo determinado. Si en el transcurso del tiempo aumenta la relación entre el valor agregado y la magnitud de la población ocupada total incorporada, ello significa que la productividad del sector manufacturero ha mejorado, y viceversa si disminuye.

La medición de la productividad sirve para evaluar la eficacia con la cual se usan los factores de producción. Puede realizarse en el ámbito de un establecimiento, de una empresa, de una industria, de un sector o de un país. Tales mediciones se realizan mediante índices, debido a que se pretende observar su evolución a través del tiempo. Incluso los índices tienen la ventaja de proporcionar una unidad de medida uniforme para los distintos sectores productivos o entidades (INEGI, 2015).

2.2.2 Exploratory Data Analysis (EDA) – Exploratory Spatial Data Analysis (ESDA)

2.2.2.1 Análisis Exploratorio de los Datos (Exploratory Data Analysis-EDA)

Las siguientes figuras ilustran las técnicas básicas para el análisis exploratorio de los datos o EDA cubriendo la visualización de la distribución no-espacial de los datos. La ilustración tradicional del EDA se resume en un histograma. La figura 2.1 es el resultado de un histograma con la variable población ocupada en el sector manufacturero exportador 2009 (POT_09), y la figura 2.2 muestra la variable de establecimientos manufactureros de exportación 2009 (EM_09) clasificadas en 5 categorías. Al comparar los dos histogramas se puede comprobar la distribución del número de observaciones en cada categoría y los rangos de valores para cada una de ellas.

Aquí, el único aspecto diferente de las dos variables es donde se encuentran los valores, no las características no espaciales de la distribución. Es decir, que independientemente de que las observaciones de las variables no sean iguales, tienen casi la misma forma de la distribución que obviamente no corresponde a una distribución normal sino a una distribución asimétrica a la derecha.

Es importante mencionar que más del 60% de las entidades en los dos casos se encuentra dentro de los valores mínimos y medios, contrariamente de dos estados que se encuentran dentro de los valores más altos: Baja California y Chihuahua.

Figura 2.1 Histograma POT 2008

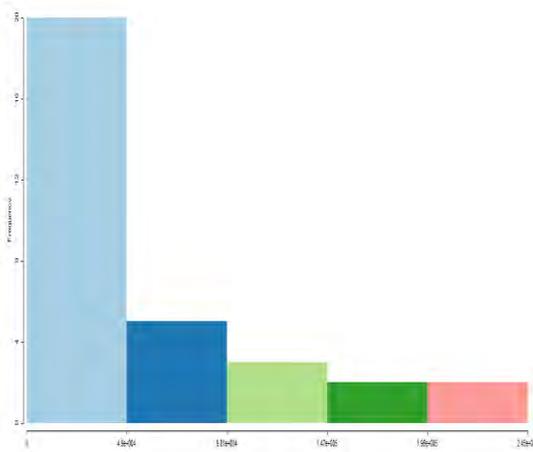
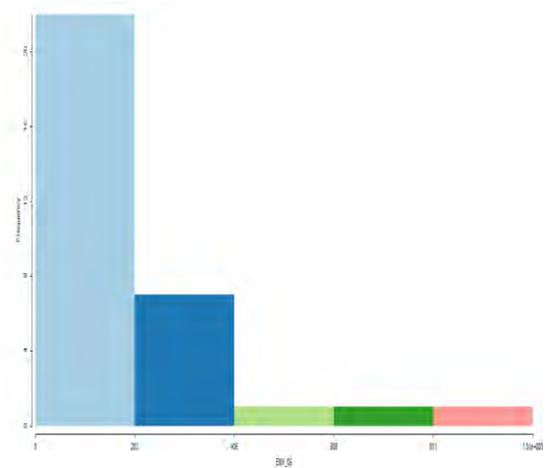


Figura 2.2 Histograma EM 2008



Fuente: Censo económico 2009 INEGI. E. Propia Fuente: Censo económico 2009 INEGI. Elaboración Propia

La segunda técnica básica EDA es el llamado “diagrama de caja” que se usa para representar la distribución no-espacial de una variable (Anselin, 2005). La figura 2.3 y 2.4 muestra este gráfico el cual se divide en cuartiles con los datos distribuidos en la misma proporción, 25% en cada cuartil (25%, 50%, 75% y 100%), la parte oscura muestra el rango intercuartílico (que va desde el cuartil 25 al cuartil 75). También se observa la mediana, la barra roja en la mitad de la caja. Así como una noción de valor atípico.

Una observación se clasifica como un valor atípico cuando se encuentra a más de un múltiplo dado del rango intercuartílico, tanto por encima como por debajo de él. Los valores atípicos para las dos figuras son Baja California, Chihuahua y Nuevo León. Se comprueba que tanto las figuras 2.1 y 2.2 (histogramas) como las de *box-plot* (2.3 y 2.4) tienen resultados similares al presentar casi los mismos valores atípicos, los únicos que no se encuentran dentro de la mayoría de los estados.

Figura 2.3 Box-Plot POT 2008

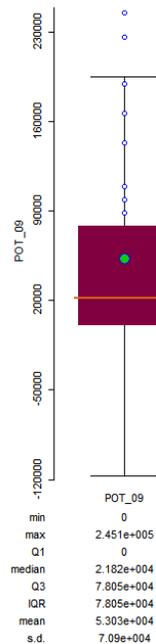
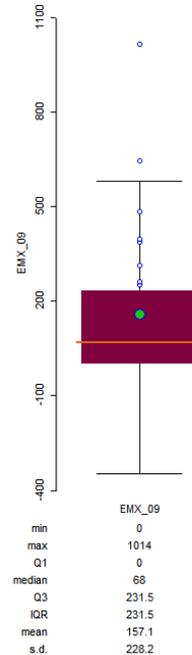


Figura 2.4 Box-Plot EMX 2008



Fuente: Censo económico 2009 INEGI. Elaboración Propia. Fuente: Censo económico 2009 INEGI. E.P

Las figuras anteriores presentan una vista rápida de variables que ilustran características del análisis exploratorio de datos, como ya se ha mencionado es un análisis que se supone corto sin tener que referirse a una información más extensa, por el contrario, el siguiente subtema se estudia con base en cartografía para resaltar el tema espacial con información visual más precisa.

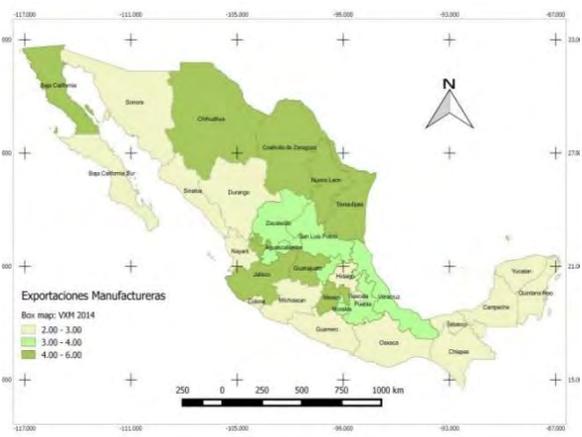
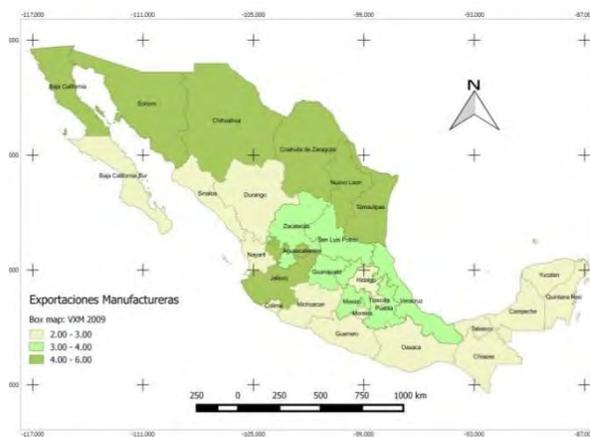
2.2.2.2 Análisis Exploratorio Espacial de los Datos (Exploratory Spatial Data Analysis-EDA)

Con la ayuda de diferentes tipos de mapas se comienza a tratar a los datos para ser analizados de manera espacial. En la figura 2.5 se muestran los valores atípicos en el primer y cuarto cuartil representado por el *box map* con la variable del valor de las exportaciones manufactureras, el *Upper Outlier* del 75% para el año 2008 describe a Baja California, Coahuila, Chihuahua, Jalisco, Nuevo León, Sonora, Tamaulipas y Aguascalientes como las que sus bienes manufactureros

generan más valor por el contrario el *Lower Outlier* del 25% muestra a Baja California Sur, Campeche, Colima, Guerrero, Nayarit, Quintana Roo y Tabasco como las de menos valor en manufacturas de exportación. Para el año 2013 el *upper outlier* cambió a solamente dos entidades; Baja California, Chihuahua y Coahuila. Sorprende que Sonora descendiera a un nivel antes de *lower outlier* cuando en el 2008 era de los más concentrados.

Figura 2.5 Box map para VXM8 - 1.5
hinge

Figura 2.6 Box map para VXM13 - 1.5
hinge.



Fuente: Censo económico 2009 INEGI. Elaboración Propia

Fuente: Censo económico 2014 INEGI. Elaboración Propia

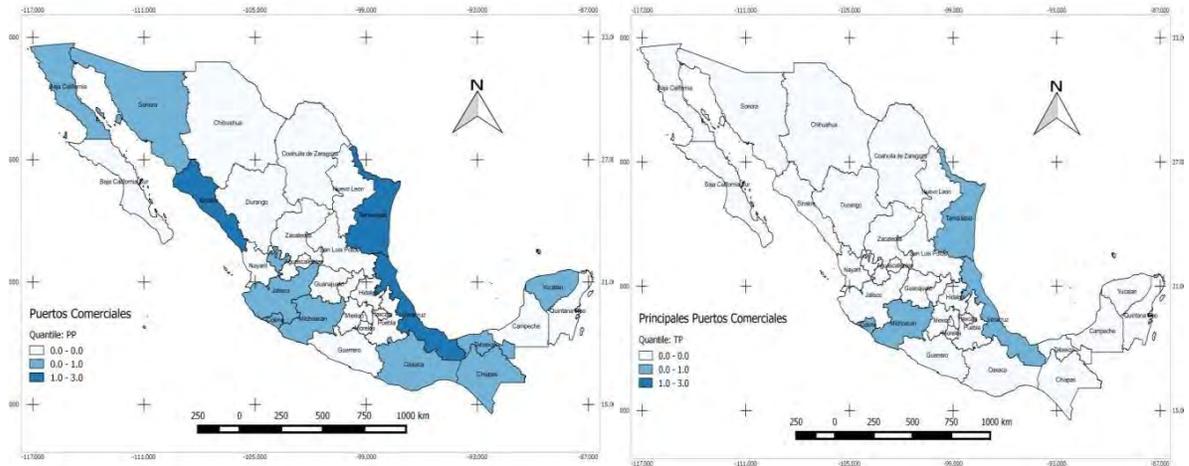
A lo largo de la investigación se ha intentado recalcar la importancia de lo que es la infraestructura para la movilidad de flujos de mercancías para que una región o nación sea productiva, este análisis en específico trabaja con bienes manufactureros. En primer lugar, desde la perspectiva de los puertos⁸ la figura 2.6 representada por un mapa de cuantiles en dos clases muestra el 67% del movimiento de carga concentrado solo en 16 puertos comercialesⁱⁱ, de los cuales en la figura 2.7 se muestran únicamente los 6 más importantes, Manzanillo en Colima, Lázaro Cárdenas en Michoacán, Altamira en Tamaulipas y Veracruz en

⁸ Puerto: El lugar de la costa o ribera habilitado como tal por el Ejecutivo Federal para la recepción, abrigo y atención de embarcaciones, compuesto por el recinto portuario y, en su caso, por la zona de desarrollo, así como por accesos y áreas de uso común para la navegación interna y afectas a su funcionamiento; con servicios, terminales e instalaciones, públicos y particulares, para la transferencia de bienes y transbordo de personas entre los modos de transporte que enlaza (Ley de Puertos de la Secretaría de Comunicaciones y Transportes, 1993).

Veracruz, operando el 96% de la carga contenerizada del 67% que representan los 16 puertos.

Figura 2.6 Puertos Comerciales A

Figura 2.7 Puertos Comerciales B



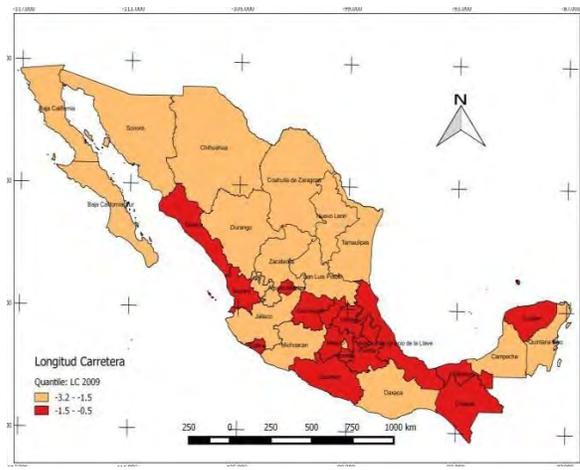
Fuente: Anuarios Estadísticos INEGI.

Fuente: Anuarios Estadísticos INEGI.

El transporte terrestre es fundamental para el comercio entre regiones, a diferencia del transporte marítimo éste juega el papel de la conectividad en todo el país para satisfacer un flujo de mercancías que proviene de diferentes puntos y que a su vez se dirigen a diferentes destinos. La figura 2.8 y 2.9 muestran a las regiones más vivas con la mayor longitud carretera, se compara al año 2008 y 2013 obteniendo a estados como Sinaloa, Jalisco y Guerrero como los que en 4 años incrementaron su infraestructura de carreteras.

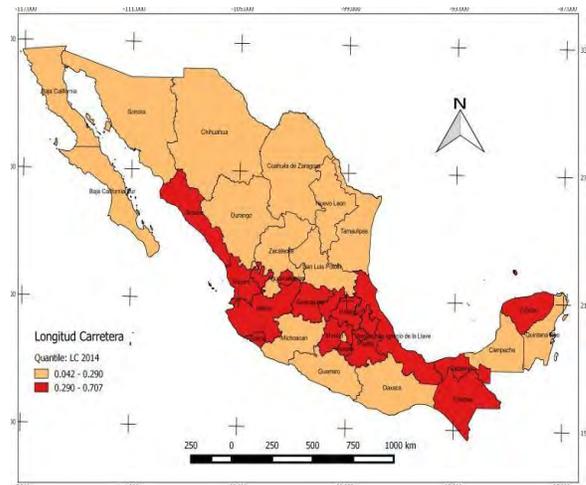
Estas dos figuras comprueban los supuestos de la econometría espacial al proyectar a la zona centro del país como la más concentrada al tener vías terrestres para su comunicación con regiones vecinas.

Figura 2.8 Longitud Carretera de México 2008



Fuente: Anuarios Estadísticos INEGI. Elaboración Propia

Figura 2.9 Longitud Carretera de México 2014

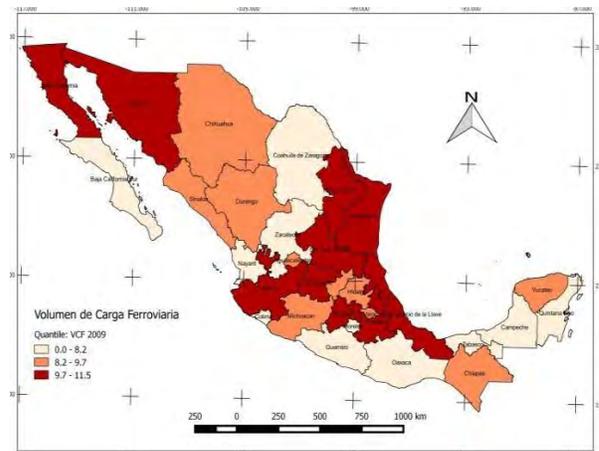
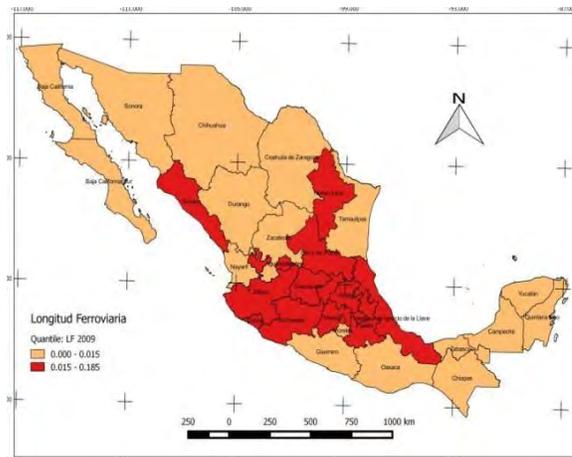


Fuente: Anuarios Estadísticos INEGI. Elaboración Propia.

Si de infraestructura terrestre se trata, otro elemento importante es la vía ferroviaria y pese a que México cuenta con vías férreas desde aproximadamente inicios de los años 1800, actualmente son vías mejoradas, renovadas e incluso existen nuevas. Como podemos ver en la figura 2.10, la región del centro y parte del norte son las que cuentan con más kilómetros de vías. Estas regiones son similares a las de la figura 2.10 que muestra las de mayor volumen de carga en toneladas transportadas por ferrocarril a excepción de Sonora y Baja California como entidades que su volumen transportado no está en función de las vías férreas con las que cuentan, finalmente Sinaloa que, aunque se encuentra dentro de la región con más vías para ferrocarril no refleja un volumen de carga en proporción a ellas.

Figura 2.10 Longitud Ferroviaria 2008

Figura 2.11. Volumen Carga Ferroviaria 2009

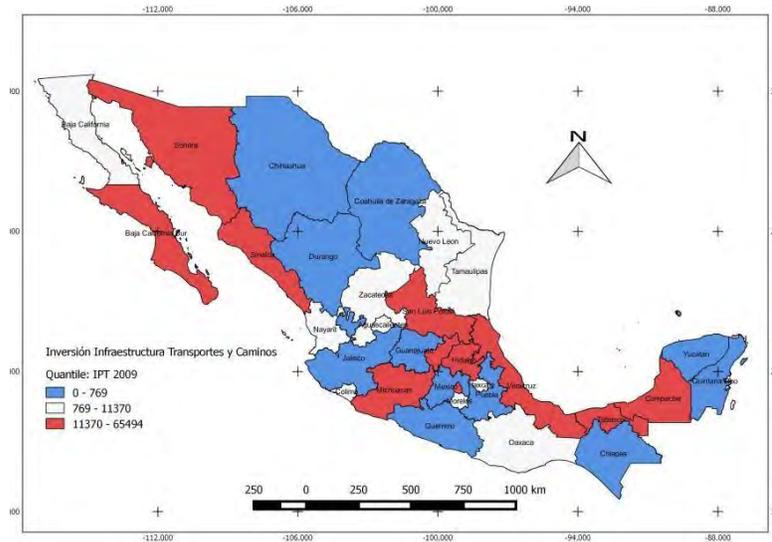


Fuente: Anuarios Estadísticos INEGI. Elaboración Propia

Fuente: Anuarios Estadísticos INEGI. Elaboración Propia

La Figura 2.12 que se presenta a continuación es la más significativa e idónea para contrastar las figuras anteriores. El mapa muestra la Inversión Pública en infraestructura en el año 2008. Como se aprecia en todo el país, las regiones más oscuras son las que presentan mayor participación de la inversión en transporte, como en los casos del estado de México, Chihuahua y Chiapas. Por el contrario la Ciudad de México, Sonora y Campeche son los de menor inversión pública. Esto apoya la necesidad de hacer uso de las figuras anteriores para un análisis más completo, en el caso del mapa de longitud carretera (figura 2.8) un supuesto del aumento de vías carreteras en Sinaloa, Querétaro, Hidalgo es que se destinó más presupuesto a la infraestructura en el año 2008 a esas entidades.

Figura 2.12 Inversión Pública para Transportes y Comunicaciones 2008.

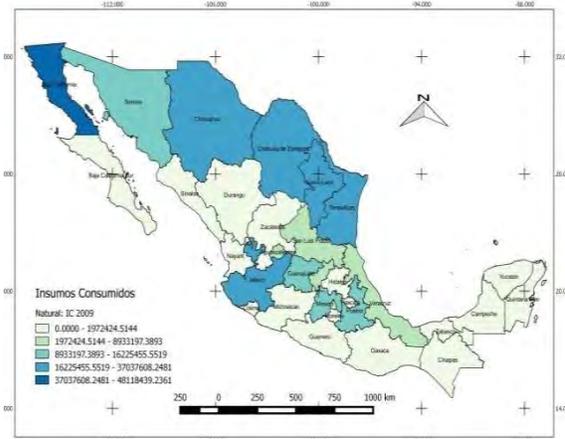
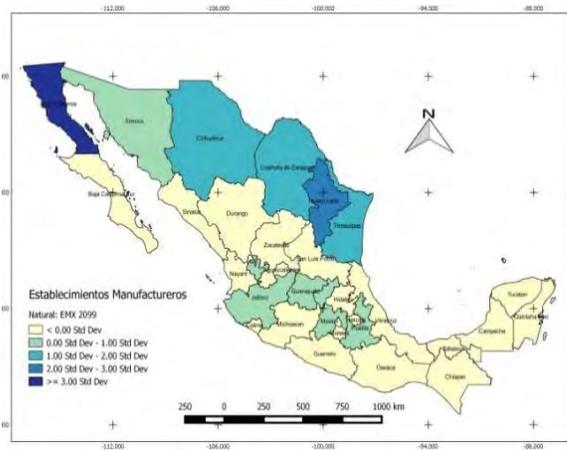


Fuente: Anuarios Estadísticos INEGI. Elaboración Propia

Hasta este punto el análisis ESDA y con ayuda de los datos que provienen del programa industria manufacturera maquiladora (INEGI, 2009) se relacionan las figuras; 2.5 (valor de las exportaciones de mercancías), 2.10 (longitud ferroviaria) y 2.12 (Inversión pública para transportes). Todas muestran que las regiones; Ciudad de México, Estado de México, Sinaloa, Sonora, Baja California, Chihuahua, Nuevo León, por mencionar algunas son las que a partir de un alto nivel de inversión pública pueden dotar de más y/o mejor infraestructura terrestre para que su nivel de valor de exportaciones aumente.

Figura 2.13 E. Manufactureros 2008

Figura 2.14 Insumo Consumidos 2008



Fuente: Programa IMMEX INEGI. Elaboración Propia

Fuente: Programa IMMEX INEGI. Elaboración Propia

A modo de conclusión para el análisis exploratorio, la variedad de técnicas posibles que da el EDA se toman en cuenta como herramientas que permiten realizar un análisis inicial de los datos en aproximaciones gráficas, es el paso previo a procedimientos estadísticos más complejos como las regresiones pero también es la antesala del ESDA, este estudio tienen como objetivo formular supuestos e hipótesis a partir de la información de los datos representados cartográficamente, de esta forma la evolución de obtener datos estadísticos a tener una visualización más dinámica e interactiva. Mencionaba Buzai (2007), la estadística básica lleva a resultados a partir de la aplicación de procedimientos aritméticos, mientras que a través del análisis exploratorio de datos estas características se hacen evidentes de forma visual mediante la realización de gráficos específicos, los cuales permiten resaltar sus principales cualidades.

2.2.3 Dependencia Espacial

El índice I de Moran permite obtener una visión global de los procesos de interrelación espacial de las regiones, pero no ofrece la estructura regional de la autocorrelación. En particular, podemos preguntarnos cuáles son las regiones que contribuyen en mayor medida a la autocorrelación espacial global, o si hay concentraciones locales, o incluso en qué medida la evolución global de la autocorrelación espacial enmascara localizaciones atípicas.

El análisis de la autocorrelación espacial local se efectúa a través de dos instrumentos: el gráfico de Mora (Anselin, 1996) y los indicadores locales de asociación espacial. Estos últimos están diseñados para contrastar la hipótesis nula de distribución espacial aleatoria comparando los valores de cada localización específica con los valores obtenidos en las localizaciones vecinas. Diseñado por P. A. Moran (1950).

Su notación es:

$$\frac{\sum \sum}{\quad}$$

Es la ponderación que cuantifica la asociación espacial hipotética o proximidad entre las observaciones , en la posición *i* y *j*, es la varianza de *n* observaciones y *W* es la suma de todos los valores de de En el cociente de la fórmula del estadístico, el numerador es un término que muestra la covarianza, mientras que el denominador indica la varianza. Esto hace que *I* se comporte como un coeficiente de correlación de Pearson y se encuentre entre +1 y -1, aunque puede superar ambos límites (Cliff y Ord, 1981; Upton & Fingleton, 1985).

Si los puntos situados a una distancia concreta muestran valores similares o muy parecidos para la variable medida, el numerador se aproxima al producto de dos desviaciones positivas o de dos negativas, obteniéndose una correlación positiva. Por el contrario, si los puntos contempla valores muy diferentes, el numerador será una correlación negativa (Castillejo, 2007). El valor del índice está condicionado a las siguientes condiciones:

Tabla 2.2 Tipo de autocorrelación para el valor de Moran

| VALOR DEL ÍNDICE DE MORAN | SIGNIFICADO |
|---------------------------|-------------|
| | |

Elaboración Propia

El estadístico I resulta funcionalmente equivalente al coeficiente (pendiente) de una regresión lineal simple de Wy sobre Y (no de Y sobre W y, que sería la forma más natural de especificar un proceso espacial). La interpretación del test I de Moran como pendiente de una línea de regresión es lo que permite visualizar la asociación lineal entre una variable (en desviaciones a la media) y su correspondiente retardo espacial en forma de diagrama de dispersión bivalente de W y sobre Y (Chasco, 2003).

Intuitivamente el test se pregunta si los valores de la variable que presente en los vecinos se mueven en dirección similar a su propio valor (Bailey, 1994). Se toman como vecinos si comparten frontera, no obstante, se pueden obtener diferentes órdenes de contigüidad cuando se considera como vecino áreas o polígonos separados por otra área o polígono (Chasco, 2003).

Otra de los estadísticos de dependencia espacial es el C de Geary que fue diseñado por R.C. Geary (Geary, 1954). Su notación es la siguiente:

$$\frac{\sum \sum}{\sum \sum} \frac{\sum \sum}{\sum}$$

El índice c varía desde 0 hasta 3 (Legendre & Fortín, 1989), la tabla 2.3 presenta el significado de este estadístico.

Tabla 2.3 INDICE DE GEARY

| VALOR DEL ÍNDICE DE GEARY | SIGNIFICADO |
|---------------------------|---|
| | >>Hay poca diferencia entre las regiones<< >>Hay mucha diferencia entre las regiones<< |

Elaboración Propia

La interpretación de c es opuesta a la I de Moran, ya que se basa en la idea de que la varianza de la diferencia entre dos variables aleatorias dependerá de su covarianza (Dutilleul, 1998), mientras que la I de Moran es un coeficiente de autocorrelación de la muestra. Esto implica ventajas y desventajas de cada índice según las circunstancias que rodeen a los datos en el análisis, ya que I se encuentra menos afectada por el tipo de distribución de los datos que pueda soportar c (Clift & Ord, 1981). Mientras que c es menos sensible que I , ante la presencia de valores extremos (Legendre & Fortín, 1989). La elección de cualquiera dependerá de la naturaleza de los datos y de lo que se pretenda.

Y el último índice a presentar pero no menos importante es la G de *Getis y Ord (1992)*, que se define como:

$$\frac{\sum \sum}{\sum \sum}$$

Este índice según Quintana y Andrés (2014) considera vecindades cuando pares de regiones se encuentran dentro de la distancia (d). Valores significativos y positivos para el índice son evidencia de la existencia de agrupamientos de valores altos por encima de la media y viceversa cuando los valores son significativos y negativos.

Los tres índices de dependencia espacial comentados deben utilizarse de manera conjunta dado que proporcionan evidencia complementaria sobre la dependencia espacial; el I -Moran y la C son medidas de correlación de una región con las demás a las que se vincula, en tanto la G es un indicador del grado de concentración de la variable en el espacio (Moreno y Vayá, 2000). Existe el caso donde existan clústeres o patrones de agrupación espacial y de acuerdo con Quintana y Andrés (2014), los tres índices anteriores no los pueden detectar, porque son índices globales de todas las regiones. Entonces se hace uso de los estadísticos locales.

2.2.3.1 Índice de Moran Local

El I. de Moran global ofrece una medida resumen de la intensidad de la autocorrelación de los territorios considerados. Sin embargo, carece del detalle de las correlaciones entre las unidades geográficas constituyentes del territorio sean vecinas o no. Por lo que se desarrolló el I. de Moran local, que muestra una indicación del patrón de esta relación, agrupada o dispersa. Permite identificar la localización de los conglomerados espaciales, cuyas presencias fueron definidas por el I de Moran global (Céspedes *et al.*, 2014).

$$\frac{x_i - \bar{x}}{\sigma} \sum_{j \in N_i} \frac{x_j - \bar{x}}{\sigma}$$

Donde x_i es el valor de la variable correspondiente en la región i ; N_i , es el conjunto de regiones vecinas a i . Un valor elevado, positivo (negativo) y significativo del estadístico da lugar a la existencia de un clúster alrededor de la región i de valores similares elevados (bajos). El diagrama 3.1 representa cuatro tipos de conglomerados espaciales, el índice identifica unidades territoriales donde valores de análisis altos o bajos se agrupan espacialmente, así como también unidades territoriales con valores muy distintos a los de las áreas circundantes. i) alto-alto: una unidad territorial con un valor de análisis por encima del promedio, rodeada significativamente por áreas vecinas que también se encuentran sobre la media con respecto a la variable de interés.

Estas unidades territoriales corresponden a los denominados conglomerados calientes (*hot spots*); ii) alto-bajo: presencia de una unidad territorial con un valor de análisis alto, rodeada significativamente por áreas vecinas con valores que se encuentran bajo la media de la variable de interés; iii) bajo-alto: presencia de una unidad territorial con un valor de análisis bajo, rodeada significativamente por áreas vecinas con valores que se encuentran por sobre la media de la variable de interés; iv) bajo-bajo: una unidad territorial con un valor de análisis inferior al

promedio, rodeada por áreas vecinas que también se encuentran bajo la media en relación con la variable de interés. Estas unidades territoriales corresponden a los denominados conglomerados fríos (*cold spots*) (Céspedes *et al.*, 2014).

Diagrama 3.1 de dispersión de Moran



Fuente: Elaboración propia.

Índice G Local; $I = \frac{\sum \{ \frac{1}{\dots} \}}{\dots}$. Donde $s(i)$ es la desviación

estándar y \bar{y} es la media y \sum

Con base en los índices locales, es posible encontrar su contribución al índice global correspondiente y detectar sus valores extremos, lo que lo convierte en una LISA (Índice Local de Asociación Espacial), en Anselin (1995) se encuentra en la proposición básica de dicha composición. Lo relevante del LISA es que permite detectar clústeres espaciales utilizando la tipología de los cuadrantes del diagrama de dispersión de Moran⁹, que ya se visualizó con anterioridad. Es decir, en particular el uso del índice de Moran permite obtener una visión global de los procesos de interrelación espacial de las regiones, pero no ofrece la estructura regional de la autocorrelación. El análisis de la autocorrelación espacial local se efectúa a través de dos instrumentos: el gráfico de Moran y los indicadores locales de asociación espacial (Anselin, 1995).

⁹ El diagrama de dispersión de Moran es, como su nombre indica, un diagrama de dispersión en el que se representa también la línea de regresión, cuya pendiente, en este caso, será el valor del test I de Moran que, por este motivo, puede ser utilizado como indicador del grado de ajuste, así como de la presencia de valores atípicos en la nube de puntos (Chasco, 2003).

El *análisis de dependencia espacial* puede definirse como la existencia de una relación funcional entre lo que ocurre en un punto del espacio y lo que sucede en otro lugar, lo cual se explica fundamentalmente por razones de interacción humana (Anselin, 1988). En este sentido y llevado a la práctica, además del mapeo de los datos se pueden analizar las relaciones entre regiones que expresen un determinado valor y su relación con la región vecina. Lo anterior sigue siendo parte del análisis exploratorio de datos espaciales, pero ahora del lado de la autocorrelación espacial.

Empíricamente se ha comprobado que los datos económicos en su mayoría generalmente presentan algún tipo de dependencia o autocorrelación espacial como supone la Ley geográfica de Tobler (1970), que todo está relacionado con todo lo demás, pero que las cosas cercanas están más relacionadas que las cosas distantes, de esta manera los fenómenos humanos o sociales de los mapas siguientes que se registran en la región o regiones se suponen no están aislados de lo que pasa en las áreas vecinas.

Aunque en realidad para confirmar la existencia de *autocorrelación espacial* no es útil el valor de la pendiente o el estadístico de Moran que arrojaron las gráficas de dispersiónⁱⁱⁱ sino tiene que ver con la generación de una hipótesis con las siguientes particulares según Buzai (2007):

H₀ (hipótesis nula); establece que la configuración espacial de los datos se produce de forma aleatoria, y

H₁ (hipótesis alternativa); establece que la configuración espacial no se produce al azar y que existe autocorrelación espacial.

Para ello se aplica un proceso de *randomization* (generación de valores aleatorios o aleatorización) con 999 permutaciones de los datos espaciales en la figura 2.17 y 2.18. Los dos valores significativos son los de la productividad 2008 y 2013 con un *p-value* de 0.0130 y 0.040 menor al criterio del 5% o 0.05 respectivamente. Para este año se concluye la existencia de una importante autocorrelación espacial ya que su *p* valor indica que esta autocorrelación se puede producir aleatoriamente solamente en el 0.1% de los casos.

Figura 2.17 Aleatoriedad en la Autocorrelación Productividad 2008

Figura 2.18 Aleatoriedad en la Autocorrelación Productividad 2013

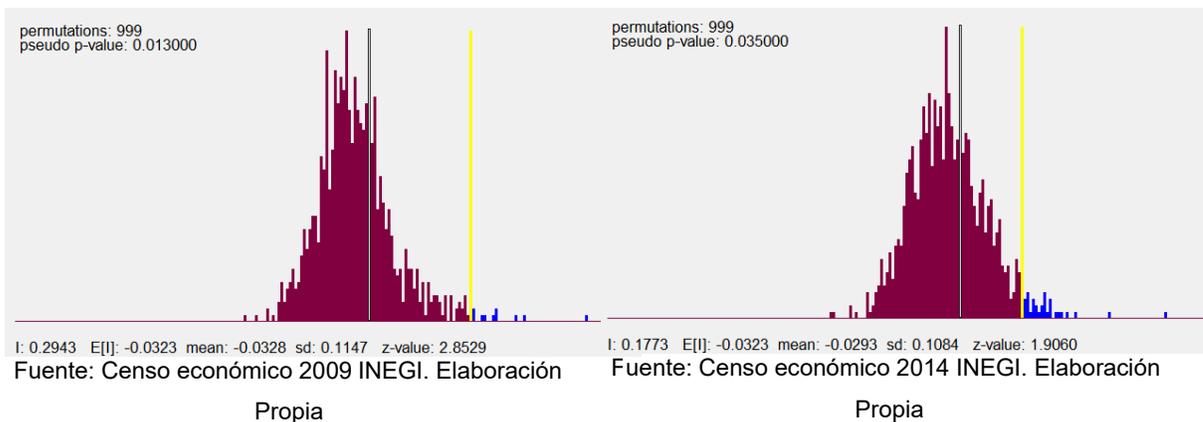
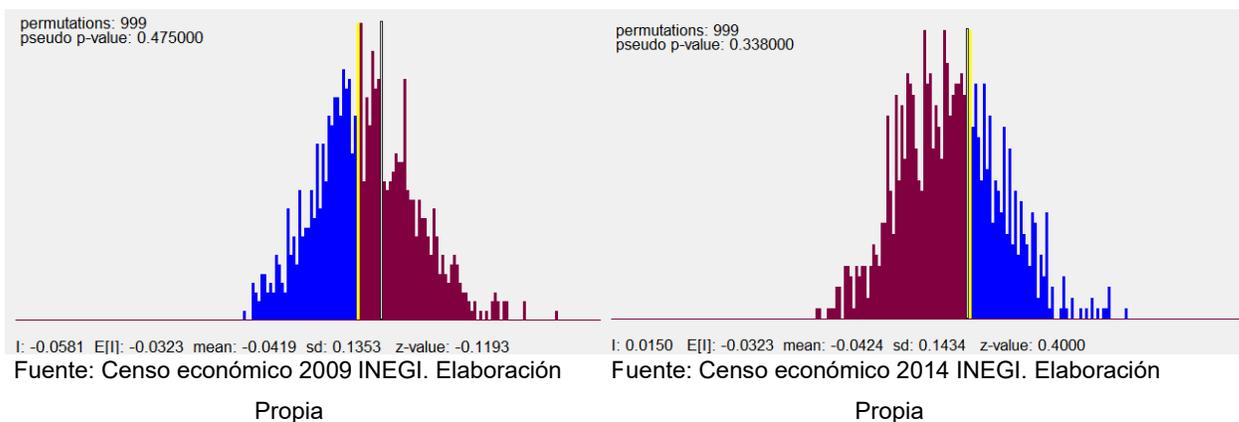


Figura 2.17.A Aleatoriedad en la Autocorrelación FMM 2008

Figura 2.18.A Aleatoriedad en la Autocorrelación FMM 2013

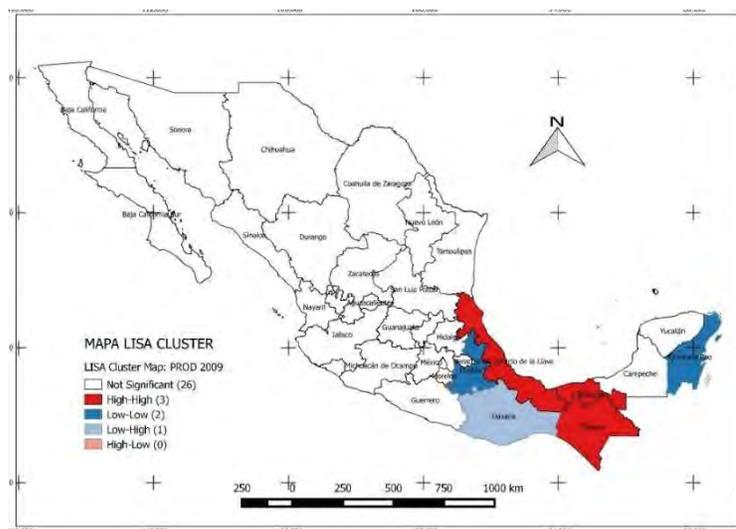


En la perspectiva que aquí se adopta, el análisis de dependencia espacial finaliza con la medición *local* y para ello se utiliza el cálculo de los Indicadores Locales de Autocorrelación Espacial (LISA) (Anselin, 1995). Son dos las principales características de LISA:

(a) Obtiene un valor de significatividad para cada clúster formado por los similares valores entre cada unidad espacial y sus vecinos, y (b) que el valor de LISA entre todas las unidades espaciales genera valores proporcionales al indicador global. Por otro lado, los *clusters* espaciales locales, también denominados hot spots, se identifican ante la Significatividad estadística obtenido por el valor de LISA mediante el test básico de H0 como la existencia de una distribución aleatoria mencionada para el cálculo de Significatividad mediante permutaciones (Buzai, 2007).

En la figura 2.19.A refleja el mapa de clusters Finalmente la cartografía que se presenta en la figura 2.19.B surge a partir del cálculo de LISA univariado, el rojo presenta las unidades espaciales con relaciones significativas de altos valores en las relaciones de contigüidad (High-High) como Veracruz, Tabasco y Chiapas, el color azul oscuro la relación inversa (*Low-Low*) como Puebla y Quintana Roo y finalmente el azul claro una relación significativa de bajo valor central y alto valor contiguo (*Low-High*) como Oaxaca para el periodo 2013.

Figura 2.19.A Cálculo de LISA – Mapa de hot-spots – Productividad 2009



Fuente: Censo económico 2009 INEGI. Elaboración Propia

Figura 2.19.B Cálculo de LISA – Mapa de hot-spots – Productividad 2013



Fuente: Censo económico 2009 INEGI. Elaboración Propia

CAPITULO 3 EVIDENCIA EMPÍRICA SOBRE EL EFECTO DEL TRANSPORTE Y EL COMERCIO EN LAS REGIONES

A lo largo de la historia de la econometría se han mostrado diferentes posturas con la intención de probar la eficiencia de la herramienta. A la par de esto se han desarrollado modelos econométricos que dan respuesta empírica a todo tipo de problemas económicos, con la intención de mostrar a los escépticos que la econometría es una ventaja de darle sustento a la teoría y no un adverso.

En los siguientes apartados de este capítulo se intenta contrastar la hipótesis de la investigación. La infraestructura está en función de la productividad y del flujo de mercancías entre regiones. De acuerdo a su ubicación geográfica es que las regiones se benefician de tal factor (infraestructura). Es importante mencionar que el modelo econométrico está basado en una recopilación de diferentes modelos, que no necesariamente usan las mismas variables de estudio, pero si tienen la misma base metodológica.

3.1 Importancia Del Análisis Econométrico

Entre los primeros estudios se encuentra Haavelmo (1944) quien señala que la econometría se dirige, esencialmente a alcanzar una conjunción entre teoría económica y medición real, utilizando como soporte la teoría y la técnica de la inferencia estadística. Aportaciones como la de Frisch (1933), establece que su primer objeto es promover estudios que se dirijan a una unificación de la aproximación teórico-cuantitativa y empírico-cuantitativa a los problemas económicos, y que constituyan reflexiones constructivas y rigurosas similares a las que han llegado a dominar las ciencias naturales. Por su parte Tintner (1968) y Stock & Watson, (2012) plantean que la econometría es el resultado de cierta perspectiva sobre el papel que desempeña la economía, que consiste en la aplicación de la estadística matemática a los datos económicos para dar soporte empírico a los modelos construidos por la economía matemática y obtener

resultados numéricos con el objeto de determinar vínculos presentes entre variables, siempre basado en un modelo conceptual de teoría económica.

Frisch (1933), Wooldrige (2008) y Christ (1983) agregan que la econometría está destinada a estimar las relaciones económicas, contrastar teorías, evaluar y poner en práctica políticas gubernamentales y de negocio. A darle contenido empírico a las relaciones económicas. Es una disciplina distinta a la estadística matemática ya que se centra en los problemas inherentes a la recopilación y al análisis de datos económicos no experimentales¹⁰ tal como menciona Haavelmo (1944), el método de la investigación econométrica pretende, esencialmente, en una conjunción de la teoría económica y las medidas reales, utilizando la teoría y la técnica de la inferencia estadística como un pilar de un puente. Por su parte, quien aplica la ciencia es el econométra que parte de una teoría, proposición racional y abstracta, que no discute en cuanto tal, sino que se limita a medirla (Álvarez, 2003), por su parte Malinvaud (1966) considera que el arte del econometrista consiste en encontrar un conjunto de supuestos lo bastante específicos y realistas para que le permitan aprovechar de la mejor manera los datos con que cuenta. Un econometrista tiene que ser un matemático y estadístico competente. Los conocimientos fundamentales de las matemáticas, estadísticas y la teoría económica son un requisito previo necesario para este campo (Baltagi, 2011).

Para Portillo (2006) una de las definiciones más completas es la descrita por Hill, *et al.*, (1988) apuntando que la econometría utiliza la teoría económica, economía matemática e inferencia estadística como fundamentos analíticos y los datos como fuente de información, proporciona a la ciencia económica una base para: 1) modificar, refinar o posiblemente refutar las conclusiones contenidas en el cuerpo de conocimientos, conocido como teoría económica; y 2) conseguir signos, magnitudes y proposiciones fiables acerca de los coeficientes de las variables en las relaciones económicas, de modo que esta información pueda servir de base

¹⁰ A diferencia de los datos experimentales que se recopilan en laboratorios los datos experimentales no se recogen mediante experimentos controlados con individuo, empresas o segmentos de la economía. También son denominados a veces *datos de observación* para enfatizar el hecho de que el investigador recopila datos de forma pasiva (Wooldrige, 2008).

para la toma de decisiones y la elección. Aunque las definiciones se encuentran en años y contextos diferentes, todo apunta hacia el mismo camino considerando en todo momento la teoría de esta disciplina pero también el método con el que se fundamenta.

Gujarati (2009) argumenta que la econometría merece un estudio separado por lo siguiente: a) La teoría económica hacen afirmaciones o hipótesis de naturaleza cualitativa. Pero la teoría por sí sola no proporciona medida numérica alguna, no dice cuánto aumentará o se reducirá la cantidad como resultado de un cambio determinado. La econometría es la que da contenido empírico a gran parte de la teoría económica, b) El interés principal de la economía matemática es expresar la teoría económica en una forma matemática (ecuaciones) sin preocuparse por la capacidad de medición o de verificación empírica de la teoría. Por su lado, la econometría se interesa en la verificación empírica de la teoría económica, y las ecuaciones econométricas requieren una gran dosis de ingenio y destreza, y c) La estadística económica se relaciona en primer lugar con la recopilación, procesamiento y presentación de cifras económicas en forma de gráficos y tablas. Los datos así reunidos constituyen la materia prima del trabajo econométrico.

Pero el estadístico económico no va más allá de la recolección de información, pues no le conciernen las cifras recopiladas para probar las teorías económicas. Sin duda, el econométrico a menudo necesita métodos especiales por la naturaleza única de la mayoría de las cifras económicas, pues no se generan como resultado de un experimento controlado.

No todos los teóricos consideran a la econometría como una ventaja o como una forma de enfoque. Tal como enfatiza Núñez (2005) cuando trae la crítica de Lucas (1978), quien considera que la econometría se encontraba inmersa en una crisis estructural debido a su incapacidad para incorporar en sus modelos tanto los choques económicos aleatorios como los ajustes de los agentes económicos privados ante los cambios de política económica. Sin embargo, lo más notable de su crítica a la econometría tradicional fue que el enfoque econométrico tradicional

no reaccionó sino apenas un poco, indicando que, en efecto, la econometría había recibido críticas incluyendo la suya, pero que también ésta se encontraba en condiciones de superar los problemas indicados por los escépticos.

Mientras que algunos econométricos como Pesaran (1987), McAleer (1983) y Volker *et al.* (1985) argumentan que no hay duda de que la econometría está sujeta a limitaciones importantes, que se derivan en gran parte de lo incompleto de la teoría económica y la naturaleza no experimental de los datos económicos. Sin embargo, estas limitaciones no deben distraernos para reconocer el papel fundamental que la econometría ha llegado a desempeñar en el desarrollo de la economía como disciplina científica.

Puede que no sea posible de manera concluyente comprobar si se cumple o no la teoría regional mediante métodos econométricos para determinada región, pero eso no quiere decir que no hay nada útil, se puede aprender de los intentos de probar formulaciones particulares de una teoría dada con posibles alternativas.

Del mismo modo, el hecho de que el modelado econométrico esté inevitablemente sujeto al problema de las búsquedas de especificación, no significa que toda la actividad no tiene sentido. Los modelos econométricos son herramientas importantes para el pronóstico y análisis de políticas, y es poco probable que vayan a ser desechados en el futuro. El reto consiste en reconocer sus limitaciones y trabajar para convertirlos en herramientas más fiables y eficaces ya que no existe otra alternativa fiable.

3.1.1 Modelo Lineal Clásico

Para considerar las relaciones inexactas entre las variables del mundo económico surgen los modelos econométricos. Éstos, además de relacionar una variable dependiente con otras independientes o explicativas (relación de comportamiento), introducen una componente aleatorio o término de error. Ésta tiene un comportamiento estocástico y representa factores determinantes del comportamiento de la variable endógena que los modelos no pueden recoger de

forma explícita. Así, el comportamiento de la variable (y) viene explicado por un modelo o relación en la que se puede distinguir una parte determinista (integrada por las variables explicativas) y una parte aleatoria (Sancho, *et al.*, 2001)

Un modelo lineal trata de explicar el comportamiento de una variable aleatoria mediante su relación lineal con los valores de otras variables que podrían influirla. Es decir, estamos asumiendo la hipótesis de que los datos siguen un patrón lineal implícito teniendo en cuenta que el ajuste no es perfecto y contiene errores. Siguiendo la notación de Baltagi (2011) para la relación lineal:

[3.1.1]

Donde indica la observación i -ésima de la variable dependiente que es la productividad del sector manufacturero (PROD) o el flujo de mercancías del sector manufacturero (FMM) dentro de las regiones. Por su parte son los flujos de bienes que abastecen los establecimientos manufactureros al mercado interno (FBI), flujos de bienes que abastecen los establecimientos manufactureros al mercado externo (FBE), longitudes de líneas carreteras (LC) (incluye redes federales, estatales, caminos rurales y brechas mejoradas) y longitudes ferroviarias (LF). α y β son el origen y la pendiente de esta simple relación lineal entre Y y X . Por lo tanto, el modelo se representa de las siguientes dos formas;

[3.1.2]

[3.1.3]

Las observaciones utilizadas son recompiladas del periodo 2008 y 2013 de las 32 entidades federativas. Al agregar el término de error la ecuación 3.1.2 y 3.1.3

establece que no todas las observaciones se encuentran dentro de la relación lineal, Baltagi (2011) menciona que el error aleatorio puede ser causado por (i) la omisión de los factores relevantes que podrían influir en la productividad manufacturera o en caso del flujo de mercancías en las regiones, (ii) el error de medición, (iii) equivocada elección de una relación lineal entre las variables dependientes e independientes, cuando la verdadera relación puede ser no lineal.

Lo anterior nos lleva al método de mínimos cuadrados ordinarios. La paternidad de este método se reparte entre Legendre (1805) y Gauss (1809) quienes exponen que cuanto menores son los residuos, mejor es el ajuste.

Sancho *et al.* (2001) y Torrelavega (2005) coinciden en que el método más efectivo para determinar los parámetros y se conoce como técnica de mínimos cuadrados ordinarios, lo plantean como la búsqueda del valor de los parámetros de la recta que mejor se ajusta a la nube de puntos que forma cada par de observaciones (y_i , x_i). Es decir, localiza los parámetros que mejor se ajustan a los datos experimentales.

El método de mínimos cuadrados asume que, al fijar las condiciones experimentales, los valores (PROD o FMM) de la variable independiente se conocen con precisión absoluta. Aunque, las mediciones de la variable x , irán afectadas de sus errores correspondientes (Torrelavega 2005).

3.1.2 Lag y Error

En el análisis exploratorio de los datos en el capítulo 2, las I de Moran para la variable de productividad fueron positivas y altamente significativas en los dos años, 2008 y 2013, muestra la presencia de autocorrelación espacial. Caso contrario para la variable flujo de mercancías manufactureras (FMM) que en el 2008 presenta un *pseudo* valor de 0.4750 y en el 2013 uno de 0.3380. Se tiene nula dependencia espacial para la variable en los dos años.

Tabla 3.1.2; Índices de autocorrelación espacial de la productividad del sector manufacturero de las regiones mexicanas 2008 y 2013

| VARIABLE | I DE MORAN (I-M) | P VALOR |
|---------------------------------|-----------------------------|----------------|
| PRODUCTIVIDAD 2008 | (0.2943) ^{***} | 0.0130 |
| PRODUCTIVIDAD 2013 | (0.1773) ^{**} | 0.0440 |
| FLUJO DE MERCANCÍAS 2008 | - | 0.4750 |
| FLUJO DE MERCANCÍAS 2013 | - | 0.3380 |

Fuente: INEGI Censos económicos 2009 y 2014.¹¹
Elaboración propia con ayuda del software GEODA

Sin embargo, pasa lo que plantea Sánchez (2008), la I de Moran provee solo una mirada limitada de este fenómeno, ya que los clústeres de valores similares (bajo-bajo, alto-alto) podrían reflejar solo la distribución geográfica de las variables explicativas de la productividad manufacturera.

Es posible que existan otros fenómenos que contribuyen a la formación de clústeres más allá de las que muestra la variable endógena de una región. Como ya se ha mencionado, la dependencia espacial es una síntesis de procesos de todo tipo (económicos, sociales, etc.) que aportan a la dependencia entre las unidades analizadas. Sánchez (2008) recalca que con datos de sección cruzada no es posible dar cuenta de los procesos que llevan a la dependencia espacial entre unidades geográficas, pero es posible identificar la existencia de dichas interacciones y que si las medidas exploratorias revelan la presencia de dependencia espacial, entonces se deben implementar dichos modelos, porque ignorarla podría comprometer la fiabilidad del análisis (Anselin, 1992).

La autocorrelación en el término del error puede ser manejada a través de un Modelo Espacial del Error que supone que la dependencia espacial encontrada en

¹¹Nivel de significación: Rechazo de la hipótesis nula de ausencia de dependencia espacial con un nivel de significación (***): $\alpha=0.01$; (**): $\alpha=0.05$ y (*): $\alpha=0.10$.

la variable dependiente es resultado de la distribución geográfica de nuestras variables explicativas y de la autocorrelación del término del error (Anselin, 2002; Baller et al., 2001). Siguiendo la metodología que analiza Sánchez (2008), el modelo es similar a un modelo de regresión estándar, pero se estima el parámetro λ que calcula el grado de autocorrelación de los errores, dada una matriz de pesos geográficos (W):

[3.1.4]

Por otro lado, la autocorrelación espacial también se presenta cuando el valor de la dependiente en cada unidad geográfica está determinado por el valor que dicha variable ocupa en las unidades vecinas. La alternativa para modelar este tipo de autocorrelación es el Modelo Espacial Lag, que considera la dependencia espacial introduciendo una variable espacial lag Sánchez (2008).

De acuerdo con Baller *et al.* (2001), un Modelo Espacial Lag representa la relación interactiva entre la variable dependiente y las variables independientes en las unidades vecinas.

La ecuación que retoma Sánchez (2008) permite entender cómo se estima un modelo de rezago espacial, donde p es el parámetro espacial de rezago que se estima, W es la matriz de pesos geográficos, X es la matriz de covarianzas, y ε es el vector de errores no correlacionados:

[3.1.5]

GeoDa ofrece una vía para determinar qué modelo necesitamos implementar. Este software permite estimar una regresión estándar de mínimos cuadrados y una serie de diagnósticos espaciales, con los cuales podemos decidir qué tipo de autocorrelación espacial está presente en nuestros datos y, por tanto, qué modelo utilizar siguiendo la regla de decisión de Anselin y Rey (1991).

Tabla 3.1 “Resultados del Modelo, Econometría Clásica y Espacial”

| ECONOMETRIA CLÁSICA/ESPACIAL | | | | |
|--|-------------------------|---------------------------------|-------------------------|-------------------------------|
| I.- Primera Parte - Estimación Regresión clásica | | | | |
| Y | 2009 | | 2014 | |
| | PRODUCTIVIDAD (PROD) | FLUJO COMERCIAL (FMM) | PRODUCTIVIDAD (PROD) | FLUJO COMERCIAL(FMM) |
| FBI | 14.77 (0.25) | 0.13 (1.54) | 12.2 (1.32) | 0.38 (3.56)** |
| FBE | -14.37 (0.25) | 0.82 (9.62)*** | -11.9 (-1.30) | 0.58 (5.43)** |
| LF | 1.37 (1.27) | -0.01 (-2.58)** | 1.02 (1.47) | -0.02 (3.67)** |
| LC | -2.99 (-0.52) | 0.11 (2.78)*** | -6.72 (-1.24) | 0.23 (3.67)** |
| Jarque-Bera | 94.099 (0.000) | 0.2808 (0.86)/ S=0.01, K=3.4 | 161.62 (0.00) | 3.3 (0.19) S=-0.58, K=4.04 |
| D-Watson | 2.10 | 2.33 | 1.79 | 1.86 |
| Breusch- Pagan | 0.49 | 0.22 | 0.61 | 0.06 |
| II.- Segunda parte - Diagnóstico para la dependencia espacial | | | | |
| TEST | PRODUCTIVIDAD (PROD) | FLUJO COMERCIAL (FMM) | PRODUCTIVIDAD (PROD) | FLUJO COMERCIAL(FMM) |
| I de Moran | 0.0123 | 0.0069 | 0.0909 | 0.2260 |
| LAG | 0.0388 | 0.9244 | 0.2165 | 0.6128 |
| LAG Robusto | 0.5233 | 0.9689 | 0.6513 | 0.5884 |
| ERROR | 0.0488 | 0.0051 | 0.2438 | 0.1296 |
| ERR-Robusto | 0.8837 | 0.0051 | 0.8518 | 0.1266 |
| SARMA | 0.1175 | 0.0198 | 0.4579 | 0.2739 |
| MODELO LAG | 0.3687 (2.2067)** | - | - | - |
| MODELO ERR | - | -0.7999 (-7.0144)*** | - | - |

Nivel de significación: Rechazo de la hipótesis nula de ausencia de dependencia espacial con un nivel de significación (***): $\alpha=0.01$; (**): $\alpha=0.05$ y (*): $\alpha=0.10$.

Elaboración propia en Eviews 6 y GeoDa.

La información de la tabla 3.1 es tomada en logaritmos para los dos periodos 2009 y 2014. En la primera parte de la tabla (Estimación Mínimos Cuadrados Ordinarios) se representan los coeficientes, en paréntesis los valores *t-Statistic* y los asteriscos muestran el nivel de significancia: 0 „***“, 0.001 „**“, 0.01 „*“, 0.05 „.“.

Después de la estimación de los parámetros conviene efectuar un conjunto de pruebas de validación de los resultados. Un modelo de regresión lineal mínimos cuadrados ordinarios (MCO) tiene que cumplir con los supuestos sobre la perturbación aleatoria o errores. Murillo y González (2000) discuten que las hipótesis relativas de las características de la distribución de probabilidad de las variables aleatorias que constituyen los términos de perturbación aleatoria del modelo de regresión pueden ser revisadas mediante dos enfoques distintos. Por una parte, con el recurso de ilustraciones gráficas del comportamiento de los residuos o errores de la estimación MCO, y por otra, con pruebas estadísticas específicamente diseñadas para estos objetivos.

La Jarque-Bera dice Baltagi (2011) una prueba asintótica simple para la hipótesis de normalidad (normalidad de los disturbios $\sim (\mu, \sigma^2)$ si se viola se acepta la presencia de *outliers*) se basa en el hecho de que la distribución normal tiene una medida de asimetría de cero y una curtosis de 3 (también llamados momentos de tercer y cuarto orden) mientras que su hipótesis nula (H_0) es normal y la hipótesis alternativa (H_a) los datos no se distribuyen como una normal el criterio para aceptar la H_0 tiene que tender o ser mayor a 3.

Otra prueba elemental es la de ausencia de correlación serial en el término de perturbación medida por la prueba Durbin-Watson, la H_0 plantea que no hay autocorrelación y la H_a que existe autocorrelación, el criterio para rechazar la hipótesis nula según Murillo y González (2000) es cuando $0 < d < 1$, si $0 < d < 2 - 4d_u < d < 4$, si $2 < d < 4$, el estadístico d toma valores en el rango comprendido entre 0 y 4.

Por el momento bastará con interpretar este contraste como una prueba de especificación errónea. La prueba Breusch-Pagan, la hipótesis nula considera la homocedasticidad ($\text{Var}(\epsilon) = \sigma^2 = \text{constante}$) mientras que la hipótesis alternativa marca presencia de heteroscedasticidad, la H_0 se rechaza cuando el p valor es mayor a 0.05.

Finalmente, la primera prueba que consiste en una regresión simple tomando únicamente el periodo 2009 se le aplica el método de mínimos cuadrados ordinarios. Posteriormente se plantea la hipótesis nula; la productividad (PROD) se explica ante cambios en la infraestructura y el comercio en el año 2009 para las entidades y sus vecinos. Con los resultados de la primera parte de la tabla (regresión clásica) se rechaza la hipótesis y se concluye que existe heterogeneidad en el comportamiento de los datos de cada estado. Por lo tanto, se debe seguir el análisis rumbo a la evidencia empírica espacial.

La segunda parte de la tabla (Diagnóstico para la dependencia espacial) muestra datos de una regresión espacial estimada con el paquete GeoDa, recordemos que si el análisis es espacial se considera una matriz de pesos espaciales. Sin embargo, en este caso se lleva a cabo con una matriz de distancias euclidianas. El valor que interesa es el estadístico índice de Moran (I-M) cuando la H_0 especifica que hay ausencia de dependencia espacial y la H_a que existencia dependencia espacial que con un criterio de rechazo menor al 5% (0.05).

Las pruebas de rezago y error espacial se anexan a esta última parte; por un lado, el diagnóstico de dependencia espacial LAG se presenta como $y = \rho W y + U$ la $H_0: \rho = 0$ y $H_a: \rho \neq 0$ donde ρ es rho, W es la matriz espacial y U los residuos.

Por otro lado, el segundo diagnóstico de dependencia espacial es ERROR y se modela como $U = \rho U + \epsilon$ con hipótesis nula $\rho = 0$ e hipótesis alternativa $\rho \neq 0$. La suma de estos dos diagnósticos se convierte en un SARMA (LAG y ERROR); es decir, una prueba simultánea de los dos modelos donde la $H_0: \rho = 0$ y la $H_a: \rho \neq 0$. Es muy importante que para la discriminación de métodos se tenga que seguir la *regla de decisión* de Anselin y Rey (1991).

3.2 Datos De Panel

Tal y como mencionan Labra y Torrecillas (2014), tradicionalmente los métodos estadísticos que intentan explicar un fenómeno observado a través de una serie de variables han sido tratados mediante regresiones lineales, usando el método de MCO. Si una muestra presenta efectos individuales y las variaciones en el tiempo también explican el comportamiento de las variables, entonces debemos aplicar la metodología de datos de panel. Por el contrario, de observarse este tipo de condición donde no existen efectos individuales un análisis utilizando MCO sería consistente y el más eficiente.

Los datos de panel ofrecen una mayor riqueza de información en relación a los otros dos tipos de datos, las series de tiempo y corte transversal. Elhorst (2003) indica que los datos de panel son generalmente más informativos y contienen más variación y menos colinealidad entre las variables. El uso de los resultados de datos de panel es una mayor disponibilidad de grados de libertad que aumenta la eficiencia en la estimación. Al igual que para Wooldrige (2010), una ventaja importante de datos de panel es que se incrementa la precisión en la estimación. Esto es el resultado de un aumento en el número de observaciones debido a la combinación de la agrupación o varios períodos de tiempo de datos para cada individuo. Elhorst (2003) proporciona un análisis de las cuestiones que se plantean en la estimación de modelos de datos de panel. Los comúnmente utilizados son cinco: efectos comunes, efectos fijos, efectos aleatorios, coeficientes fijos, y los modelos de coeficientes aleatorios.

Los datos de panel proporcionan información sobre el comportamiento individual, tanto en el tiempo y en todos los individuos. Incluso en el caso de regresión lineal, el análisis de datos de panel estándar utiliza una gama mucho más amplia de modelos y estimadores que es el caso con datos de corte transversal (Wooldrige, 2010).

Cobacho y Bosch (2005) citan a Baltagi (2001) y enumeran algunas de las ventajas e inconvenientes del uso de los datos de panel. Entre las ventajas están el control sobre la heterogeneidad individual; más variabilidad, menos colinealidad entre las variables, más grados de libertad y mayor eficiencia; mejor adecuación al estudio de las dinámicas de ajuste; mejor capacidad de identificar y medir efectos que no son detectables en datos puros de sección cruzada o de series temporales y también mejor capacidad de análisis en comportamientos más complicados.

En resumen, según Mahía (2000), la utilización de datos de panel en lugar de series temporales se justifica por aprovechar la variabilidad transversal. La identificación y estimación de los parámetros de una función de respuesta explota la variación de las variables incluidas. Si las variables no presentan excesiva variabilidad temporal pero sí transversal, la aproximación con datos de panel aportaría capacidad extra para esa estimación.

Siguiendo la metodología de Wooldrige (2010) un modelo de datos de panel podría formularse en términos plenamente genéricos como:

[3.1.2]

Donde y_{it} es una variable dependiente del individuo i en el tiempo t , α constante del modelo, β_1 coeficiente 1 de la variable x_{it} , β_2 es un vector $k \times 1$ de la variable ingresos de las empresas manufactureras que comercian sus productos al mercado interno en el tiempo t , β_3 coeficiente 2 de la variable z_{it} , ingresos de las empresas manufactureras que comercian sus productos al mercado externo en el tiempo t , β_4 coeficiente 3 de la variable l_{it} , longitudes de líneas carreteras (incluye redes federales, estatales, caminos rurales y brechas mejoradas)

en el tiempo t , α_i coeficiente 4 de la variable i , β_i longitudes ferroviarias en el tiempo t . ϵ_{it} es un término de perturbación escalar, i índices individuales (de las entidades federativas) en una sección transversal, y t índices de tiempo.

Este modelo aún es demasiado general y no es estimable, ya que hay más parámetros a estimar que las observaciones. Más restricciones necesitan ser colocadas en la medida en que α_i y β_i varían con i y t , y en el comportamiento del error (Wooldrige, 2010). A partir de este modelo general, y con base en ciertos supuestos y restricciones acerca del valor de algunos de los parámetros, se pueden derivar algunas otras variantes de modelos de datos de panel (Labra y Torrecillas, 2014). Los efectos individuales (α_i) pueden ser tratados como aleatorios o fijos, estos últimos son en los que se apoyará el análisis de datos de panel para México con sus 32 entidades federativas. Para poder llevar a cabo esta estimación, se asume que los β_i son constantes a lo largo del tiempo.

Aunque uno de los modelos clásicos más utilizados en la técnica de datos de panel es el de efectos comunes o también llamado modelo pool porque se siguen manteniendo los supuestos de homocedasticidad, normalidad y no autocorrelación serial, no será el método de análisis para la investigación ya que éste se rige bajo el supuesto que los estados o las entidades son homogéneos. Dado que la heterogeneidad no se puede detectar ni con estudio de series temporales ni con los de corte transversal, la técnica de panel permite realizar un análisis más dinámico al incorporar a la dimensión temporal de los datos, lo que enriquece el estudio, en especial en periodos de grandes cambios (Mayorga y Muñoz, 2000).

Wooldrige (2010) menciona que otra ventaja de usar panel es realizar estimaciones con efectos fijos ya que permite visualizar la heterogeneidad de los datos que podrían estar correlacionados con los regresores. Por lo tanto, uno de los principales objetivos en la investigación es capturar la heterogeneidad no observable entre las entidades federativas, para eso cada estado deberá de tener diferentes impactos. De entrada, se sabe que los estados contribuyen de una manera heterogénea a una variable y que tienen grandes cambios en el tiempo.

3.2.1 Modelo de Efectos Fijos: Within

El modelo de efectos fijos considera que existe un término constante diferente para cada individuo y supone que los efectos individuales son independientes entre sí. Con este modelo se considera que las variables explicativas afectan por igual a las unidades de corte transversal y que éstas se diferencian por características propias de cada una de ellas, medidas por medio del intercepto.

Es por ello que los N interceptos se asocian con variables *dummy* con coeficientes específicos para cada unidad, los cuales se deben estimar para la i -ésima unidad de corte transversal, además la heterogeneidad no observable se incorpora en la ordenada al origen del modelo Mayorga y Muñoz (2000). Siguiendo la notación de Mayorga y Muñoz (2000) la relación es la siguiente:

$$\mu \quad [3.2.1]$$

Donde i se refiere al individuo o a la unidad de estudio (corte transversal), t a la dimensión en el tiempo, μ_i es un vector de interceptos de n parámetros, X_{it} es un vector de k parámetros y y_{it} es la i -ésima observación al momento t para las k variables explicativas. En este caso, la muestra total de las observaciones en el modelo vendría dado por $N \times T$. Como es una extensión de los datos de panel la especificación del modelo es casi la misma. Debe hacerse notar que en este modelo se presenta una pérdida importante de grados de libertad. Estos autores mencionan que mientras el interés sea limitado a una muestra que se ha seleccionado a conveniencia o bien que se esté trabajando con la población, la estimación de efectos fijos será la correcta. En la investigación el modelo se ve como un caso de efectos fijos, respecto a los efectos que se encuentran en la muestra.

Antes es importante identificar que el tipo de panel que se utiliza es estático. Es decir, el modelo de regresión de un solo factor supone que el error aleatorio se descompone en $\alpha_i + \epsilon_{it}$ donde cada α_i es el efecto individual (inobservado) de cada unidad de sección cruzada, invariante en el tiempo. La presencia del efecto fijo en esta ecuación hace que una estimación de β por MCO no sea consistente. Los métodos que se utilizan para solventar ese problema son la estimación de efectos fijos o *within* (y por estimación de efectos aleatorios). La estimación *within* puede llevarse a cabo transformando el modelo (3.1.2) en otro en términos de las medias de grupo (Wooldrige, 2010). Se trabaja con un panel corto que mide la relación entre las desviaciones individuales específicas de los valores de sus regresores promediados en el tiempo, esto se hace usando la variación de los datos en el tiempo.

En concreto, el modelo *within* de efectos fijos asume que cada variable explicativa tiene un solo coeficiente; es decir tiene el mismo impacto sobre la variable dependiente, pero en donde cada individuo tiene distinta constante. Para tratar los efectos fijos se emplea el estimador intragrupos (*within*), el cual asume que el efecto individual está correlacionado con las variables explicativas. Este supuesto relaja la condición impuesta por el estimador de efectos aleatorios, tratando el efecto individual separadamente del término de error (ϵ_{it}), este estimador tiene la ventaja de que permite conocer los separadamente, lo que contribuye a entender de mejor forma el modelo. Además, evita una sobrestimación del parámetro β , lo que ocurre cuando se aplica el estimador de efectos aleatorios (Labra y Torrecillas, 2014).

Finalmente cuando se tiene que elegir cuál es la estimación más adecuada a nuestra situación de estudio con datos de panel, los contrastes de hipótesis suponen una herramienta útil que permite dar soporte o rechazar determinados supuestos. La doble dimensionalidad de los conjuntos de datos de panel permite en algunos casos adaptar los contrastes de hipótesis ya planteados para alguna de sus dos dimensiones, y en otros, ha requerido el planteamiento de *tests* propios

de los datos de panel (Cobacho y Bosch, 2005). En este caso, se contrasta¹² el modelo within vs pool con el test f para efectos individuales ($pFtest$) para concluir con un choque entre efectos fijos y efectos aleatorios con base en la prueba de Hausman. Montero (2005) aborda ampliamente la prueba de Hausman y la define como un test *chi-cuadrado* que determina si las diferencias son sistemáticas y significativas entre dos estimaciones, también describe la hipótesis, si es que p valor < 0.05 se rechaza la hipótesis nula de igualdad al 95% de confianza y se deben asumir las estimaciones de efectos fijos.

Por el mismo criterio, si p valor < 0.05 se rechaza la hipótesis nula de igualdad al 95% de confianza y se debe rechazar la hipótesis de independencia o irrelevancia de las variables. Por el contrario, si p-valor > 0.05 se debe admitir la hipótesis nula de igualdad de estimaciones y, entonces el estimador más eficiente, el del efectos variables debe ser seleccionado. Igualmente, si el p-valor > 0.05 debe asumirse con el 95% de confianza, que la variable introducida en el modelo de contraste no es irrelevante. En específico, Mayorga y Muñoz (2000) definen a esta prueba como la que permite determinar qué modelo es el más adecuado para el panel de datos que se está analizando, si el de efectos fijos o de efectos aleatorios. Utiliza para ello una prueba chi-cuadrado con la hipótesis nula de que el modelo de efectos aleatorios es el que mejor explica la relación de la variable dependiente con las explicativas, y por tanto se tiene la hipótesis alternativa de que el mejor método que se ajusta es el de efectos fijos.

3.3 La Matriz De Distancias y El Modelo Durbin

Según Viton (2010), la base para la mayoría de los modelos es un indicador de si una región es un vecino espacial de otro; o lo que es equivalente, que las regiones son vecinos de una región determinada. Se trata de una matriz simétrica $R \times R$ (i, j) con elemento igual a 1 si las regiones i y j son vecinos de los otros (o, más generalmente, son espacialmente relacionada), y cero en caso contrario.

¹² RStudio es un entorno de desarrollo integrado (IDE) para R (lenguaje de programación), <https://www.rstudio.com/products/rstudio/>.

Como señala LeSage (1998), hay un gran número de formas de construir una matriz. Ésta podría ser por contigüidad reina, torre, alfil o por distancias, de primer o segundo orden, etc. El enfoque que da Viton (2010) y que está basado en la investigación es el que se basa en la distancia, considera dos regiones como vecinos si, por ejemplo, sus centroides ponderados por población son dentro de una cierta distancia determinada uno de otro. Esto, obviamente se puede ampliar de varias maneras: se podría utilizar diferentes distancias, o diferentes pesos para el cálculo de los centroides. Es claro que no se debe proceder tan mecánicamente como aquí específica, sino solo considerar los problemas prácticos que se están estudiando.

En el modelo Durbin, según Viton (2010) sólo se suma los valores promedio al vecino de las variables independientes a la especificación. Ejemplo: el nivel de flujos comerciales en la región j depende de la intensidad de la infraestructura carretera en la región j , así como también de la intensidad en las regiones vecinas. Aparte de los problemas potenciales de multicolinealidad (recordemos que por filas, X y W X son para diferentes regiones debido a que los elementos diagonales de W son cero), este modelo solo incluye rezagos de las variables explicativas.

Un modelo Durbin es representado de la siguiente manera:

$$[3.3]$$

Donde al igual que en panel, las primeras cuatro variables exógenas son un vector en el tiempo, pero los cuatro últimas exógenas () se multiplican la matriz de distancias (W). El modelo Durbin es sumamente efectivo al agregar rezagos a las variables independientes.

Sin embargo, este modelo se clasifica previamente en el panel dinámico. Arellano y Bond (1991) construyen un estimador basado en el Método Generalizado de los Momentos (GMM), que utiliza variables instrumentales basadas en retardos y diferencias de todas las variables del modelo y que está especialmente propuesto para paneles con muchos individuos y pocos periodos. Las posibles variables instrumentales y sus retardos las obtienen del método desarrollado por Hansen (1982).

3.4 Evidencia Empírica Del Modelo

Tabla 3.4 “Resultados del Modelo Panel con Modelo Durbin”

| DATOS DE PANEL Y DURBIN | | | | | | | |
|-------------------------|---------------|-----------------|----------------|---------------|-----------------|---------------|----------------|
| PANEL (WITHIN) | | | | DURBIN | | | |
| | PRODUCTIVIDAD | FLUJO COMERCIAL | | PRODUCTIVIDAD | FLUJO COMERCIAL | | |
| FBI | -2.23 (-0.29) | 0.05(0.45) | | 5.48 (0.61) | 0.46 (3.45)** | | |
| FBE | 2.23 (0.30) | 0.94 (7.77)*** | | -5.35 (-0.61) | 0.51 (0.13)*** | | |
| LC | -0.71(-0.82) | -0.003(-0.26) | | -0.30(-0.30) | -0.04(-3.1)** | | |
| LF | -7.7 (-1.04) | 0.09(0.79) | | -4.01(-0.47) | 0.13(0.81) | | |
| WFBI | - | - | | 6.17(0.41) | -0.12(0.50) | | |
| WFBE | - | - | | -7.56(0.60) | 0.11(0.52) | | |
| WLC | - | - | | -3.23(0.17) | 0.02(1.06) | | |
| WLF | -- | - | | -2.80(0.87) | 1.34(3.73)** | | |
| - | | pFtest | Hausman | PFtest | Hausman | pFtest | Hausman |
| - | | 8.871e- | | 6.743e | | | |
| - | | 05 | 3.148e-06 | -08 | 0.7818 | 1.67e06 | 0.00001 |

Elaboración propia en RStudio.

La tabla 3.1 (Primera Parte-Estimación Mínimos Cuadrados Ordinarios) se refiere al caso en que no existe heterogeneidad no observable ya que las variables se modelaron en corte transversal y no son significativas en todos los casos.

Por lo tanto, debe buscarse una especificación que capture la heterogeneidad no observable ya sea a través del tiempo, entre unidades de estudio (individuos) o en ambos sentidos en forma apropiada, con el fin de evitar el problema de sesgo sobre los estimadores de los parámetros de las variables explicativas que se cometería si se emplea el caso de mínimos cuadrados ordinarios.

Existen dos procedimientos adicionales para estimar el modelo en un sistema de datos de panel: uno de ellos implica el reconocimiento de que las variables omitidas pueden generar cambios en los interceptos ya sea a través del tiempo o entre unidades de corte transversal, en este caso el modelo de efectos fijos trata de aproximar estos cambios con variables dummy Mayorga y Muñoz (2000). El estimador *within* como ya se mencionó es un estimador del modelo de efectos fijos que, a diferencia del de *efectos comunes*, utiliza las características espaciales de los datos de panel.

La tabla 3.4 de los datos de panel (aun sin Durbin) muestra los resultados. La información corresponde a dos periodos (2008 y 2013) para 32 entidades federativas. El objetivo es analizar a la productividad del sector manufacturero (PROD) o el flujo de las mercancías manufactureras (FMM) en función de la infraestructura (LC y LF) y el comercio (mercancías creadas por establecimientos manufactureros para el mercado interno y externo). Los resultados asumen que cada variable explicativa tiene un solo coeficiente para todas las entidades y un intercepto diferente para cada una. Como resultado para la variable endógena que nos ocupa (FMM), el modelo estima una relación positiva entre el comercio interregional y el comercio con el exterior, dicha relación es de 0.05 y 0.94 respectivamente.

Lo anterior implica que ante un aumento en una unidad en la productividad manufacturera produce un incremento del 0.05 y del 0.94 en el comercio interno y externo. De la misma manera es interesante el resultado de una relación negativa entre la variable endógena y la infraestructura carretera (LC) de 0.003.

Dentro de la misma tabla 3.4 en la cuarta y quinta columna se muestran los datos obtenidos para el modelo Durbin para el que las variables exógenas (FBI, FBE, LC y LF) fueron promediadas con la matriz de distancias, la variable endógena para este caso de estudio es el flujo de mercancías (FMM). Los resultados derivados de una estimación within o de efectos fijos con respecto a la variable FMM son significativamente diferentes a los anteriores en los que no se integró a la matriz de distancias. El modelo estima una relación positiva entre FBI, FBE, LF, WFBE, WLC y WLF lo que significa que ante aumentos en una unidad de la variable FMM estas también aumentan cada una con respecto a su coeficiente. Los más altos cambios se aprecian en FBE con el 0.51 y en la WLF con 1.34 con una significancia al 0.05 y al 0.001 respectivamente. De la misma manera es interesante el resultado de una relación negativa de la variable endógena con la LC (infraestructura carretera) (-0.04) dentro del 0.05 del área de significancia.

Siguiendo la teoría, Elhorst (2003) indica que los datos de panel son generalmente más informativos, Wooldrige (2010) menciona que se incrementa la precisión en la estimación y Baltagi (2001) encuentra ventajas importantes para el uso de los datos de panel. Por lo que la tabla 3.4 resume el trabajo de investigación con datos de panel y el modelo Durbin. Como se ha mencionado, el modelo de datos de panel en esta investigación se centra en el de “efectos fijos o within”, ya que se suponen los datos son heterogéneos debido a que se trabaja con estados o entidades federativas con un comportamiento que genera diferentes impactos. Es importante mencionar que a lo largo de la investigación se hizo un juego de cambio con la variable endógena y es en este punto en el que gracias a los resultados obtenidos en panel se puede elegir a la de flujos de comercio como la única y más significativa que muestra mejores resultados contra las variables exógenas.

De manera que con el modelo within considera que las variables explicativas: ingresos de las empresas del sector interno y del sector externo como la infraestructura carretera y ferroviaria afectan por igual a las unidades de corte transversal aunque éstas se diferencian por características propias de cada una de ellas, medidas por medio del intercepto (debe hacerse notar que en este modelo se presenta una pérdida importante de grados de libertad ya que los interceptos están asociados con variables dummy). La única variable significativa fue la de los ingresos del sector externo al 0.0001 con un coeficiente de 0.94. Para elegir entre el modelo within y el modelo pool el test rechaza al de efectos comunes con un valor de $8.871e-05$ se elige el modelo de within. Por lo tanto, la prueba de Hausman nos obliga a elegir efectos fijos sobre efectos aleatorios con un p valor de $3.148e-06 < 0.05$ se rechaza la hipótesis nula de igualdad al 95% de confianza y se deben asumir las estimaciones de efectos fijos.

Para concluir los resultados de la investigación que finalizan con un modelo Durbin, tres de las primeras 4 variables exógenas, ingresos de las empresas manufactureras del mercado externo y del interno y la infraestructura carretera sin rezagos de la matriz de distancias fueron significativas al 1% y una exógena con la integración de la matriz de distancias solo fue significativa al 5% que es la de infraestructura ferroviaria. Lo que significa que el nivel de flujos comerciales en la región j depende de la intensidad de la infraestructura ferroviaria en la región j, así como también de la intensidad en las regiones vecinas.

El cambio de significancia del modelo de datos de panel sin haber incluido el modelo Durbin, hasta el momento en el que se agregó la matriz de distancias para promediar a las variables exógenas, fue fundamental y definitivo para concluir con las estimaciones ya que le dio la relevancia a las variables al hacerlas más significativas. Y es por ello que se culmina el análisis empírico con el modelo Durbin (no espacial) después de haber intentado una estimación espacial y no haberse logrado el ajuste de las variables seleccionadas para poder explicarse espacialmente.

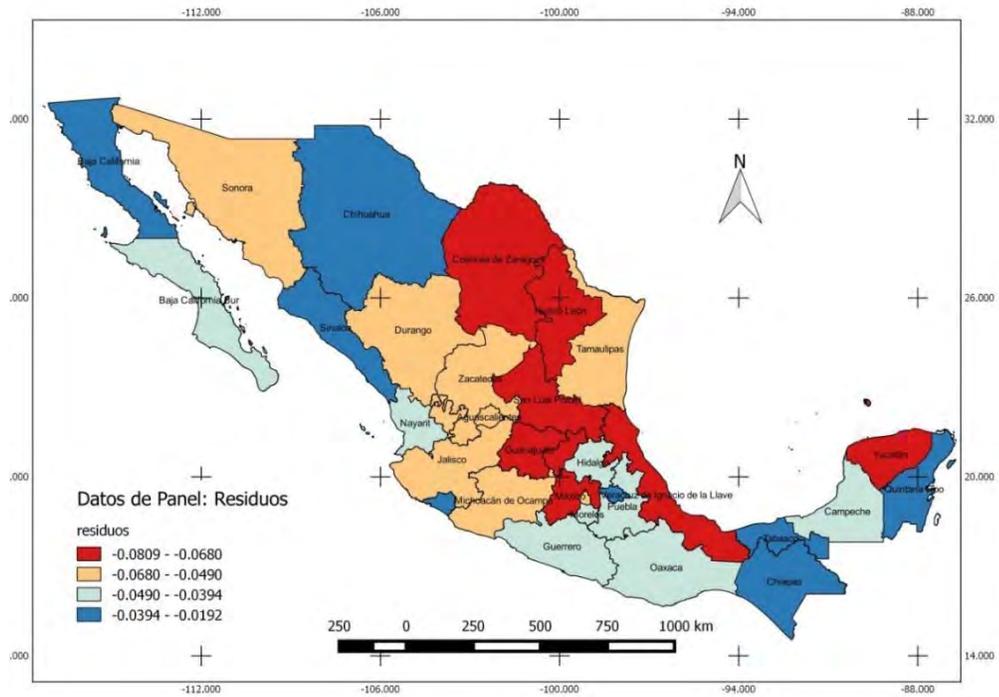
3.5 Errores del Modelo

Los resultados obtenidos en la tabla 3.4 tras la estimación del panel de datos según un modelo Durbin no espacial, con la desventaja de que las naturalezas de los datos de panel no terminan siendo considerados realmente relevantes. Es por ello que se estimó la misma ecuación teniendo en cuenta la modelización espacial, dado que permite integrar la heterogeneidad espacial ya que se trabaja con unidades espaciales heterogéneas, Moreno y Vayá (2000) dan un ejemplo, las regiones avanzadas del norte y las atrasadas del sur, pueden estar, por tanto, directamente relacionada con la localización en el espacio.

Sin embargo, a pesar de los esfuerzos por continuar el análisis econométrico no se obtuvo el resultado de presencia espacial en las variables. Por tanto, para proceder con un análisis del tipo espacial se necesita una especificación para la variabilidad en el espacio que sea parsimoniosa en el sentido de que sólo pueden utilizarse unos cuantos parámetros en dicha especificación. Es decir, resulta necesario limitar el número de diferentes regímenes, evitando problemas de grados de libertad, y esperando poder conseguir así unas estimaciones eficientes (Moreno y Vayá, 2000).

La investigación se limita a presentar un mapeo de los errores del modelo Durbin con base en la ecuación 3.3. La figura 3.5 muestra los residuos de las distintas variables tomadas en el tiempo. Ésta muestra que, la dispersión de los residuos alrededor de su media muestral, aumenta a mayor valor en los estados de Baja California, Colima, Chiapas, Chihuahua, Quintana Roo, Sinaloa, Tabasco y Tlaxcala y en menor valor estados como Coahuila, Guanajuato, Veracruz, Estado de México, San Luis Potosí, Yucatán.

Figura 3.5 Residuos del Modelo Durbin



Elaboración propia con datos del programa nacional de comercio 2008-2013

CONCLUSIONES

Con el fin de incorporar las diferencias regionales de la infraestructura del transporte y el comercio en el análisis de productividad, en esta investigación se estimó la significancia de las variables exógenas (FBI, FBE, LC y LM) respecto a la variable endógena (PROD/FMM) para la industria manufacturera mexicana para cada una de las 32 entidades federativas en los años 2008 y 2013

Para lograr este objetivo se utilizó la metodología de datos de panel, específicamente efectos fijos a través del estimador within con ayuda del software RStudio. Los resultados obtenidos de la prueba F y Hausman nos obligan a basar el análisis del estudio productividad-transporte-comercio en la regresión del modelo Durbin (promedio de las variables exógenas con la matriz de distancias) se observa que el coeficiente de FBE es positivo y el más significativo de todas las variables exógenas. Por lo tanto, el comercio externo tiene cambios más importantes con respecto al flujo de mercancías al igual que la WLF (longitud ferroviaria multiplicada por la matriz) lo que significa que la infraestructura ferroviaria responde positivamente ante cambios en la variable endógena. Seguido de la variable del comercio interno (FBI) que no es menos importante pero si menos representativa que la del comercio externo de acuerdo a su coeficiente y su nivel de significancia. Sin embargo, al evaluar a la variable LC (longitud carretera) resulta significativa pero responde negativamente al flujo de mercancías lo que sugiere el deterioro o ineficiencia de la infraestructura carretera.

A nivel regional (considerando que todas las entidades son heterogéneas) los indicadores si resultan más significativos que si se considerara un análisis homogéneo en la productividad de cada estado. Es decir, con la incorporación del producto de la matriz de distancias con cada variable exógena resulta una evidencia empírica de la relación entre las entidades de acuerdo a su

infraestructura y su comercio con el resto de las otras entidades. Las regiones; centro y norte, fueron las que representaron mayor productividad, así como alta competitividad, estados como Nuevo León, Baja California, Ciudad de México y Querétaro. Éstos son beneficiados por el área geográfica donde se encuentran ubicados, la región norte con frontera a Estados Unidos y la región centro por la conexión que tiene de fácil acceso con la mayoría de los estados.

Por otro lado, la región sur (este-oeste) está representada en términos generales en su mayoría de los puertos marítimos, como Veracruz, Oaxaca y Michoacán. Sus índices no se encuentran característicos en competitividad y productividad pero no se puede afirmar que la infraestructura en forma de puerto marítimo sea el único factor que disminuya el nivel de crecimiento de la región. Por lo tanto, se concluye que la infraestructura y el flujo de mercancías están correlacionados con las regiones, al contrario de la productividad del sector manufacturero. Lo que significa únicamente que ante cambios en la proporción de infraestructura el flujo de mercancías se ve afectado.

Se debe aceptar la presencia de variables omitidas a causa de la falta de información a nivel regional y por periodo en los sistemas de información de México, con las que se pudo haber tenido un modelo más robusto. Adicionalmente para investigaciones futuras, el modelo productividad-transporte-comercio se profundizará ya que puede ser de gran ayuda para la comprensión de las disparidades regionales y la creación de políticas comerciales.

Resulta innegable que la infraestructura influya en el comercio, pero si es cuestionable que influya en la productividad de un sector (manufacturero). Aunque el impacto que ésta (la productividad) ejerce es de forma espacial, la productividad de un estado afecta a la productividad de sus estados vecinos. Es decir, la variable está condicionada a las variables que se incorporan y al espacio geográfico que se evalúa.

Mientras se necesita adaptar las variables exógenas correctas para la explicación del desarrollo de la productividad, que se traduce en el nivel de competitividad logrado por la conexión de las entidades a través de la infraestructura que genere un crecimiento potencial del comercio en los estados.

Finalmente, se debe de considerar el tema del transporte como prioritario ya que México geográficamente se encuentra en un lugar privilegiado respecto a otros países de Latinoamérica; con salidas al mar para conectar con Asia, suficiente extensión territorial, colindancia con uno de los más importantes países del mundo (Estados Unidos). Promover la realización de proyectos de infraestructura productiva con impacto regional, que ayuden a mejorar la conectividad al interior de las entidades con el resto del país, así como con los mercados internacionales. Se exige continuar con este estudio sin olvidar que se cumpla con los objetivos nacionales llevado a cabo por investigadores, inversionistas y servidores públicos juntos con un mismo fin, el desarrollo y crecimiento de la nación.

BIBLIOGRAFÍA

- Álvarez, N. (2003) *Introducción a la Econometría*. España: Ed. Académicas
- Anderson, et al. (1999) *Estadística para Administración y Economía*. México: Ed. International Thomson Editores.
- Anselin, L. (1995) *Las medidas locales de autocorrelación espacial*. Estados Unidos: Ed. Geographical analysis.
- Anselin, L. (2005) *GeoDa a 0.9 User's Guide. Spatial Analysis Laboratory*. Estados Unidos: Ed. Department of Agricultural and consumer economics.
- Anselin, L. (1999) *The future of spatial analysis in the social sciences*. Estados Unidos: Ed. Geographic Information Sciences.
- Arrow K. y Debreu G. (1954) *Existence of an equilibrium for a competitive economy*. Estados Unidos: Ed. Econometrica, Journal of the Econometric Society.
- Aschauer D. (1989) *Es productivo el gasto público*, Países bajos: Ed. Journal of Monetary Economics.
- Asuad N. (2013) *Desarrollo regional y política en México 1994-2014*. México.
- Asuad N. (2001) *Importancia actual del estudio de la economía regional y del desarrollo urbano*. México: Ed. Colegio de Puebla, A. C y BUAP
- Baltagi B. (2001) *El análisis econométrico de datos de panel*. Estados Unidos: Ed. John Wiley & Sons.
- Baltagui B. (2011) *Econometrics*. Estados Unidos: Ed. Springer.
- Batanero et.al. (1991) *Efecto del modelo combinatorio implícito en el razonamiento combinatorio en escolares*. Estados Unidos: Ed. Educational Studies in Mathematics.
- Best M. (1990) *La nueva competencia: Instituciones de la reestructuración industrial*. Estados Unidos: Ed. Harvard University Press.
- Black J. (1997) *Oxford diccionario de economía*. Estados Unidos: Ed. Oxford University Press.

- Boudeville J. (1970) *Las zonas económicas*. París: Ed. PUF.
- Boudeville J. (1966) *Problemas de planificación de economía regional*. Reino Unido: Ed. Edinburgh at the University Press.
- Buendía E. (2013) *El papel de la ventaja competitiva en el desarrollo económico del país*. México: Ed. Universidad Autónoma Metropolitana.
- Buzai, G. (2007) *Métodos Cuantitativos en Geografía de la Salud*. Argentina: Ed. Universidad Nacional de Luján.
- Canning y Bennathan (1999) *La tasa de retorno social de las inversiones en infraestructura*. Estados Unidos: Ed. Policy Research Working Papers.
- Capello R. (2011) *Globalización y crecimiento regional en Europa: tendencias del pasado y escenarios del futuro*, Alemania: Ed. Fratesi and L. Resmini.
- Casas L. (2006) *Teorías del crecimiento regional y el desarrollo divergente. Propuesta de un marco de referencia*. México: Nóesis. Revista de Ciencias Sociales y Humanidades, vol. 15, núm. 30.
- Castells y Borja (1998), *La gestión local de la economía global: ¿ciudades globales o redes globales de nodos urbanos? en Local y Global, la gestión de las ciudades en la era de la información*. España: Ed. Santillana
- Cendrero M. (2012) *El análisis de la respuesta del sistema geomorfológico al cambio en las unidades naturales y humanos*. España: Ed. Geomorphology in press.
- CEPAL (2004) *Unidad de Inversiones y Estrategias Empresariales, La inversión extranjera directa en América Latina*. Chile: Ed. CEPAL.
- Céspedes et al. (2014) *Guía para estimar la pobreza infantil*. Ed. UNICEF-CEPAL.
- Chasco C. (2003) *Econometría espacial aplicada, Consejería de economía e Innovación tecnológica*. España: tesis doctoral.
- Chasco et al. (2000) *Métodos gráficos del análisis exploratorio de datos espaciales*, España: Ed. Universidad Autónoma de Madrid.

- Christ, C. (1983) *The Founding of the Econometric Society and Econometrica*. Estados Unidos: Ed. Econometrica.
- Christaller, W. (1966) *Lugares centrales en el sur de Alemania*, Prentice Hall. Estados Unidos.
- Cliff y Ord (1981) *Proceso espacial: modelos y aplicaciones*. Londres: Ed. Taylor & Francis.
- Correa, G. (2009) *Competencia por productos de México y China en el mercado de los Estados Unidos*. México: Ed. H. Cámara de Diputados, LX Legislatura y la Universidad Autónoma Metropolitana.
- Cuadrado, J. (2012) *¿Es tan “nueva” la “Nueva Geografía Económica”? Sus aportaciones, sus límites y su relación con las políticas*. España: Serie documentos de trabajo Universidad de Alcalá.
- Cuadrado, J. (2006) *El desarrollo de los estudios de Economía Regional en España*. España: revista de estudios regionales nº 75.
- De Arce y Mahía (2012) *Introducción a la econometría. Un enfoque moderno: Interpretación de los Parámetros de un Modelo Básico de Regresión Lineal*. México: Ed. Thomson.
- Duch, N. (1998) *Localización industrial*. España: Ed. Mella Márquez.
- Dutilleul, P. (1998) *Incorporating scale in ecological experiments: Study design, Scale Issues in Ecology*. Estados Unidos.
- E, Malinvaud. (1966) *Statistical Methods of Econometrics*. Estados Unidos: Ed. Rand McNally.
- Elhorst, J. (2003) *Specification and estimation of spatial panel data models*. Países bajos: Ed. International Regional Science Review.
- Forbes (2015), *Manufacturas el “arma secreta” de México contra China*. México.
- Fotheringham, S. (2000) *Perspectivas sobre el análisis de datos espaciales*. Londres: Ed. SAGE.
- Frisch, R. (1993) *Econometrica*. Noruega: Ed. Econométrica. Noruega.
- Fujita, M. (2000) *Economía espacial*. Inglaterra: Ed. Ariel S.A.

- Gannon, C. y Liu, Z. (1997) *Poverty and transport. Estudio de discusión TWU-30 Transportation, Water and Urban Development Department*. Estados Unidos: Ed. Banco Mundial.
- Gaus, C. (1809) *Theoria motus corporum coelestium, translation*. Estados Unidos: Ed. C.H. Davis.
- Geary, R. (1954) *La relación de contigüidad y asignación estadística*. Irlanda: Ed. The Incorporated Statistician.
- Good, I. (1983) *La filosofía del análisis exploratorio de datos*. Estados Unidos: Ed. La Filosofía de la Ciencia.
- Gujarati, D. (2009) *Econometría básica*. Estados Unidos: Ed. McGraw-Hill.
- Haavelmo, T. (1944) *The Probability Approach in Econometrics*. Noruega: Ed. Econometrica
- Hartwig y Dearing. (1979) *Exploratory Data Analysis*. Estados Unidos: Ed. California: Sage Publications.
- Horber y Ladiray. (1995) *Análisis Exploratorio de los datos, II Seminario de Capacitación de docentes en Producción y tratamiento de datos de Investigación en ciencias humanas*, Venezuela.
- Iregui et al., Productividad regional y sectorial en Colombia: un análisis utilizando datos de panel, Ensayos sobre Política Económica, 25(53), Colombia, 2006.
- Kaldor, N. (1966) *Causes of the Slow Rate of Economic Growth of the United Kingdom*. Estados Unidos: Ed. Universidad de Cambridge.
- Keeble, et al. (1986) *Peripheral regions in a Community of Twelve Member States*. Estados Unidos.
- Krieger, L. (1999) *Competition, Innovation and Competitiveness in Developing Countries*. París: Ed. Organisation for Economic Co-operation and Development.
- Krugman, et al. (2000) *Economía Internacional: teoría y política*. Estados Unidos: Ed. Addison Wesley.

- Krugman, P. (1995) *La globalización y la desigualdad de las naciones*. Estados Unidos: Ed. CX.
- Labra y Torrecillas (2014) *Guía Cero para Datos de Panel. Un enfoque práctico*. España: Ed. Universidad Autónoma de Madrid.
- Legendre & Fortin (1989) *Patrón espacial y análisis ecológico*. Francia: Ed. International Journal.
- León, J. (2004) *Modelo de Competitividad Global de la Industria de Piel de Cocodrilo Moreletii*. México: Ed. Universidad Autónoma de Sinaloa, Escuela de Economía.
- Losch A. (1954) *La economía de localización*. Estados Unidos: Ed. Yale University Press.
- Losch A. (1944) *La regulación espacial de los derechos económicos*. Estados Unidos.
- Lucas, R. (1976) “*Econometric Policy Evaluation: A Critique*” Estados Unidos: Ed. the Phillips Curve and Labor Markets.
- Mahía R. (2000) *Introducción a la Especificación y Estimación de Modelos con Datos de Panel*. España: Ed. Universidad Autónoma de Madrid.
- Malinvaud, E. (1966) *Statistical Methods of Econometrics*. Holanda.
- Mayorga y Muñoz, (2000) *La Técnica de Datos de Panel una Guía para su uso e Interpretación*. Costa Rica: Ed. Banco Central de Costa Rica
- McAleer, et al.(1985) *What Will Take the Con Out of Econometrics*. Estados Unidos: Ed. American Economic Review.
- Michael P. (1999) *Clusters y Competencia: nuevas agendas de empresas, gobiernos e instituciones*. Estados Unidos: Ed. Harvard Business School,
- Miguel. A. (2004) *Principios de Economía y Desarrollo*. México: Ed. Ciencia Regional.
- Montero. R. (2005) *Test de Hausman*. España: Ed. Universidad de Granada. España.
- Moran P. (1950) *Notes on continuous stochastic phenomena*. Reino Unido: Ed. Oxford University.

- Moreno y Vayá (2000) *Técnicas Econométricas para el Tratamiento de Datos Espaciales: La Econometría Espacial*. España: Ed. Universidad de Barcelona.
- Muñetón y Vanegas (2014) *Análisis espacial de la pobreza en Antioquia*. Colombia: Ed. Equidad & Desarrollo.
- North, D. (1970) *Las Posibilidades de Detonar a Partir de la Capacidad Exportadora*. Estados Unidos.
- North, D. (1977) *Markets and other allocation systems in history: the challenge of Karl Polanyi*. Reino Unido: Ed. Journal of European Economic History.
- Núñez y Lucas Jr. (2005) *Crítica a la Econometría Tradicional, Origen de los Enfoques Económicos Contemporáneos*. México: Ed. Carta de Políticas Públicas, UNAM Facultad de Economía.
- Ocegueda, J. (2003) *El Sector Manufacturero y la Restricción Externa al Crecimiento Económico de México*, México: Ed. Facultad de Economía UNAM.
- Pagan, A. (1987) *Twenty Years After: Econometrics, 1966–1986*. Francia: Ed. Anniversary Conference.
- Parra, J. (2002) *Análisis Exploratorio y Análisis Confirmatorio de Datos*. Venezuela: Ed. Espacio Abierto.
- Parra, J. (1999) *Contenidos de Estadística en Universidades de Latinoamérica*. Venezuela: Ed. Centro de Estadística e Investigación de Operaciones de La Universidad del Zulia.
- Pérez, J. (2006) *Econometría Espacial y Ciencia Regional Investigación Económica*. México
- Perroux, F. (1949) *Las macrodecisiones, Economía aplicada*.
- Pesaran, M. (1990) *Econometrics, in J. Eatwell, M. Milgate and P. Newman; the New Palgrave: Econometrics*. Estados Unidos
- Polése, M. (1998) *Economía urbana y regional: introducción a la relación entre territorio y desarrollo*. Francia: Ed. UCOL.

- Porter, M. (2011) *Estrategia de competitividad: Técnicas para analizar industrias y competidores*. Estados Unidos: The free press.
- Quintana y Andrés (2014). *Técnicas Modernas de Análisis Regional*. México.
- Richardson, H. *Economía regional y urbana*. Madrid: Ed. Alianza Editorial.
- Rodrik, D. (2006) *¿Qué es tan especial en las exportaciones de China?* Estados Unidos: Ed. NBER
- Sánchez, P. (2008) *Alcances y límites de los métodos de análisis espacial para el estudio de la pobreza urbana*. México: Ed. Papeles de Población
- Sancho et al. (2001) *Comparativa de indicadores de sostenibilidad para destinos desarrollados, en desarrollo y con poblaciones vulnerables*. España: Ed. Annals of tourism research.
- Smith, et al. (2008) *La mejora de la competitividad internacional de la industria de productos forestales a través de proporcionar métodos de transporte*. Virginia: USDA, Estados Unidos.
- Thirlwall, A. (2003) *La Naturaleza del Crecimiento Económico. Un marco alternativo para comprender el desempeño de las naciones*. México: Ed. Fondo de Cultura Económica (FCE).
- Tintner, G. (1968) *Methodology of Mathematical Economics and Econometrics*. Estados Unidos: Ed. The University of Chicago Press.
- Tintner, G. (1953) *The Definition of Econometrics*. Ed. Econometrica.
- Tobler, W. (1970) *Un crecimiento urbano simulado en la Región de Detroit*. Estados Unidos: Ed. Economic Geography.
- Tomassian, et al. (2010) *Políticas integradas de infraestructura, transporte y logística: experiencias internacionales y propuestas iniciales*. Chile: Ed. División de recursos naturales e infraestructura.
- Torrelavega (2005) *Ajuste por Mínimos Cuadrados*. México: Ed. Escuela Politécnica de Ingeniería de Minas y Energía.

- Tukey, J. (1977) *Análisis exploratorio de los datos*. Estados Unidos: Ed. Reading PA.
 - Tukey, J. (1980) *We need both Exploratory and Confirmatory*. Estados Unidos: Ed. The American Statistician.
 - Upton y Fingleton (1985) *Análisis espacial de los datos*. Estados Unidos: Ed. Point pattern and quantitative.
 - Vázquez, A. (2005) *Desarrollo endógeno. Teorías y políticas de desarrollo territorial*. España: Ed. Investigaciones Regionales.
 - Vázquez, A. (2002) *Las nuevas fuerzas del desarrollo*. España: Ed. UAM-Antoni Bosch.
 - Viego, V. (2004) *Desigualdades regionales y costos en transporte en Argentina*. Argentina: Ed. Universidad Nacional del Sur.
 - Viton, P. (2010) *Notes on Spatial Econometric Models*. Estados Unidos: Universidad de Ohio.
 - Volker, et al. (1985) *Lo que se llevará a cabo la estafa de Econometría*. Estados Unidos: Ed. American Economic Review.
 - Weber, A. (2009) *Teoría de la localización de las industrias*. Estados Unidos: Press Chicago.
 - Wooldrige, J. (2010) *El análisis econométrico de sección transversal y datos de panel*. Estados Unidos: Ed. Universidad de Michigan.
 - Wooldrige, J. (2008) *Métodos de datos de panel para las variables de respuesta fraccionarios con una aplicación para probar los índices de aprobación*. Estados Unidos: Ed. Journal of Econometrics.
 - Zamora y Pedraza (2013) *Competitividad del transporte en el marco del comercio internacional*. México: Ed, Universidad Michoacana de San Nicolás de Hidalgo.
-