UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE ECONOMÍA
DIVISION DE ESTUDIOS DE POSGRADO

"ENSAYOS SOBRE CRECIMIENTO Y DESIGUALDAD: ALTERNATIVAS EN EL MANEJO Y ADQUISICIÓN DE DATOS"

Por

DAVID ROBERTO VALENZUELA VEGA

Tesis presentada como requisito parcial para obtener el grado de Doctorado en Ciencias Económicas

JUNIO 2022

“ENSAYOS SOBRE CRECIMIENTO Y DESIGUALDAD: ALTERNATIVAS EN EL MANEJO Y ADQUISICIÓN DE DATOS”

DAVID ROBERTO VALENZUELA VEGA

Aprobación de Tesis:

Asesor de la Tesis **Dr. Edgar Mauricio Luna Domínguez**

Dra. Cinthya Guadalupe Caamal Olvera

Dra. Joana Cecilia Chapa Cantú

Dr. Vicente Germán Soto

Dr. Ernesto Aguayo Téllez

DR. PEDRO ANTONIO VILLEZCA BECERRA
Director de la División de Estudios de Posgrado
De la Facultad de Economía, UANL
Junio, 2022

Dedicatoria

A mi madre, que me enseñó a superarme.
A mi padre, que me enseñó a sonreír.

Agradecimientos

Quiero agradecer en primer lugar a mi asesor el Dr. Edgar Luna, por su apoyo y guía en la realización de este trabajo de tesis. A mis profesoras la Dra. Joana Chapa y la Dra. Cinthya Caamal por brindarme los conocimientos necesarios para desarrollarla, así como transmitirme el gusto e interés en los temas desarrollados en el presente trabajo.

Agradezco al CONACYT por el apoyo económico y de infraestructura brindado durante la realización mis estudios e investigación.

De igual manera agradezco a mis sinodales, a la Universidad Autónoma de Nuevo León y en general a toda la Facultad de Economía, incluidos mis profesores y administrativos quienes con su trabajo y dedicación hicieron posible la realización de este trabajo.

También agradezco a mi familia, a Maggie y a mis amigos por el interés, apoyo, paciencia y ánimos que me brindaron durante todo el tiempo que tomó cumplir esta meta.

Por último agradezco a mis compañeros y compañeras por las incontables horas de interesantes pláticas, reflexiones y críticas que me enriquecieron tanto personal como profesionalmente.

Índice general

Introducción **1**

1. Construcción de variables de agrupamiento estructural: Un enfoque de selección endógena en el marco del crecimiento económico. **3**
1.1. Introducción . . . 3
1.1.1. La interconexión subyacente en el planteamiento teórico. . . . 4
1.1.2. Un enfoque estructural. . . . 6
1.2. Índice de Agrupamiento Estructural. . . . 6
1.2.1. Construcción del conjunto estructural. . . . 7
1.2.2. Proceso de Agrupamiento. . . . 8
1.2.3. Estimación del Modelo. . . . 8
1.3. Datos . . . 9
1.4. Resultados . . . 10
1.4.1. Del proceso de Agrupamiento . . . 11
1.4.2. De la regresión . . . 13
1.5. Conclusiones . . . 18

2. Desigualdad espacial metropolitana, una vista desde el espacio. **20**
2.1. Introducción . . . 20
2.1.1. Desigualdad espacial en México y el mundo. . . . 21
2.2. Metodología . . . 24
2.3. Datos . . . 25
2.4. Resultados . . . 27
2.5. Conclusiones . . . 30

3. Clubes de convergencia metropolitana: un análisis a través del índice lumínico. **32**
3.1. Introducción . . . 32
3.1.1. Estudios regionales de convergencia estocástica para México . . . 35
3.2. Metodología . . . 37
3.2.1. Algoritmo de convergencia . . . 39
3.3. Índice Lumínico per cápita por área metropolitana . . . 40
3.3.1. Estadísticos Descriptivos del Índice Lumínico . . . 45
3.4. Resultados . . . 48
3.4.1. Heterogeneidad interestatal. . . . 50

3.4.2. Distribución geográfica de los Clubes. 52
3.4.3. Análisis de las sendas de transición relativas. 52
3.5. Conclusiones 56

Conclusiones finales **58**

Bibliografía **59**

Anexo **65**

Introducción

La economía positiva busca explicar y describir los fenómenos económicos por medio de la aceptación o rechazo de hipótesis asociadas a relaciones causa efecto. Si bien el mero planteamiento teórico de los determinantes de la dinámica económica representa un reto considerable, este no es el único problema al que se enfrenta, existe además el problema de disponibilidad de los datos. La disponibilidad de datos puede jugar dos papeles en cuanto a la estimación de un modelo representativo. Por una parte la falta de información suficiente en tiempo y/o en forma es un problema evidente respecto al tema de disponibilidad de datos. Esta situación se agrava cuando la falta de información se presenta en situaciones donde podría tener más impacto la exploración de los determinantes económicos, como es el caso de regiones pobres sin instituciones capaces de generar dichos datos. Por otra parte, la disponibilidad de grandes cantidades de información no resulta siempre beneficiosa, puesto que la selección de las variables es en sí un problema complejo. Esta situación lleva a un problema de especificación del modelo, ya que la inclusion de un gran número de variables puede llevar a la consecuente pérdida de grados de libertad del modelo y por ende afectar la representatividad del mismo. Motivado por dichas situaciones la presente tesis expone soluciones alternativas a estos problemas en los artículos que componen este trabajo. En el primer artículo se presenta una metodología para la construcción de un índice que concentra una gran cantidad de información, mientras que en los dos artículos restantes se utiliza una fuente de información alternativa para llevar a cabo el estudio de las dinámicas económicas a un nivel de detalle que no es posible realizar con las fuentes de información tradicionales.

El documento está compuesto por tres artículos donde cada uno de ellos corresponde a un capítulo del trabajo de tesis. En el capítulo uno se plantea que el estudio del crecimiento económico ha pasado por varias fases, desde el planteamiento de modelos teóricos hasta la búsqueda de evidencia empírica que los sustenten. Sin embargo la complejidad inherente al proceso, expresada por la gran cantidad de determinantes que parecen influir en el crecimiento económico, dificulta el establecimiento de una metodología bien definida para la construcción de los modelos empíricos, y en consecuencia los resultados de éstos suelen ser sensibles a la selección de variables explicativas, es decir, se genera un problema de identificación del modelo. Por tal motivo este capítulo tiene el objetivo de aportar una metodología para la construcción de modelos representativos, donde se sintetiza una gran cantidad de información a través de la construcción de una variable de agrupamiento que refleja al menos parte del efecto que tienen las características económicas, tecnológicas, geográficas y demográficas en la senda de crecimiento económico de los países. La inclusión de la que llamamos *Índice de Agrupamiento Estructural* resulta en una mejora de ajuste de modelo del 52 % con respecto a un modelo de agrupación clásico regional. De tal forma que se abona a una comprensión

más detallada del camino de crecimiento de las naciones.

En el capítulo dos se busca evidencia de la relación de U invertida entre el ingreso y la desigualdad espacial a un nivel de zonas metropolitanas en México. Tomando como medida de desigualdad el coeficiente de variación ponderado calculado sobre las municipalidades que conforman una zona urbana y generando así una medida de desigualdad espacial. El problema de disponibilidad de los datos a este nivel se enfrenta utilizando un proxy del ingreso, a través de las luces captadas desde el espacio como lo propuso Henderson (2012). Los resultados señalan que los efectos del ingreso en la desigualdad son de signo negativo, es decir que el incremento en el ingreso siempre disminuye la desigualdad espacial, presentando solo diferencias de nivel en un par de regiones.

El capítulo tres se enfoca en identificar las zonas metropolitanas que convergen y divergen económicamente en México. Para estimar la convergencia entre las diferentes zonas, se utilizan el índice lumínico y la metodología propuesta por Phillips y Sul (2007), la cual permite la conformación de subgrupos de zonas en los que se puede presentar convergencia o divergencia entre los elementos del subgrupo generado. Al encontrarse cinco clubes de convergencia entre las diversas zonas metropolitanas de México, se demuestra que la dinámica de convergencia económica es heterogénea, por lo que resulta relevante un estudio a nivel de cada zona. De igual manera, se identifica que existe un club de convergencia con zonas metropolitanas predominantemente en el norte del país, sin embargo, la división generada por los procesos de convergencia es más compleja que una simple división Norte-Sur. Si bien la metodología de Phillips y Sul (2007) permite generar clubes de convergencia sin necesidad de hacer supuestos relacionados con las propiedades de estacionariedad de las variables, esta metodología no permite identificar los factores específicos que hacen que estos grupos se formen. El presente es el primer trabajo que analiza convergencia económica entre zonas metropolitanas de México.

Capítulo 1

Construcción de variables de agrupamiento estructural: Un enfoque de selección endógena en el marco del crecimiento económico.

1.1. Introducción

Los resultados de los modelos empíricos de crecimiento pueden ser susceptibles a las variables que se escogen para el estudio. Si bien la definición y papel de los determinantes del crecimiento están fundamentados teóricamente, el expresar dicha relevancia empíricamente no es tarea fácil. Aunque se han encontrado variables que se presentan consistentemente, debido a que representan de forma eficaz estos determinantes teóricos, existen otras que dependen de las variables incluidas.Estas variables que dependen de otras para mostrar significancia son las que Levine y Renelt (1992) refieren como un determinante *frágil* en su trabajo. Desde un punto de vista utópico bastaría con incluir todas las variables de forma simultánea e ir descartando aquellas que no resulten relevantes al modelo; sin embargo esto no es posible debido a la gran cantidad de variables que parecen afectar la dinámica del crecimiento económico. El presente trabajo aborda este problema y plantea una forma de ingresar esta información a través de un índice que refleja el estado de similitud multidimencional de los países de tal forma que mejora la calidad del modelo hasta un 52 % con respecto a los modelos con variables de agrupamiento clásicas.

La contribución del presente trabajo es presentar una metodología sencilla que brinda la posibilidad de incluir de manera indirecta la totalidad de una serie de variables que llamaremos variables descriptivas. Esta metodología se basa en la idea de que es posible extraer información de un amplio número de variables y de sus interacciones de forma tal que los modelos se vuelvan robustos y permitan identificar los efectos de los determinantes de una forma diferenciada entre diversas economías.

En lo que resta de la presente sección se aborda la evolución de los modelos que tratan el crecimiento económico y las dificultades a las que se han enfrentado. En la segunda sección se explica la metodología utilizada para construir el índice de agrupamiento estructural. La

sección tres aborda lo referente a los datos utilizados para el estudio. La sección cuatro presenta tanto los resultados del agrupamiento como los de la regresión y la quinta sección cierra con las conclusiones.

1.1.1. La interconexión subyacente en el planteamiento teórico.

Bajo un enfoque clásico del estudio del crecimiento económico se ha tratado de asociar la dinámica de crecimiento a un determinante o grupo de determinantes en específico. Esto está justificado, ya que el sentido de un modelo es extraer la mayor información con el menor número de variables, de otra forma el modelo sería tan complejo como el fenómeno mismo. Sin embargo la complejidad del fenómeno, ha 'jalado' los modelos hacia otros más refinados. En ese refinamiento se pueden ver por lo menos dos fenómenos, el incremento en el número de variables utilizadas y el de la especialización de las variables. El primero se explica fácilmente ya que al tratarse de un fenómeno complejo en el que toman relevancia muchos factores mientras más información de distintos rubros se agregue es mejor. El segundo fenómeno no se explica tan directamente, creemos obedece en parte a un efecto de interacción entre las variables, una especie de correlación por umbrales. Para mostrar este punto tomemos los trabajos seminales de Swan (1956) y Solow (1957) , en estos trabajos se hace notar la importancia de la acumulación de capital físico en la dinámica de crecimiento, aún y cuando dentro del modelo se hace referencia a la población, es la acumulación de capital físico lo que mueve principalmente el modelo. Posteriormente vinieron trabajos como los de Mankiw et al. (1992), Kortum(1997) y Segerstrom(1998) entre otros, en los que la variable principal pasó a ser el capital humano, para el trabajo de Mankiw, y la investigación y el desarrollo de nuevas tecnologías (I&D) para los dos restantes. Como se puede observar los modelos se fueron 'refinando', de capital a capital humano y a su vez a investigación. En este sentido resulta difícil pensar en un nivel de I&D tal que tenga efecto en la senda de crecimiento de un país, que no esté soportado por una base sólida de investigadores (capital humano) e infraestructura (capital físico) dispuestos para alcanzarlo. Si bien la selección de trabajos mencionados tiene la intención de hacer evidente la conexión entre las variables utilizadas para explicar el crecimiento económico, la conexión se podría hacer hacia frentes menos evidentes, como el nivel de apertura económica, los niveles de desempleo o la seguridad pública. Cada uno de estos aspectos no se pueden concebir como aislados y se relacionan entre sí, en este sentido no resultaría sorprendente que la combinación de niveles de estas variables generen una dinámica particular, y que esta dinámica no dependa únicamente del nivel individual de cada variable.

El enfoque empírico.

Adicionalmente a los esfuerzos realizados en favor de plantear un modelo teórico se inició una corriente que criticaba por medio de la validación empírica de las teorías de crecimiento. El trabajo de Barro (1991) en el que busca evidencia de que los países pobres crecen a tasas más altas que los países ricos como lo plantea el modelo de Solow es el primer trabajo en esta dirección. Entre los resultados que obtiene se encuentra el hecho de que no solo importa que el capital inicial en una economía sea menor para experimentar tasas más altas de crecimiento, sino que es necesario que este esté acompañado de un nivel de capital humano considerable

para experimentar altas tasas de crecimiento. Debido a estos resultados resalta el hecho de que una variable no puede explicar por sí misma el crecimiento de los países sino que es la interacción entre ellas lo que puede modelarlo de mejor manera. De aquí se inició una carrera por incluir un sin número de variables que presumiblemente explicarían de forma más precisa el crecimiento de los países, validando estas suposiciones con evidencia empírica representada por la significancia dentro de una regresión.

Una vez hecha evidente la necesidad de utilizar más de una variable como determinante del crecimiento económico surgieron varios trabajos en los que se incluían diversas variables que lo explicaban, hasta llegar al punto en el que se utilizaban más de cincuenta variables a las que se les atribuían un poder explicativo en la construcción de los modelos. Es entonces cuando surge el trabajo Levine y Renelt (1992), en donde se busca determinar si las variables explicativas son robustas al modelo, en este trabajo se utilizó un *extreme-bounds analysis* para probar la robustez de los coeficientes estimados al someterlo a cambios en el conjunto de información. Lo que hace es evaluar si el coeficiente de una variable permanece significativo y sin cambio de signo ante un cambio en el conjunto de covariantes, de ser así se califica como robusta a la variable y frágil en caso contrario. Uno de sus resultados fue que muy prácticamente ninguna de las variables calificaban como robusta. De esta forma el trabajo implicaba que no existía una relación consistente entre cada una de las variables y el crecimiento económico. Sin embargo mas tarde llegaría el trabajo de Sala-i-Martin (1997) que se contrapone a la visión pesimista de Levine y Renelt. En su trabajo Sala-i-Martin incluye una ponderación a los coeficientes calculados, de tal forma que aquellos coeficientes presentes en un modelo con mejor ajuste tienen un mayor peso en la estimación del coeficiente promedio. De esta forma Sala-i-Martin ya no solo busca de forma individual las variables que deberían ser utilizadas, sino que además da un paso en busca del "verdadero modelo". Cabe mencionar que en este trabajo se encontraron 21 variables significativas de las casi 60 analizadas, dejando un panorama mucho más prometedor en el estudio empírico del crecimiento económico.

Como se vuelve evidente en estos trabajos el problema en el estudio del crecimiento económico se vuelve uno de identificación del modelo, donde el punto crucial se enfoca en la selección de la combinación de variables pertinentes. Entre los trabajos que siguieron esta vertiente se encuentran el de Doppelhofer, Miller y Sala-i-Martin (2004) que proponen un método de selección de variables con bases bayesianas. Posteriormente Masanjala y Papageorgiou (2005) utilizan este método bayesiano para una subconjunto de datos que presumiblemente comparten características y donde al aplicarlo ponen de manifiesto características relevantes del grupo en particular. Más adelante Basturk, Paap y Van Dijk (2012) permiten la autoselección de grupos de países con base en los coeficientes significativos endogeneizando el proceso de agrupación. Es entonces que podemos ver el problema de especificación del modelo como uno que se desarrolla en dos etapas, una en la que se busca un nivel de agregación de países tal que las características que comparten ayuden a extraer mayor información por parte del modelo y una segunda etapa en la que se identifican las variables relevantes al modelo. En el presente trabajo se aborda la primera etapa, consistente en la identificación de los niveles de agregación.

1.1.2. Un enfoque estructural.

Respecto al problema de especificación del modelo en el contexto de los modelos de crecimiento nos enfrentamos de entrada a dos situaciones. La primera es la problemática de la inclusión de variables, ya que quisiéramos evitar al menos parcialmente el sesgo por variable omitida incluyendo el set de información lo más grande posible, pero nos vemos limitados por la pérdida de grados de libertad asociada. La segunda es la posible heterogeneidad de los efectos de los determinantes entre los distintos países. De esta forma los efectos 'verdaderos' de un determinante podrían estar escondidos detrás de esa heterogeneidad y de la falta de información en el modelo. En este tenor el trabajo de Masanjala y Papageorgiou (2005) hizo grandes avances y encontró que los efectos comúnmente aceptados como válidos para el conjunto completo de países no resultaban así para el subconjunto formado por los países pertenecientes al África subsahariana. Al aislar este grupo salieron a relucir los efectos propios de la dinámica en la que se encuentran estos países. Esta selección del grupo de estudio habría que decir que fue *a priori* puesto que está basada en una percepción (en este caso evidente) de que la delimitación geográfica tenía un impacto en el efecto de las variables, o mejor dicho, que la adscripción geográfica funcionaba como una proxi de las similitudes de la dinámica que experimentan dichos países.

El presente estudio aborda dichas problemáticas de forma simultánea, construyendo una variable categórica que identifica las similitudes proporcionales entre países, generando una agrupación basada en las características estructurales de los mismos. Cabe mencionar que la construcción de la variable es mediante un proceso no supervisado, es decir, se desarrolla bajo la guía de los datos en sí y no sobre una construcción predefinida. Esta agrupación que llamamos Índice de Agrupamiento Estructural mejora la estimación realizada con los determinantes utilizados comúnmente de forma tal que permite comprender un poco mejor la dinámica de crecimiento en la que se desenvuelven los países de una forma diferenciada.

1.2. Índice de Agrupamiento Estructural.

El índice de agrupamiento estructural se basa en la idea de que los efectos de estructura son más importantes que los efectos de nivel. Los efectos de nivel son aquellos que representan el valor real o relativo de alguna variable, por ejemplo, la tasa de inversión promedio de Mongolia fue del 40 % del PIB para el periodo de 1971 a 2014 mientras que la de Sierra Leona fue del 12 % para el mismo periodo, estos países presentan una gran diferencia en el nivel de inversión, fenómeno que puede llevar también a una diferencia en los rendimientos de la tasa de inversión sobre el crecimiento, como bien lo aborda Durlauf (1995) en su trabajo. La diferencia en el rendimiento de la inversión inducida por la diferencia en el nivel se puede observar al calcular el modelo a ambos niveles (altos y bajos), obviamente es necesario incluir dicha variable en niveles en la regresión, y lo mismo aplica para cada una de las variables de las que se desee tomar en cuenta un efecto diferenciado en niveles, con la subsecuente restricción de la pérdida en los grados de libertad como lo menciona Durlauf en su trabajo.

Si bien el enfoque anterior es de suma importancia, nace la inquietud de si existen diversos factores que cambien su efecto dependiendo su nivel, y más aún si dichos efectos se ven afectados por uno o varios niveles en otras variables. Debido a que el estudio individual (y aparte por niveles) resulta en un incremento de la complejidad del modelo se opta por hacer

un análisis de forma agregada, de tal forma que se pueda sintetizar un gran conjunto de información de forma concisa y lo más compacta posible, es así como nacen los efectos de estructura. Los efectos de estructura son aquellos asociados a la combinación de factores vistos desde una perspectiva proporcional. Estos se forman por la combinación de todos los niveles relativos de las variables incluidas, de tal forma que se caracteriza la estructura asociada a dichas variables, posteriormente se busca un nivel de similitud entre las estructuras caracterizadas y se traduce en un índice que representa un nivel de similitud estructural entre los países que lo comparten. De tal forma se agrupan los países que presumiblemente cuentan con características similares y se espera ver efecto de esta estructura en el desempeño promedio de dichos países, en cuanto al crecimiento se refiere.

Por ejemplo una combinación de factores podría ser un porcentaje de gasto de gobierno del 39 % con respecto al PIB, combinado con un nivel de cobertura educativa del 53 % y una participación del 19 % de inversión extranjera directa con respecto al PIB, datos correspondientes a Chile[1],esta combinación de factores se plantea genera una dinámica dentro de la economía de tal forma que acota los efectos de los determinantes comúnmente utilizados para explicar el crecimiento económico y sería diferente al efecto provocado por una combinación 35 %, 49 % y 9 % para las mismas variables, valores de su país vecino Perú. De esta forma se le da importancia a la combinación de los factores dimensionados al tamaño de dicha economía, y no a los valores individuales de dichas variables. Así se plantea que pueden existir dos economías muy diferentes en términos absolutos pero similares en una combinación de términos relativos.

La estrategia que se seguirá es correr una regresión de corte transversal tal que incluya los determinantes del desarrollo comúnmente utilizados en los estudios de crecimiento económico para posteriormente incluir información de la dinámica económica, demográfica y tecnológica por medio del Índice de Agrupamiento Estructural (IAE). Lo que se busca es tener evidencia de que el IAE proporciona información adicional al modelo, de forma eficaz, aumentando el nivel de ajuste del modelo, y eficiente, ingresando variables que resultan en una mejora del modelo sin perder drásticamente los grados de libertad.

1.2.1. Construcción del conjunto estructural.

El conjunto de variables informativas o conjunto estructural se construye con la finalidad de aportar información de diversos rubros. Así se seleccionan un conjunto de variables del rubro económico, tecnológico, demográfico y geográfico y se procede a verificar que estén en términos proporcionales, las únicas variables que no tienen proporción respecto al propio país, en el ejercicio expuesto, son las de población y extensión territorial, estas se ingresarán como proporcionales al conjunto de observaciones.

La información obtenida está relacionada con un período de tiempo, por lo tanto cada variable es tratada para representar un valor promedio del período, esto porque se trata de extraer una aproximación del escenario más comúnmente observado del país, así tenemos:

$$z_{ij} = \sum_{t=1}^{T} z_{ijt} \cdot \frac{1}{T}$$

[1]Datos promedio para el periodo 1971-2014

Donde T es el número de periodos, z_{ijt} es el valor de la serie i para el país j en el periodo t y z_{ij} es el valor promedio de la serie i para el país j.

El siguiente punto a abordar es la distancia, debido a que se utilizará un método que minimiza la distancia hacia los centroides, es necesario tratar las variables para que sean comparables ya que no todas las variables presentan el mismo nivel y rango. Con esta finalidad definiremos:

$$w_{ij} = z_{ij}/max(z_i)$$

Donde z_{ij} es el valor de la serie i para el país j; z_i es el vector de valores de la serie i y w_{ij} es el valor normalizado entre cero y uno. Este tratamiento permitirá apreciar el efecto del índice cuando se le da la misma importancia a las distancias de las distintas variables. Una vez que se aplica el proceso al total de los datos se obtiene la matriz W que es a la que se aplicará el proceso de agrupamiento con el fin de detectar la similitud estructural entre los países.

1.2.2. Proceso de Agrupamiento.

El proceso de agrupamiento se realiza de acuerdo al algoritmo planteado por Hartigan y Wong (1979) denominado K-Medias. Este es un algoritmo de agrupamiento no supervisado, en el cual se tienen una serie de elementos ubicados en un espacio n-dimencional que se desea sean asignados a una serie de K grupos. Por cada uno de estos K grupos se definirá un centro y el algoritmo iterará en busca de minimizar la distancia entre el centro de grupo y los elementos asignados a este grupo. Este proceso de minimización intragrupal resulta de forma simultánea en la dispersión de los centros a lo largo del espacio. Es importante recalcar el hecho de que no es algoritmo supervisado por lo que los mismos datos son los que guían el proceso de agrupamiento.

En nuestro caso la posición de cada país la define el vector w_i de dimensión j, de tal forma que se posicionará cada país en un hiperplano j-dimensional. De esta forma se generarán centroides que agruparán a los países con las posiciones más cercanas a estos. Un punto importante es la dispersión de estos centros, ya que permite reflejar la disimilitud de los elementos pertenecientes a dos grupos diferentes mientras que acentúa la similitud entre elementos del mismo grupo.

Una gran ventaja al utilizar este método es que el número de grados de libertad que se perderán al agregar información adicional está únicamente ligado a K, el número de grupos, y no al total de variables utilizadas en el cálculo del agrupamiento. Debido a lo anterior resulta relevante la selección de K.

1.2.3. Estimación del Modelo.

El modelo a estimar es de la forma:

$$g_i = x_i + z_i + \epsilon_i$$

donde $\epsilon_i \sim NID(0, \sigma^2)$. Con g_i que representa el crecimiento per capita, x_i es el conjunto de variables explicativas y z_i es la variable categórica de agrupamiento. El conjunto de variables explicativas x_i está conformado por los determinantes utilizados comúnmente en los

estudios de crecimiento, específicamente es muy similar al utilizado por Basturk(2012) , con la diferencia que se agrega la conocida variable de convergencia.

El conjunto de variables z_i podrá tomar valores de agrupamiento procedentes de la agrupación geográfica clásica o de el proceso endógeno de agrupación. Las variables de agrupamiento del proceso endógeno se toman una vez que han sido asignados los países a cada uno de los K grupos, tomando esta asignación como una variable categórica de agrupación, representando de forma abstracta las similitudes de las variables de estructura.

Una vez estimado el modelo se busca evidencia de la mejoras generadas por la inclusión de las variables de agrupamiento.

1.3. Datos

La información utilizada en el estudio comprende datos de 104 países en un periodo que abarca de 1971 a 2014, los países y el periodo se tomaron basándose en la disponibilidad de la información, seleccionando aquellas series que contaran con al menos 80 % de la información disponible por país. Estos países abarcan 7 regiones según la clasificación del Banco Mundial[2], misma clasificación que es utilizada más adelante como criterio de agrupamiento.

Las fuentes de información utilizadas son la base de datos del Banco Mundial [25], particularmente la serie 'Indicadores del Desarrollo Mundial' (WDI por sus siglas en inglés) y la base de datos Penn World Table (PWT)[21]. Las series utilizadas se pueden dividir en dos grandes grupos, primero las que están relacionadas con la regresión, llamadas explicativas y segundo las que están relacionadas con el proceso de agrupamiento, llamadas informativas o estructurales.

Las variables utilizadas en la regresión son las que comúnmente se encuentran en la literatura del estudio del crecimiento, en primer lugar se tiene la variable de interés, en este caso el crecimiento, calculado con el PIB real ajustado por la paridad de poder compra en base a dólares americanos del 2011; como variables independientes se toman el índice de capital humano basado en la razón de los años promedio de estudio (Barro and Lee 2013)[4] y una tasa de retorno estimada (Psacharopoulos,1994)[50], el gasto de gobierno ajustado por la paridad de poder de compra y el precio relativo a la inversión tomando como base el nivel de precio de la inversión de EUA en el año 2011, todas las anteriores disponibles en la PWT. La inflación, el crecimiento anual de la población, el comercio como porcentaje del PIB, la formación bruta de capital como porcentaje del PIB y las distorsiones del tipo de cambio fueron tomadas de la base de datos de WDI.

Por otra parte la información utilizada para el proceso de agrupamiento se compone por 21 variables informativas relacionadas al rubro económico, tecnológico, demográfico y geográfico. Las variables del rubro económico tratan de reflejar la dinámica de la economía del país en base a las participaciones que tienen diversos factores en el PIB nacional y en las características relacionadas a su fuerza laboral.

Por su parte los indicadores tecnológicos tratan de reflejar la asimilación de tecnología en la vida diaria, en el sector productivo y la relación que se tiene con la innovación.

Los indicadores demográficos tratan de reflejar la dinámica socio-demográfica en la que se ve inmerso un individuo y trata de capturar las condiciones asociadas a la calidad de vida.

[2]En la versión del año 2015, http://data.worldbank.org/

Cuadro 1.1: Indicadores Estructurales

Indicadores Económicos
Gasto general de gobierno (% del PIB)
Desempleo total (% de la población activa total)
Fuerza laboral femenina (% de la fuerza laboral total)
Inversión extranjera directa(% del PIB)
Rentas totales de recursos naturales (% del PIB)
Suelo agrícola (% de la superficie terrestre)
Exportaciones de bienes y servicios (% del PIB)
Formación bruta de capital (% del PIB)
Gasto en consumo final de los hogares, etc. (% del PIB)
Importaciones de bienes y servicios (% del PIB)
Indicadores Tecnológicos
Suscripciones de telefonía celular móvil (por cada 100 personas)
Artículos de revistas científicas y técnicas (% de la población total)
Emisiones de CO2 (kg por PPC $ del PIB)
Emisiones de CO2 (toneladas métricas per cápita)
Indicadores Demográficos
Matrícula escolar secundaria (% bruto)
Población rural (% de la población total)
Tasa bruta de mortalidad (por cada 1000 habitantes)
Población total
Inmunización TDT (% de niños de 12 a 23 meses)
Tasa de natalidad bruta (por 1000 personas)
Indicador Geográfico
Extención territorial (kilómetros cuadrados)

Por último está el indicador geográfico que trata de capturar los niveles de complejidad administrativa, este es también el caso de la variable de población total dentro de los indicadores demográficos. Todas las variables informativas fueron recolectadas del WDI y se muestran en el cuadro 1.1.

1.4. Resultados

Los resultados obtenidos se pueden dividir en dos partes, los referentes al proceso de agrupamiento y los resultados de la regresión. Si bien los resultados del agrupamiento cobran su mayor relevancia al aplicarlos a la regresión, es interesante resaltar algunos de los resultados del agrupamiento en sí, ya que en cierto modo validan los resultados obtenidos.

1.4.1. Del proceso de Agrupamiento

El proceso de agrupamiento se aplicó a un total de 104 países, que corresponden a 7 regiones geográficas, según los cataloga el Banco Mundial en su base de datos.

El primer paso en el proceso de agrupación es seleccionar el número de centros (K), para esto se utilizó el criterio del "codo". Bajo este criterio la selección se hace graficando la suma de la variabilidad intragrupal (WSS por sus siglas en inglés) respecto a K y se selecciona el número de centros que corresponden a un cambio marcado en la pendiente de la curva, es decir, donde se forma un "codo". La gráfica resultante se muestra en la figura 1.1, donde el primer punto corresponde a la variabilidad intragrupal en el caso que hubiera un solo grupo, es decir, el caso en el que no hay agrupamiento, el segundo punto a la variabilidad intragrupal con dos grupos y así sucesivamente. Con la finalidad de facilitar la visualización de los cambios en la variabilidad intragrupal al incrementar el número de centros se muestra también la gráfica de las diferencias de las variabilidades (figura 1.2).

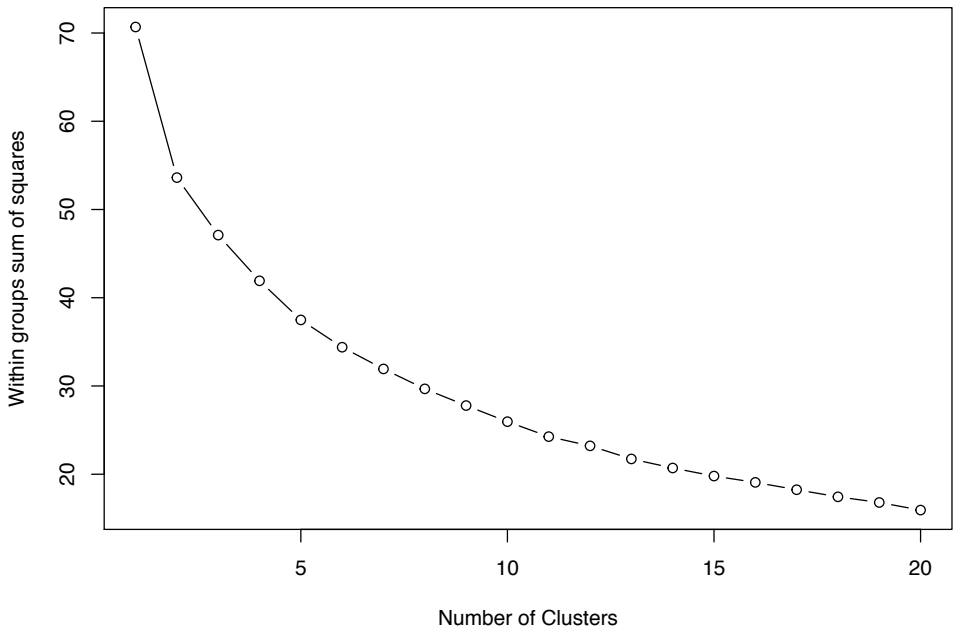


Figura 1.1: Variabilidad intragrupal por número de centros.

En la figura 1.2 se observa la formación de dos "codos", sin embargo el primer "codo" no se considera puesto que se refiere al cambio de la variabilidad entre usar y no usar agrupamiento y lo que se busca es que dado un proceso de agrupamiento se obtenga la mejor relación entre el número de centros y la homogeneidad del grupo. Así según este criterio se toma el segundo "codo" el cuál indica que el número de centros K es seis.

En la tabla 1.2 se muestra la clasificación resultante del proceso de agrupamiento con seis centros, donde cada columna pertenece a un grupo.

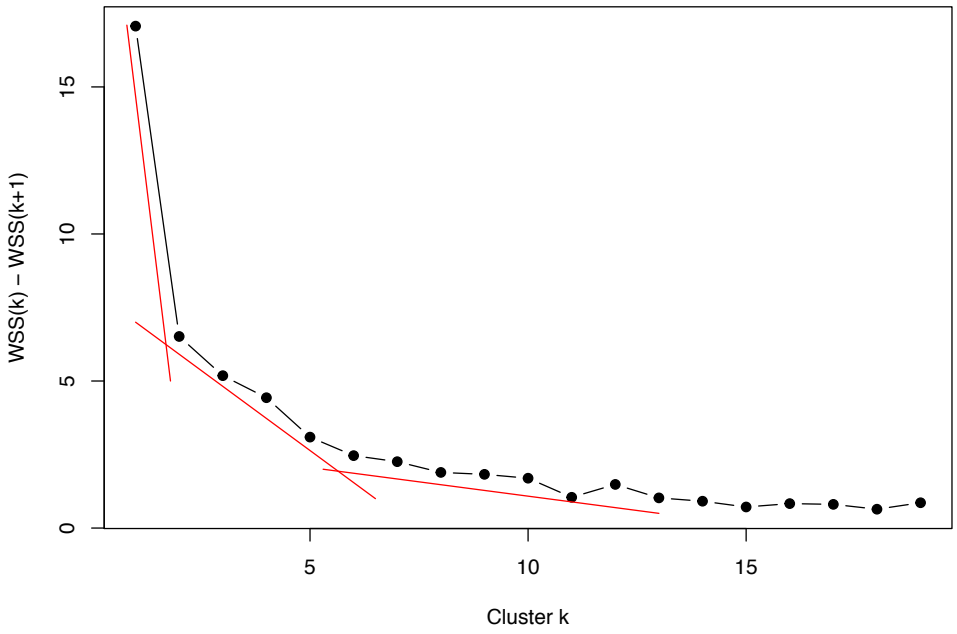


Figura 1.2: Diferencia de la variabilidad intragrupal.

Por otro lado la tabla 1.3 muestra el número de elementos en cada agrupación, la región más observada dentro de la agrupación y el porcentaje que representa esta última con respecto al total de los elementos en la agrupación.

De entrada se puede observar en la tabla 1.3 que el número de elementos por grupo en general no está balanceado, sin embargo es notoria la presencia de tres grandes grupos: el A, el B y el F, que concentran al 79 % de los países en la muestra y que además están relativamente balanceados entre sí. El resto de los grupos parecen identificar composiciones estructurales más específicas, llegando al caso extremo de Hong Kong siendo el único elemento del grupo G.

Desde un punto de vista intragrupal resulta muy interesante la composición del Grupo A, que en su mayoría agrupa países del África Sub Sahariana, estos resultados van en línea con lo expuesto por Masanjala y Papageorgiou que encuentran una dinámica diferente en los determinantes de crecimiento para el África Subsahariana respecto a los que se toman de forma global, en ese trabajo ya encuentran una diferenciación respecto al grueso de los países por medio de los efectos en los determinantes, mientras que el presente estudio identifica esta diferenciación de forma previa, desde un punto de vista enfocado en las causas y no en los efectos.

Por otra parte el Grupo F concentra 17 de los 19 países pertenecientes a Europa y Asia Central, que confirma una agrupación que podríamos llamar 'natural' dado el nivel de integración entre los países pertenecientes o cercanos al bloque 'Euro', solo dejando fuera a Albania y Turquía que se concentran en el grupo con predominancia de países Latinoamericanos, que

Cuadro 1.2: Resultado de clustering con 6 centros.

Grupo A	Grupo B	Grupo C	Grupo D	Grupo E	Grupo F
Angola	Australia	Argentina	Albania	United Arab Emirates	Hong Kong
Burundi	Brazil	Austria	Belize	Bahrain	
Benin	Canada	Bulgaria	Bolivia	Congo	
Burkina Faso	China	Barbados	Botswana	Algeria	
Bangladesh	United States	Chile	Colombia	Gabon	
Central African Republic		Cyprus	Costa Rica	Iran	
Côte d'Ivoire		Germany	Dominican Republic	Iraq	
Cameroon		Denmark	Ecuador	Qatar	
Ghana		Spain	Egypt	Saudi Arabia	
The Gambia		Finland	Guatemala	Trinidad and Tobago	
Haiti		France	Honduras	Venezuela	
Kenya		United Kingdom	Indonesia		
Cambodia		Greece	India		
Liberia		Hungary	Jamaica		
Lesotho		Ireland	Jordan		
Madagascar		Iceland	Sri Lanka		
Mali		Israel	Morocco		
Malawi		Italy	Mexico		
Niger		Japan	Mauritius		
Nigeria		Korea	Malaysia		
Nepal		Malta	Nicaragua		
Rwanda		Mongolia	Pakistan		
Senegal		Netherlands	Peru		
Sierra Leone		Norway	Philippines		
Togo		Portugal	Paraguay		
Tanzania		Uruguay	El Salvador		
Uganda		South Africa	Syrian Arab Republic		
Zambia			Thailand		
Zimbabwe			Tunisia		
			Turkey		
			Vietnam		

Fuente: Elaboración propia.

parecen haber intercambiado papeles con Chile y Uruguay.

Con la finalidad de hacer evidentes las similitudes de las variables estructurales entre países se graficaron las 21 variables de un país contra las 21 variables de otro en su forma normalizada, esto entre países del mismo grupo. Retomando el caso de Turquía, al graficarlo contra uno de los países Latinoamericanos de su grupo (México) se hacen evidentes las similitudes como se muestra en la figura 1.4.1. De la misma forma se observa con el caso de Uruguay que fue incluido en el grupo Euro que si bien no hay una correspondencia perfecta de puntos, sí denota un patrón similar.

Cabe recalcar que la clasificación está basada en condiciones de los países que no tienen en cuenta el desempeño de dichas naciones de forma directa (como el PIB), sin embargo ya se encuentran coincidencias no del todo sorprendentes en cuanto a desempeño, como la inclusión de Haití en el Grupo *A* dominado por países Africanos.

1.4.2. De la regresión

El modelo base comprende las variables mostradas en la Tabla 1.4, que son aquellas que suelen presentarse en los estudios de crecimiento y que han sido ampliamente analizadas. Estas variables cuentan con un amplio consenso respecto a la relevancia que tienen en el crecimiento económico razón por la cual fueron elegidas como modelo de referencia y con

Cuadro 1.3: Resumen del Agrupamiento

Grupo	Elem.	Moda	Obs. Moda	%Moda
Grupo *A*	29	Africa Sub-Sahariana	25	0.86
Grupo *B*	5	Norte América	2	0.40
Grupo *C*	27	Europa y Asia Central	17	0.63
Grupo *D*	31	América Latina y el Caribe	14	0.45
Grupo *E*	11	Medio Oriente y África del Norte	7	0.64
Grupo *F*	1	Asia y Pacífico	1	1

Fuente: Elaboración propia.

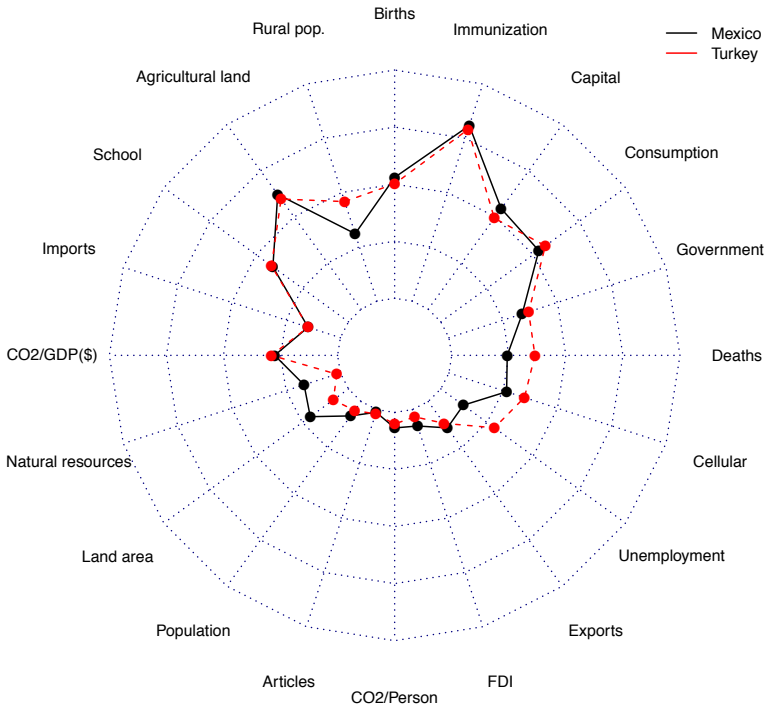


Figura 1.3: Similitud estructural entre elementos del grupo *D*: México-Turquía, datos del Banco Mundial (WDI)

fines de comparación.

En la Tabla 1.5 se muestran los resultados del modelo de mínimos cuadrados en tres

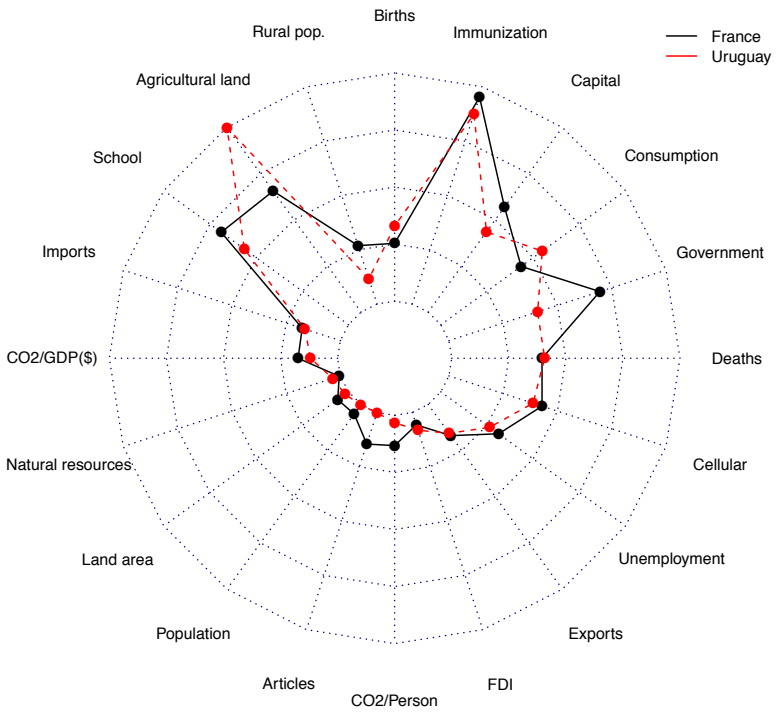


Figura 1.4: Similitud estructural entre elementos del grupo C: Francia-Uruguay, datos del Banco Mundial (WDI)

versiones, la primera (1) no usa variable alguna para diferenciar entre grupos de países, esta es la estimación de referencia. Las dos regresiones siguientes corresponden al modelo con variables regionales (2) y al modelo con dummys estructurales (3).

Cuadro 1.4: Variables utilizadas en el modelo MCO.

Variable	Descripción
Growth	Crecimiento del PIB, diferencia de los logaritmos del PIB de 2014 y 1971.
PIB70	PIB en 1970
Human_Cap70	Índice de Capital Humano Inicial (1970)
Gov_Share	Gasto de gobierno promedio 1971-2014
Inflation	Inflación Promedio (1971-2014)
Investment	Promedio de la inversión (1971-2014)
Inv_Price	Precio de la Inversión Promedio (1971-2014)
Pop_growth	Promedio del Crecimiento de la población (1971-2014)
Trade	Promedio del comercio como porcentaje del PIB (1971-2014)
RER_Dist	Distorsiones del tipo de cambio real promedio (1971-2014)

Como ya se mencionó con anterioridad el objetivo de este trabajo no es analizar en detalle los efectos de los determinantes que se utilizan en el estudio del crecimiento, sino el encontrar la forma de integrar al modelo información adicional que abone en el desempeño del modelo en su conjunto. Dicho lo anterior podemos ver que el modelo que integra las variables regionales tiene un mejor ajuste que la versión de referencia, por lo que podemos decir que el agregar las variables regionales si abona al desempeño del modelo. Por otra parte al interpretar las variables regionales podemos observar que la región *Este de Asia y Pacífico* es significativamente diferente a la región base, que en este caso es *Oriente Medio y Norte de África* cuestión que era de esperarse. Lo que si es de llamar la atención es que apenas se diferencia respecto a la región *Latinoamérica y Caribe* y la parte Sur de Asia. Mientras que la diferencia con respecto a *Europa y Asia Central* es significativa a un nivel de significancia del 5 %. Sin embargo, no se tiene evidencia una diferencia significativa respecto a la región de *Norteamérica* resultado que no se esperaría. Por último no existe evidencia de una diferencia significativa respecto al África Sub-Sahariana, que en este caso si podría deberse a características compartidas entre las regiones.

Por otra parte la regresión (3) que incluye las variables *IAE* muestra todas las variables de agrupamiento significativas, con excepción de la agrupación *E*, además el estadístico de ajuste $R^2adj.$ es el más alto entre los tres modelos, de hecho más del doble que el modelo base sin agrupación y 52 % mejor que el modelo con agrupamiento regional. Estos resultados plantean de inicio que las características estructurales que definen la pertenencia a un determinado grupo tiene un efecto en el crecimiento promedio que presentará dicho país.

Ahondando en los resultados de este modelo es conveniente resaltar la agrupación de referencia, que en este caso es la agrupación *A*, está constituida en su mayor parte por países del África Subsahariana. De tal forma que podemos interpretar que los países del grupo *A* tienen un crecimiento significativamente menor a los demás grupos, excepto a los pertenecientes al grupo *E*, que está formado mayoritariamente por países del *Norte de África y Medio Oriente*.

Con la finalidad de evaluar que tan robustos son los resultados de la agrupación realizada

Cuadro 1.5: Regresión MCO para los 104 países.

	(1)	(2)	(3)
PIB70	−0.0001	0.00002	0.00001
Human_Cap70	0.357*	−0.111	−0.287
Gov_Share	0.304	1.776	0.043
Inflation	−0.031	−0.018	−0.061
Investment	4.916***	3.612**	4.316***
Inv_Price	0.236	0.549	0.444
Pop_growth	0.565	0.060	−0.091
Trade	−0.009	0.158	0.116
RER_Dist	−0.322	−0.287	−0.687**
East Asia & Pacific		0.915***	
Europe & Central Asia		0.631**	
Latin America & Caribbean		0.444*	
North America		0.527	
South Asia		0.604*	
Sub-Saharan Africa		−0.198	
Group B			1.294***
Group C			1.279***
Group D			0.920***
Group E			−0.138
Group F			1.335**
Constant	−0.869**	−0.583	−0.169
Observations	104	104	104
R^2	0.275	0.416	0.552
Adjusted R^2	0.206	0.316	0.482
Note:			*p<0.1; **p<0.05; ***p<0.01

con el índice estructural se corrieron tres modelos adicionales que se muestran en la Tabla 1.6. El modelo (1) es el modelo de referencia que también se muestra en la Tabla 1.5, en el modelo (2) se remueve la variable Investment, en el modelo (3) se remueve la variable RER_Dist y por último el modelo (4) conserva solo aquellos covariantes del modelo econométrico de referencia que resultaron significativos en alguno de los modelos anteriores de la tabla.

Cuadro 1.6: Regresión MCO para pruebas de robustez.

	Dependent variable:			
	GrowthDif			
	(1)	(2)	(3)	(4)
PIB70	0.00001	0.00001	0.00001	
Human Cap70	-0.287	-0.371^{*}	-0.463^{**}	-0.282
Gov Share	0.043	1.099	0.494	
Inflation	-0.061	-0.099^{*}	-0.058	-0.071
Investment	4.316^{***}		4.260^{***}	4.188^{***}
Inv Price	0.444	0.308	0.051	
Pop growth	-0.091	0.377	-0.078	
Trade	0.116	0.129	0.127	
RER Dist	-0.687^{**}	-0.665^{**}		-0.506^{*}
Group D	0.920^{***}	1.048^{***}	0.923^{***}	0.898^{***}
Group B	1.294^{***}	1.694^{***}	1.246^{***}	1.217^{***}
Group C	1.279^{***}	1.594^{***}	1.154^{***}	1.249^{***}
Group E	-0.138	0.165	-0.122	-0.107
Group F	1.335^{**}	1.681^{**}	1.192^{*}	1.471^{***}
Constant	-0.169	0.571^{*}	-0.112	0.062
Observations	104	104	104	104
R^2	0.552	0.495	0.529	0.541
Adjusted R^2	0.482	0.422	0.461	0.497
Note:				*p<0.1; **p<0.05; ***p<0.01

La principal intención al remover las variables de los modelos es la de observar si existían cambios sobre la significancia de las variables de agrupación estructural. Como se puede observar no se pierde la significancia en ninguno de los casos, el único efecto es sobre la significancia del grupo F que llega a ser de 10 %. Por lo que se concluye que la información que proporcionan al modelo las variables de agrupamiento estructural no está fuertemente relacionada al resto de las variables explicativas.

1.5. Conclusiones

Basado en los resultados de Masanjala y Papageorgiou (2005) de que existen diferentes sendas de crecimiento dependiendo de las dinámicas dominantes en cada país e inspirado en

el enfoque de clasificación de Durlauf y Jhonson (1995) se propuso la construcción de un Índice de Agrupamiento Estructural (IAE) que concentra la información de 21 variables. Este índice clasifica estructuralmente a los países y permite utilizar dicha clasificación como variable categórica dentro de una regresión de corte clásico. La metodología propuesta permite tomar en cuenta información de los rubros económico, tecnológico, geográfico y demográfico para clasificar a los países de forma endógena. La utilización del IAE como instrumento de agrupación mejora en un 52 % el ajuste del modelo con respecto a una agrupación regional clásica y en un 133 % con respecto a un modelo sin agrupación. Estos resultados son obtenidos sin perder grados de libertad de forma proporcional al número de variables utilizadas, solo ingresando las K variables categóricas, en este caso siete.

Un resultado importante del proceso de agrupación es la generación endógena de bloques de países, en donde se generan a través de un proceso no supervisado agrupaciones naturales como la de países del África Subsahariana o la zona Euro. Pero que detecta también casos especiales como el mayor grado de similitud de Argentina y Uruguay con el grupo de países predominantemente europeos, la inclusión de Turquía con países Latinoamericanos o la inclusión de Haití con países del bloque africano. Teniendo dichos resultados respaldados por datos duros y libres de cualquier suposición o juicio a priori.

Otra ventaja de la metodología presentada es la flexibilidad que tiene para incluir variables, esta posibilidad permite en un futuro agregar índices de gobernabilidad, corrupción, desigualdad entre otros guiando la agrupación hacia temas de interés específico. Sin embargo, aunque se tiene la posibilidad de agregar variables adicionales concentradas en el índice, el efecto que tienen estas sobre el crecimiento sigue oculto en el índice. Sucede también con la importancia relativa de cada una de las variables respecto a la efectividad del índice que permanece desconocida, por lo que la metodología no puede utilizarse con fines exploratorios. Una importante extensión de este trabajo es la implementación de la metodología en un esquema de datos panel, lo cuál podría dar cuenta de la transformación estructural que pudieran presentar los países a lo largo del tiempo.

Capítulo 2

Desigualdad espacial metropolitana, una vista desde el espacio.

2.1. Introducción

El presente artículo aborda el tema de desigualdad espacial a nivel zonas metropolitanas. Tomando como base la relación descrita por Kuznets (1955), en la que relaciona el crecimiento y la desigualdad, donde el proceso de distribución de la riqueza se produce en dos etapas: una primera etapa en la que el incremento del ingreso se concentra solo en una parte de la población incrementando la desigualdad de ingreso. Y una segunda etapa en la que el sector con menor ingreso comienza a migrar hacia aquellas actividades que generan mayor ingreso, de tal forma que la desigualdad disminuye, generando un nuevo equilibrio en el que el ingreso es mayor en términos generales para la totalidad de los individuos. Esta idea fue a su vez extrapolada a nivel regional por Williamson (1965), bajo el concepto de desigualdad espacial. Bajo la premisa planteada por Williamson, de la misma manera que la desigualdad interpersonal en el ingreso se incrementa en las primeras etapas de desarrollo, para llegar a un máximo y eventualmente decrecer, las regiones dentro de los países presentarían un proceso similar. Donde la selección de las regiones con un desarrollo durante la primera etapa podría estar relacionado al descubrimiento y utilización de recursos naturales o ventajas logísticas, las cuales no suelen estar distribuidas de forma homogénea en los países. Sin embargo, un proceso migratorio o políticas redistributivas iniciarían un proceso de convergencia entre las regiones, de tal forma que la desigualdad decrecería.

La premisa es la siguiente, dada una zona metropolitana constituida por un conjunto de municipios donde uno de sus municipios funge como municipio principal, se esperaría que ante un proceso de crecimiento económico, el municipio principal concentrara inicialmente el ingreso generando un escenario de desigualdad metropolitana, para después iniciar un proceso de distribución del ingreso hacia los municipios secundarios, de esta forma se observaría un decremento en la dispersión de los datos asociados al ingreso entre los municipios. Si este proceso realmente ocurre entre los municipios de las zonas metropolitanas observaríamos el famoso patrón de U invertida planteado por Kuznets.

Sin embargo al llevar a nivel de municipios el estudio se presenta el problema de disponibilidad de los datos, es debido a esto que se utiliza una proxy del ingreso, este proxy se genera a través de las luces captadas desde el espacio y traducidas a un índice de luminosidad

espacial, tal como lo propuso Henderson (2012), donde plantea la relación entre el producto interno bruto y la luz observada desde el espacio. De esta forma se toma este índice para analizar la distribución del producto/ingreso de los municipios pertenecientes a cada zona metropolitana.

Al realizar el análisis se encontró que el efecto del ingreso en la desigualdad es consistente para todos los municipios de las zonas metropolitanas analizadas. Para todas presenta un signo negativo, es decir, que a mayor ingreso menor desigualdad espacial entre los municipios metropolitanos. Sin embargo, al controlar regionalmente, se encontró que no todas las zonas metropolitanas presentan el mismo efecto promedio en la desigualdad. De esta forma se muestra que por lo menos existen dos grupos diferentes en cuanto al efecto del ingreso en la desigualdad se refiere.

El contenido del artículo se detalla a continuación. El resto de la sección aborda principalmente los trabajos realizados sobre desigualdad espacial en México y la utilización de indicadores de desigualdad basados en datos energéticos. La segunda sección presenta la metodología utilizada para calcular el coeficiente de variación ponderado, así como la definición del modelo econométrico. La tercera sección muestra la fuente y la forma de extraer los datos utilizados en el estudio. La cuarta presenta los resultados bajo las dos especificaciones de modelo. Finalmente, la quinta cierra con las conclusiones.

2.1.1. Desigualdad espacial en México y el mundo.

En cuanto al estudio de las desigualdades regionales Barrios y Strobl (2009) analizan el vínculo entre las desigualdades regionales y el PIB per cápita a nivel país en busca de una evolución de las desigualdades regionales que describa una forma de U invertida, esto para una serie de países europeos. Según su análisis se encuentra que si se presenta una U invertida cuando se relacionan la desigualdad y el PIB per cápita para estos países y resulta interesante el hecho de que aquellos países que entraron posteriormente a la Unión Europea (en el año 2004) presentaron un proceso rápido de *catching-up* generando un incremento en las desigualdades regionales a medida que aumenta el nivel relativo del PIB nacional per cápita, para luego disminuir una vez que se alcanza un cierto nivel relativo del PIB nacional per cápita.

Este trabajo toma base los trabajos de Lessmann relacionados al estudio de la desigualdad espacial (Lessmann, 2014 y Lessmann y Seidel, 2017) donde busca evidencia de una relación en forma de U invertida entre la desigualdad espacial y el desarrollo económico. En el primer estudio encuentra evidencia a favor de la existencia de una U invertida para 56 países. Estos países se estudian a nivel subnacional, determinando así un nivel de desigualdad por país, para luego relacionarlo con la PIB per cápita a nivel nacional y buscando evidencia de una relación cuadrática entre ambos. En el segundo estudio amplía la muestra a 180 países al utilizar una estimación de los ingresos regionales basada en datos de luz nocturna satelital. De esta forma les es posible investigar las diferencias de ingresos regionales en los países en desarrollo que carecen de datos oficiales sobre ingresos. Al analizar los datos obtenidos encontraron que alrededor 67-70 % de todos los países experimentan una convergencia sigma y se encontró evidencia de una relación en forma de N entre desarrollo y desigualdad regional, donde países en niveles avanzados de desarrollo presentaban nuevamente incrementos en la desigualdad al incrementarse el PIB per cápita. Así fue posible capturar los diferentes estados

de desarrollo y la entrada nuevamente en el ciclo de generación y redistribución de los niveles de ingreso.

Respecto a los estudios para México ya trabajos como el de Sánchez 2006 toma como punto de partida el planteamiento de la curva de Kuznets para explorar el comportamiento de variables asociadas a crecimiento y desigualdad. En su estudio encuentran una correlación débil entre el coeficiente de Gini y el PIB por habitante para el periodo 1950 a 2004 e identifican las fases de la curva de Kuznets para dicho periodo. Además dentro de las reflexiones finales apuntan la importancia de acciones regionales a nivel estatal, señalando a Chiapas, Campeche, Oaxaca, Yucatán y Guerrero por ejemplo. Promoviendo el desarrollo regional con el objetivo de reducir las brechas entre los distintos niveles territoriales del país.

Respecto a las diferencias regionales basadas en análisis espacial Vilalta (2003) realizó un estudio en el que se encuentran dos procesos diferenciados temporalmente, uno de convergencia (1970 - 1980) y uno de divergencia (1970 - 1990) como ya otros trabajos habían planteado, sin embargo, al agregar la dimensión espacial al análisis encuentra la existencia de clubes de convergencia y concluye que es indispensable la inclusión de técnicas de análisis espacial en el estudio de las desigualdades regionales además de la recomendación de realizar estudios a niveles más detallados como ciudades.

En lo que respecta a la desigualdad regional a nivel municipal López (2008) hace un análisis para la Región Centro de México en el que mide la autocorrelación espacial de la productividad laboral bajo un enfoque espacial. Entre los resultados encuentra una fuerte dependencia espacial entre las municipalidades y el surgimiento de nuevas regiones o epicentros productivos con características propias. Esta dependencia espacial, según el estudio, ha sido alta y creció durante el periodo de estudio. Si bien el estudio se enfoca solo en la región centro de México, expone un proceso de polarización similar al conocido Norte-Sur, al aplicar un nivel más alto de desagregación, de estados a municipios. En este sentido se encontró evidencia de un proceso de desigualdad territorial que bien podría extrapolarse al resto del país.

Relacionado a la concentración económica espacial en las entidades federativas de México Sanén y Romero (2010) identifican las formación de clubes de convergencia y divergencia en los que resultan relevantes los factores geo-espaciales asociados. Los resultados sugieren la existencia de procesos de convergencia regional, lo cual en nuestra opinión tendría asociados procesos de desigualdad también diferenciados regionalmente. Además basado en los resultados de divergencia también encontrados en dicho trabajo se pone de manifiesto la posibilidad de diversos estados de "madurez" económica, subrayando la importancia de realizar un análisis a un mayor nivel de detalle.

Por su parte Campos y Monroy-Gómez-Franco (2016) analizan la relación entre el crecimiento económico y la pobreza para las 32 entidades federativas Mexicanas, encontrando que la elasticidad crecimiento-pobreza no es homogénea entre las entidades federativas. Y si bien, para 25 entidades se muestra una relación negativa entre crecimiento y pobreza, solo en nueve se observan reducciones más que proporcionales, en el que se sugiere un perfil geográfico asociado a dicha heterogeneidad en los efectos. Este efecto menos que proporcional se concentra en las entidades del sur, donde nuevamente nos hace pensar en las posibles disparidades de madurez económica o la posible falta de la estructura productiva y social necesaria para el aprovechamiento del crecimiento económico en la disminución de la pobreza.

Rodríguez y Cabrera (2017) identificaron una serie de municipios que clasificaron con pro-

cesos virtuosos, viciosos y resilentes de crecimiento. El estudio abarcó a los municipios de la república mexicana en el periodo 1999-2009 utilizando como proxy del desarrollo económico el valor agregado per capita disponible en los Censos Económicos y Censos de Población del Instituto Nacional de Estadística y Geografía. El trabajo mostró que los municipios mantienen un proceso de propagación del desarrollo económico, sin embargo este proceso no es homogéneo entre los municipios y se presenta de manera diferenciada, sobre todo en los municipios de las entidades del noroeste y centro-sureste del país. Este proceso se denomina vicioso cuando concentra el crecimiento en una municipalidad y virtuoso cuando se "contagia" a los municipios adyacentes. Además, al realizar un análisis espacio-temporal, dan cuenta que las similitudes disminuyen en el tiempo, sugiriendo un complejidad creciente en cuanto a dinámicas de crecimiento se refiere.

El trabajo de Rodriguez y Cabrera (2020) hace un estudio a nivel municipal para México en el que encuentra que existen efectos locales de interacción entre los municipios, que resultan en procesos de convergencia localmente diferenciados. De tal forma que encuentran procesos de convergencia con estados estacionarios y velocidades diferentes. Dicha heterogeneidad del proceso de convergencia no sigue una simple división Norte-Sur y muestra la necesidad de un análisis acotado regionalmente.

Márquez y Ramírez (2020) abordaron los efectos del crecimiento económico en la desigualdad y la pobreza para los municipios del estado de Nayarit encontrando que el avance en materia económica ha reducido tres tipos de pobreza de ingresos. Sin embargo, dicho efecto no es homogéneo y se produce un proceso de polarización entre los municipios con avances en temas de pobreza y desigualdad y lo que no, los primeros caracterizados enfocados de forma predominante a actividades relacionadas con el sector terciario. De esta forma se muestra nuevamente la existencia de procesos diferenciados de crecimiento y desigualdad diferenciados a niveles más profundos que el estatal.

En cuanto a la heterogeneidad presente en los procesos de crecimiento económico Valenzuela, Luna y Chapa (2021) encuentran que, al analizar a nivel zona metropolitana, se pueden encontrar cinco diferentes dinámicas de crecimiento económico o clubs de convergencia, los cuales dejan ver que si bien la división Norte-Sur está presente, esta simple división no alcanza a cubrir el mosaico diverso de dinámicas de crecimiento presentes en la República Mexicana. Si bien el estudio aborda el tema de convergencia, resultaría interesante analizar esa heterogeneidad a la luz de las dinámicas de desigualdad, y establecer si existe un proceso bien definido y unificado de la evolución de la desigualdad en México o este presenta una dinámica regionalizada.

El uso de variables asociadas al consumo energético para la elaboración de indicadores de desigualdad ya ha sido utilizado. Un ejemplo de ello es el trabajo de Duan y Chen (2018) en el que construyen una curva de Lorenz para cuantificar y revelar la desigualdad respecto al consumo de energía entre 121 países del mundo. En su trabajo encuentran una disminución en el índice de GINI de 2000 a 2013 que implica un consumo total de energía más igualitaria, sin embargo, no encuentran los mismos resultados al aislar el consumo de energía generado por fuentes renovables, donde se observa una gran desigualdad. Por su parte Lawrence, Liu y Yakovenko (2013) ven el consumo de energía generada por fuentes renovables como una oportunidad para disminuir la desigualdad en el consumo energético. Además mencionan la importancia de la heterogeneidad y desigualdad de las condiciones de vida de la población al evaluar costos y definir estrategias sociales, esto en contraste al modelo de agente represen-

tativo, donde toda la población se reduce a un solo agente "típico". Si bien en el artículo se enfocan en la mitigación del cambio climático nos parece importante tomar estas consideraciones en el contexto del crecimiento económico, en nuestro caso en el nivel de desigualdad que presentan las zonas metropolitanas en el país.

Por su parte Oswald, Owen y Steinberger (2020) plantean que la desigualdad en el consumo de energía afecta la distribución de los beneficios derivados del uso de la energía. De tal forma que la provisión de energía se convierte en un punto clave para el desarrollo. En su estudio analizan 86 países en los que miden el consumo de energía asociado a los bienes y servicios demandados, donde encuentran una alta granularidad y un rango más amplio de huellas energéticas que los estudios internacionales que solo operan en el nivel promedio nacional. De esta forma demuestran la importancia de los estudios intrapaís para entender a fondo la problemática en un contexto regional y en su caso llevándolo a un nivel de detalle aún mayor al analizar por grupos de ingresos. Según los resultados del estudio la desigualdad energética no es exclusiva de alguna categoría y se presenta por igual en el transporte, la salud o la educación, es por ello que consideramos importante definir el grado y dinámica de la desigualdad entre los municipios de las zonas metropolitanas ya que resulta un indicador de las posibilidades de desarrollo de una vida digna en las urbes.

2.2. Metodología

El presente trabajo toma como base la metodología utilizada en Lessmann (2014), donde se busca evidencia de una relación de U invertida entre la desigualdad espacial y el desarrollo. En el trabajo de Lessmann (2014) se calcula el coeficiente de variación ponderado (CVP) a nivel subnacional y se genera así una medida de desigualdad nacional para el conjunto de países definido. En el caso del artículo de referencia los elementos utilizados para el cálculo del CVP son el primer nivel de división nacional, el equivalente a estados en México. De esta forma se toma el PIB regional (estatal) y se calcula el CVP nacional. Posteriormente se busca explicar la medida de desigualdad nacional con el PIB per cápita nacional, por medio de una relación cuadrática de tal forma que, de existir, brinde evidencia de una relación de U invertida. A diferencia del artículo de referencia en este trabajo se busca evidencia de una relación de U invertida al relacionar la desigualdad metropolitana con PIB el metropolitano, esto para las zonas metropolitanas de la República Mexicana.

Ya que no existen datos anuales del PIB municipal para generar el CVP y calcular el PIB metropolitano se utilizan los datos de luz nocturna satelital para estimar un proxy de los ingresos regionales como se propone en Henderson et al. (2012), en este caso a nivel municipal. De esta forma se genera el coeficiente de variación ponderado que está definido como:

$$CVP_j = \frac{1}{\hat{y_j}} \left[\sum_{i=1}^{n} p_{ij} (\hat{y_j} - y_{ij})^2 \right]^{1/2}$$

donde:

$\bar{y_j}$: Es el promedio del ILpc de la zona metropolitana j.

y_{ij} : Es el ILpc del municipio i de la zona metropolitana j.

p_{ij} : Es la fracción de población del municipio i de la zona metropolitana j.

Una vez que se obtuvo el CVP por área metropolitana, al igual que en Lessmann (2014), se realizan dos estimaciones, una de corte transversal y una de datos panel con efectos fijos, definido el modelo como:

$$CVP_{i,t} = \alpha_i + \sum_{j=1}^{k} \beta_j Y_{i,t}^{j} + \sum_{m=1}^{q} \gamma_m X_{m,i,t} + \mu_t + \epsilon_{i,t}$$

donde:

$Y_{i,t}^{j}$**:** Es el ILpc de la zona metropolitana i en el tiempo t en grado j.

$X_{m,i,t}$**:** Son las Q variables de control.

Entre las variables de control que se incluyen siguiendo el artículo de referencia se encuentran aquellas que controlan la heterogeneidad territorial, que aplicadas al nivel de estudio de este trabajo son: (1) el logaritmo de el número de unidades municipales que integran la zona metropolitana, (2) el logaritmo de la extensión territorial y (3) el logaritmo de la razón del área entre el número de unidades municipales. Además en el trabajo de Lessmann utilizan variables de control para controlar la heterogeneidad entre características de los individuos, que incluyen variables como diversidad étnica, apertura comercial, porcentaje de la población en áreas urbanas y una dummy de federalismo. Estas variables no son incluidas en el presente estudio ya que tienen a nuestro parecer poca relevancia a nuestro nivel de agregación ya porque son características presumiblemente homogéneas dentro de un país o por definición cubiertas, como es el caso del porcentaje población en áreas urbanas, ya que al tratarse de zonas metropolitanas no se incluye población de áreas rurales. Sin embargo, si se incluyen una serie de dummys para captar las posibles diferencias regionales, como pueden ser las diferencias étnicas de las zonas metropolitanas.

2.3. Datos

Los datos utilizados para el presente estudio corresponden a la información de 56 zonas metropolitanas de la República Mexicana, compuestas por 401 municipios para el periodo de 1992 a 2015 con periodicidad anual. La definición y delimitación de zona metropolitana está tomada del documento "Delimitación de las zonas metropolitanas de México 2015"publicado por la Secretaría de Desarrollo Agrario, Territorial y Urbano (SEDATU), el Consejo Nacional de Población (CONAPO) y el Instituto Nacional de Estadística y Geografía (INEGI). Según este documento se contabilizan 74 zonas metropolitanas en México, sin embargo, solamente 56 de estas cuentan con dos o más municipios mexicanos en su conformación, los casos restantes corresponden a capitales estatales que por definición se contabilizan como zona metropolitana y zonas metropolitanas fronterizas que están formadas por más de una entidad geopolítica pero que solo una de estas corresponde a México. Una vez definidos los elementos que conforman a las zonas metropolitanas se construyó una variable proxy del PIB municipal basada en la relación que existe entre la luz nocturna satelital y la actividad económica (Sutton, Elvidge

y Ghosh, 2007). Siguiendo la metodología descrita en Henderson et al. (2012) se realizó la extracción de datos geolocalizados para la generación de un índice lumínico, el procedimiento es el siguiente: como primer paso se obtienen los datos lumínicos que son presentados como un conjunto de pixeles, donde cada pixel tiene asignando un valor numérico que corresponde al nivel de luz captado por el sistema satelital, además cada pixel está asociado a un área ($0{,}44m^2$ o $0{,}88m^2$ según la tecnología utilizada) debidamente geolocalizada. Como segundo paso se delimita la zona geográfica de interés y se suman los valores de los pixeles contenidos en dicha delimitación, por último se divide el total obtenido entre la población de la región delimitada para generar un Índice Lumínico per cápita (ILpc). Los datos para la construcción de este índice se obtuvieron por una parte del Centro Nacional de Datos Geofísicos (National Geophysical Data Center, 2010) que depende de la Oficina Nacional de Administración Oceánica y Atmosférica (NOAA, por sus siglas en inglés) de los Estados Unidos de América. Los datos capturados por medio de estos satélites corresponden a algún instante entre las 8:30 p.m. y las 10:00 p.m. (hora local) con lo que se genera un promedio anual para los años de 1992 a 2015. Estos datos fueron generados bajo dos tecnologías, la primera (OLS, Operational Linescan System) genera datos con pixeles que corresponden a un área aproximada de 0.88 km2 para los años 1992 a 2013. La segunda (VIIRS, Visible Infrared Imaging Radiometer Suite; Elvidge et al., 2017) genera datos por pixel que corresponden aproximadamente a 0.44 km2 y cubren los años 2013, 2014 y 2015. Bajo esta tecnología los datos disponibles para el 2013 y 2014 se presentan con frecuencia mensual por lo que se calculó el promedio anual de forma similar a lo realizado por Chen y Nordhaus (2019). Los datos relacionados a los límites geopolítico-espaciales se obtuvieron del mapa geoestadístico del INEGI, donde se obtuvieron los límites geopolíticos de los municipios correspondientes a las zonas metropolitanas bajo estudio. Por último la información poblacional para cada uno de los municipios se obtuvieron de los documentos censales e intercensales publicados también por el INEGI. Debido a que estos se presentan para los años 1990, 1995, 2005, 2010 y 2015, se interpolaron linealmente para obtener una estimación anual de la población municipal.

El uso de la luz nocturna satelital si bien ha resultado útil para la estimación de un proxy de la actividad económica, presenta limitaciones y puntos de debate. Como primer punto la estimación por este método de la actividad económica tiene un pobre desempeño en territorios con baja densidad poblacional donde se realizan actividades del sector primario como agricultura o pesca (Chen y Nordhaus, 2015b), por lo tanto la estimación para áreas rurales no suele ser tan precisa como para las zonas urbanas, sobre todo aquellos datos obtenidos mediante la tecnología OLS, que de hecho comprende 87 % de la muestra. Sin embargo, al centrar el estudio en áreas metropolitanas este efecto adverso se ve disminuido ya que la relación entre el índice lumínico y el PIB es fuerte a nivel zona metropolitana como lo muestran Chen y Nordhaus, 2019 en su estudio a nivel zona metropolitana (área estadística metropolitana según su nomenclatura) para los Estados Unidos de América, donde existen datos para comparar a nivel municipal este indicador. De hecho, concluyen que este indicador es más fuerte a nivel área metropolitana que a nivel estado. El segundo punto que suele prestarse a debate es la relación log-lineal entre la densidad de luz y los ingresos. Y plantea que una elasticidad luz-ingreso constante implica que la relación entre los niveles de luz y los niveles de ingresos depende del nivel de luz en sí. De tal forma que la diferencia absoluta de luminosidad entre dos regiones no implicaría la misma diferencia en los ingresos de otras dos regiones que son, en promedio, más ricas o más pobres. Por lo tanto, las medidas de desigualdad calculadas

directamente a partir de datos de densidad lumínica no siempre proporcionarían una medida adecuada de desigualdad en el ingreso regional (Lessmann , 2017). Debido a este fenómeno otros trabajos tratan de controlar el componente idiosincrático de los elementos bajo estudio por medio de un modelo de efectos fijos (LLamosas et al., 2018 y Rangel, 2019) o con un modelo de efectos aleatorios con variables de control (Lessmann , 2017), sin embargo esto no es necesario en nuestro caso, en primer lugar porque se analizan conjuntos de elementos que por definición son homogéneos, es decir, la definición de zona metropolitana implica una integración en funciones, actividades y área urbana de los municipios, manteniendo un alto grado de integración socioeconómica(SEDESOL,2018). Y, en segundo lugar, el uso del coeficiente de variación ponderado no toma en el efecto en niveles, si no mas bien se enfoca en la variabilidad alrededor de este. De esta manera observamos la variabilidad intragrupal de elementos idiosincráticamente integrados; en los que se busca evidencia de un efecto de *(*spillover), más que una comparación de las diferencias absolutas de los niveles entre los grupos (zonas metropolitanas). Así, se sigue la línea de trabajos como los de Mveyange (2015) y Alesina et al. (2016) en los que utilizan directamente las luces para estimar la desigualdad espacial en África, para el primero, y entre regiones etnolinguísticas, para el segundo.

Los estadísticos descriptivos correspondientes a las 56 zonas metropolitanas para el periodo de 1992 a 2015 se muestran en el Cuadro 2.1.

2.4. Resultados

Siguiendo el trabajo de Lessmann (2014) se calcularon los valores promedio de la variable de desigualdad para las 56 zonas metropolitanas bajo estudio. Una vez obtenido el CVP promedio se calculó el modelo lineal y cuadrático utilizando como variable explicativa la estimación del ingreso por medio del índice lumínico. Además, se incluyeron las variables de control territorial: número de municipios por zona metropolitana y la extensión territorial del área metropolitana. El modelo de referencia utilizaba adicionalmente, la razón entre la extensión territorial y el número de entidades, en nuestro caso municipios del área metropolitana, sin embargo se eliminó esta variable por presentar problemas de multicolinealidad en el modelo. Siguiendo esta especificación se calculó el modelo lineal sin variables de control (1), el modelo cuadrático sin variables de control (2), el modelo lineal con variables de control (3) y el cuadrático con variables de control (4); los resultados se muestran en el cuadro 2.2.

Respecto a los resultados del cuadro 2.2 se observa que los modelos con variables de control territorial presentan un ajuste mucho mayor que los modelos que no las incluyen. La variable de extensión territorial y el número de municipios resultan significativas para ambos modelos. De esta forma se plantea una relación directamente proporcional entre la extensión territorial y la desigualdad espacial metropolitana de tal forma que los efectos de *spillover* no alcanzan a extenderse posiblemente debido a un efecto de lejanía o a la complejidad logística asociada a esta. Por otra parte el número de municipios tiene una relación directamente proporcional a la desigualdad, por lo que esperaríamos ver una menor desigualdad en zonas metropolitanas con un menor número de municipios. Por otra parte los resultados referentes a la variable de interés indican dos resultados interesantes, en primer lugar no hay evidencia estadística de un proceso de U invertida para los municipios de las zonas metropolitanas de México. En segundo lugar, la variable proxy del ingreso presenta un signo negativo para los modelos

Cuadro 2.1: Estadísticos descriptivos de los municipios de las 58 zonas metropolitanas en los años 1992 a 2015.

	N	Mean	St. Dev.	Min	Max
Centronorte					
totLights	240	26,574.690	17,735.000	2,755	73,651
pop	240	567,770.200	442,992.200	95,542	1,768,193
lightsPc	240	0.053	0.014	0.022	0.100
Centrosur					
totLights	120	98,004.560	129,166.100	5,256	530,485
pop	120	4,398,121.000	7,324,128.000	100,991	20,892,724
lightsPc	120	0.047	0.019	0.016	0.082
Noreste					
totLights	168	53,218.850	43,259.970	11,196	223,197
pop	168	1,064,347.000	1,153,875.000	123,258	4,689,601
lightsPc	168	0.062	0.018	0.033	0.117
Noroeste					
totLights	120	26,100.430	19,503.620	3,893	69,386
pop	120	537,304.200	539,244.300	101,960	1,840,710
lightsPc	120	0.062	0.019	0.031	0.109
Oeste					
totLights	192	26,093.660	32,700.190	3,427	157,025
pop	192	785,643.500	1,269,834.000	113,102	4,887,383
lightsPc	192	0.048	0.017	0.017	0.090
Oriente					
totLights	360	22,964.710	20,804.010	1,729	107,203
pop	360	493,764.700	556,633.100	81,551	2,941,988
lightsPc	360	0.053	0.027	0.013	0.161
Sureste					
totLights	72	37,750.790	14,592.850	15,109	86,919
pop	72	704,043.100	220,979.500	240,403	1,143,041
lightsPc	72	0.054	0.013	0.032	0.107
Suroeste					
totLights	120	20,828.510	8,948.922	6,368	52,984
pop	120	477,759.900	250,189.600	137,342	886,975
lightsPc	120	0.053	0.033	0.020	0.298
Total					
totLights	1,392	34,989.980	49,825.620	1,729	530,485
pop	1,392	965,479.300	2,502,742.000	81,551	20,892,724
lightsPc	1,392	0.054	0.022	0.013	0.298

Cuadro 2.2: Regresión lineal CVP contra Índice Lumínico

	Dependent variable:			
	log(wcv)			
	(1)	(2)	(3)	(4)
logLightsPc	−0.107	2.472	0.235	−0.759
	(0.278)	(3.267)	(0.253)	(2.907)
logLightsPc2		0.426		−0.166
		(0.538)		(0.483)
log(NumMun)			0.438***	0.451***
			(0.109)	(0.116)
log(ExtTerr)			0.199**	0.197**
			(0.087)	(0.088)
Constant	−0.396	3.455	−1.501	−2.976
	(0.836)	(4.932)	(1.001)	(4.416)
Observations	58	58	58	58
R^2	0.003	0.014	0.308	0.310
Adjusted R^2	−0.015	−0.022	0.270	0.258
Note:			*p<0.1; **p<0.05; ***p<0.01	

(1), (3) y (4), siendo significativo en el modelo (1). De tal forma se plantea que a mayor ingreso se observa menor desigualdad espacial metropolitana. Este resultado implica que por lo menos para el periodo de 1992 a 2015 un incremento en el ingreso resulta directamente en una disminución en las diferencias de los municipios pertenecientes a la zona metropolitana. Esto puede significar que se está en presencia de economías en fases de madurez, las cuales presentaron un aumento en la desigualdad ante un crecimiento en el ingreso en un periodo anterior a 1992 o que el proceso para las zonas metropolitanas es de transferencia directa respecto al incremento en el ingreso. Debido a que no se cuentan con datos anteriores a 1992 para esta variable no es posible validar cuál opción es la correcta, sin embargo, para trabajo futuro queda abierta la posibilidad de buscar una variable que refleje el ingreso a nivel municipio para fechas anteriores a 1992. Adicionalmente se calcularon los modelos de tercer y cuarto orden, sin embargo los coeficientes de la variable lumínica no resultan significativos en los modelos adicionales.

Con la intención de captar las posibles diferencias regionales en el efecto del ingreso sobre la desigualdad espacial en las zonas metropolitanas se incluyeron variables categóricas que identifican a las zonas metropolitanas en las regiones: Centronorte, Centrosur, Noreste, Noroeste, Oeste, Oriente, Sureste y Suroeste. Los resultados se muestran en el cuadro 2.3 donde se observa el modelo lineal (1) y el modelo cuadrático (2), ambos incluyendo variables de control territorial. Como se puede observar la componente cuadrática en el modelo (2) no resulta significativa por lo que no hay evidencia de un proceso de U invertida al diferenciar por regiones a las zonas metropolitanas. Por otra parte, en el modelo lineal, se observa una diferencia entre las zonas metropolitanas pertenecientes a la región Oeste respecto a la región Noreste de referencia. La variable dummy de esta región presenta un signo negativo por lo que se concluye que en promedio tiene una desigualdad menor al mantener sin cambio los efectos restantes.

Cuadro 2.3: Regresión lineal CVP contra Índice Lumínico con Dummy Regional

	Dependent variable:	
	log(wcv)	
	(1)	(2)
logLightsPc	0.321	−1.462
	(0.264)	(2.997)
logLightsPc2		−0.297
		(0.498)
log(NumMun)	0.425***	0.444***
	(0.138)	(0.142)
log(ExtTerr)	0.311**	0.311**
	(0.124)	(0.124)
RegionCentronorte	0.489	0.479
	(0.370)	(0.373)
RegionCentrosur	0.643	0.662
	(0.465)	(0.469)
RegionNoroeste	0.752*	0.768*
	(0.384)	(0.388)
RegionOeste	0.715*	0.724*
	(0.370)	(0.372)
RegionOriente	0.746*	0.762*
	(0.392)	(0.395)
RegionSureste	0.545	0.527
	(0.456)	(0.460)
RegionSuroeste	0.582	0.591
	(0.409)	(0.412)
Constant	−2.643**	−5.314
	(1.238)	(4.642)
Observations	58	58
R^2	0.385	0.390
Adjusted R^2	0.254	0.244
Note:	*p<0.1; **p<0.05; ***p<0.01	

2.5. Conclusiones

Al analizar el índice lumínico como proxy del ingreso para 58 de las 74 zonas metropolitanas de la República Mexicana se encontró que no existe evidencia estadística de un proceso de U invertida en la desigualdad espacial de los municipios pertenecientes a las zonas metropolitanas de la República Mexica. Además, se encontró que la variable proxy del ingreso presenta un signo negativo y significativo, de tal forma que a mayor ingreso se observa menor desigualdad espacial metropolitana. Este resultado implica que por lo menos para el periodo de 1992 a 2015 un incremento en el ingreso resulta directamente en una disminución en las diferencias de ingresos de los municipios pertenecientes a la zona metropolitana.

Por otro lado se encontró que no existe evidencia de un efecto del ingreso medido a través del ILpc sobre la desigualdad espacial entre los municipios debido a diferencias de nivel, ya que al dividirlos en nivel alto y bajo, la variable dummy que categoriza este fenómeno no resultó estadísticamente significativa.

Sin embargo, al analizar de forma regional si se presentan diferencias, ya que se encontró que 18 de las 58 zonas metropolitanas tienen en promedio un nivel de desigualdad espacial menor que el resto al controlar por los demás factores. Estas zonas metropolitanas pertenecen a las regiones Centronorte y Oeste de la República Mexicana.

En general se concluye que el efecto del ingreso sobre la desigualdad espacial tiene el mismo signo para todas las zonas metropolitanas del país. Presentando solo diferencias en cuanto a nivel para dos regiones del país. Queda como trabajo futuro el analizar estos datos por medio de técnicas que permitan analizar de forma más detallada los elementos permitiendo cambios en el coeficiente a nivel región o diferenciando por características idiosincráticas, con el fin de entender de forma mas detallada la dinámica del desarrollo de las comunidades urbanas en el país.

Capítulo 3

Clubes de convergencia metropolitana: un análisis a través del índice lumínico.

3.1. Introducción

Uno de los principales retos que enfrenta la economía mexicana, y que probablemente la pandemia del Covid-19 los magnifique, es la profunda desigualdad en los niveles de la actividad económica en las diversas entidades federativas. De acuerdo con Rodriguez-Oreggia (2005) y Chiquiar (2005) existen dos regiones en el país, una desarrollada representada por los estados del norte, y otra rezagada concentrada en la región sur, en las cuales la diferencia de desarrollo económico es persistente y hasta exacerbada por los procesos de liberación, creando una suerte de ganadores y perdedores. Estas diferencias no solo se ven reflejadas en nivel regional Norte-Sur sino que además se perciben dentro de las entidades federativas, a través de sus municipios y áreas metropolitanas. Mientras que en el agregado se genera un proceso de divergencia entre estos dos grandes grupos regionales, dentro de las regiones se produce de forma simultánea un proceso de convergencia entre aquellas entidades con la infraestructura apropiada para aprovechar cambios estructurales como el tratado de libre comercio, y generando un proceso similar, pero negativo, entre las entidades sin las características necesarias para asimilar en su favor políticas de este tipo. Es decir, la heterogeneidad de las diferentes entidades del país juega un rol primordial en el análisis de convergencia. Esta heterogeneidad, no solo está asociada a la falta de infraestructura sino que además se relaciona con la vocación predominante en la entidad, de tal manera que no se tendría que esperar convergencia entre una entidad con actividad predominantemente industrial y una con vocación turística.

Desde un punto de vista teórico neoclásico, plasmado en los trabajos seminales de Solow (1956) y Swan (1956), se plantea la posibilidad de un proceso de convergencia absoluta entre las economías donde las menos desarrolladas alcanzan a las más desarrolladas al experimentar las primeras una razón de crecimiento más grande que sus pares desarrollados a través del efecto de los rendimientos decrecientes de los factores. Esta posibilidad de convergencia inspiró trabajos empíricos como los de Barro y Sala-i-Martin (1991); Mankiw, Romer y Weil (1992) entre otros, en los que basados en sus hallazgos se propone un proceso de convergencia

condicional. En este tipo de convergencia el estado inicial juega un papel importante en la definición del estado estacionario, por lo que países con diferencias en las condiciones iniciales alcanzarán distintos estados estacionarios.

Dentro de la literatura existen diversos trabajos que abordan el tema de la convergencia regional en México tanto en la búsqueda de evidencia de la convergencia entre las entidades federativas (Esquivel, 1999; Rodriguez-Oreggia, 2002), como trabajos que buscan identificar el posible efecto de la liberación comercial en los procesos de divergencia regional en México (Cermeño, 2001; Esquivel y Messmacher, 2002; Chiquiar, 2005; Rodriguez-Oreggia, 2005; Carrion-I-Silvestre y German-Soto, 2007) encontrando evidencia de un cambio en la dinámica de convergencia de las entidades en los años posteriores a 1985 y 1994.

Si bien los trabajos antes mencionados estudian los aspectos de convergencia considerando la heterogeneidad de las diversas entidades federativas, estos trabajos plantean el análisis de manera agregada y tomando en cuenta las regiones de acuerdo con las divisiones territoriales. Es decir, solo toman en cuenta la heterogeneidad de las diversas entidades federativas. No obstante, esta diversidad no es exclusiva solo de las entidades, ya que las características de la dinámica económica de cada entidad es el resultado de la combinación de las singularidades y actividades económicas particulares de cada uno de sus municipios y zonas metropolitanas. De igual manera, las metodologías utilizadas en estos estudios solo permiten la convergencia entre las entidades o bien la convergencia de las entidades hacia una entidad líder bajo un marco de referencia de seguidor-líder.

Con la finalidad de considerar la diversidad intra-estado y de llevar el análisis de convergencia sin considerar las divisiones territoriales estatales, este trabajo estima la convergencia entre las diferentes zonas metropolitanas del país. En particular, siguiendo la metodología propuesta por Phillips y Sul (2009) se buscan los grupos o clubes de convergencia, que funcionan como una clasificación del tipo de zona metropolitana respecto a la dinámica de crecimiento y que puede utilizarse para buscar los determinantes del crecimiento de las zonas metropolitanas. A diferencia de estudios previos, bajo esta metodología se permite que se dé la convergencia heterogénea entre zonas metropolitanas. Por ejemplo, las zonas turísticas del país pueden converger a un estado estacionario diferente a la convergencia que se daría en las zonas metropolitanas industriales del país.

Así, este trabajo sigue la línea en favor de la clasificación de convergencia endógena, ya que la subdivisión geográfica estatal también se puede ver como una división *a priori*, que en el presente trabajo se evita. Esta división *a priori* no presenta problemas cuando no se cruzan fronteras geopolíticas, es decir que la zona metropolitana pertenece a un solo estado, cuando la entidad tiene un solo centro económico o cuando presenta varios centros económicos con el mismo comportamiento. Sin embargo este no es el caso para todas las entidades federativas de la República Mexicana ya que el 50 % de las entidades presentan comportamientos heterogéneos de crecimiento. De esta forma también se evitan clasificaciones basadas en la ubicación geográfica de las entidades (*i.e.* Zona Norte, Zona Centro, Bajío, etc.) y se buscan clasificaciones que, en caso de existir, se formen naturalmente de acuerdo con las características afines en los datos que las generan. De tal manera que al llevar a nivel de zonas metropolitanas el análisis, se conservan las diferencias que pueden existir entre las regiones. Además, estas no se ven afectadas por posibles divisiones territoriales que fraccionan artificialmente lo que en un ámbito socio-económico se ha desarrollado de forma conjunta y unificada.

Para llevar a cabo el análisis a nivel de zonas metropolitanas para la República Mexicana se utiliza el Índice Lumínico (*IL*) que funciona como proxy de medición de la actividad económica de dichas zonas. La utilización del *IL* se debe a la falta de datos de producción a nivel municipal por parte de las instancias oficiales como el Instituto Nacional de Geografía y Estadística (INEGI). Además, este índice presenta dos características sumamente atractivas: la versatilidad en cuanto a la delimitación geográfica al permitir el acceso a datos de las diversas zonas metropolitanas en todo el territorio nacional (Llamosas-Rosas et al., 2018); y la precisión en la medición de la actividad económica real ya que no excluye la actividad económica informal (Ghosh et al., 2009) y minimiza el efecto que pueden tener rentas extraordinarias como las petroleras en entidades como Campeche y Tabasco, las cuales podrían sesgar y/o distorsionar la medición de la actividad económica[1].

Finalmente, los resultados sugieren que en el 50 % de las entidades federativas la dinámica de la convergencia económica es heterogénea al interior, por lo que resulta relevante el estudio a nivel zonas metropolitanas. De igual manera, se encuentra que si bien existe una división entre un club de convergencia con elementos predominantemente en el norte del país, esta división es más compleja que una simple división Norte-Sur, ya que se encontraron cinco diferentes clubes o grupos de convergencia. Estos resultados contrastan con los trabajos de Rodriguez et. al (2016) y Mendoza-Velázquez et. al. (2019) en los cuales encuentran seis y dos clubes respectivamente, utilizando la misma metodología pero a un nivel estatal. Finalmente, se realizó un análisis de serie truncada, como es común en estudios de este tipo, en la cual se identificó un elemento divergente, la zona metropolitana del Valle de México. Este resultado es interesante ya que, bajo un esquema seguidor-lider[2] es posible considerar como líder a la economía más desarrollada, en este caso el Valle de México. Por lo tanto bajo este resultado sería poco probable observar un proceso de convergencia de alcance. Además resulta consistente con el rechazo de las pruebas de convergencia basadas en raíces unitarias utilizando a la Ciudad de México como líder (Ayala y Chapa, 2011), ya que se realizan contra un elemento con trayectoria divergente y que pudiera presentar condiciones históricamente diferentes a las demás regiones, aunque cabe señalar que esto no implica la posible existencia de un proceso de convergencia σ.

La organización del documento es de la siguiente manera. La sección uno aborda los trabajos realizados sobre el tema de convergencia en México. La sección dos presenta la metodología utilizada para encontrar los clubs de convergencia estocástica. La sección tres muestra lo referente a los datos utilizados para el estudio. La sección cuatro presenta tanto los resultados del agrupamiento como un análisis gráfico del mismo. Finalmente, la quinta sección cierra con las conclusiones.

[1]Los trabajos de Brida et al. (2013) y Rodriguez et al. (2016) realizan análisis de convergencia con y sin estos estados con la finalidad de evitar dicho sesgo.

[2]Entre los trabajos para México que utilizan este esquema se encuentran Carrion y German-Soto, 2007; Ayala y Chapa, 2011 y German-Soto y Salazar, 2016; en los que se utiliza al Valle de México y/o el promedio nacional, afectando en gran medida los resultados de la prueba para el primer caso y parcialmente al utilizar el promedio, sobre todo en periodos en los que la actividad económica se concentraba principalmente en la capital.

3.1.1. Estudios regionales de convergencia estocástica para México

Dentro de la literatura, existen diversos estudios que prueban la presencia de un proceso de convergencia entre las diversas regiones y entidades federativas de México. Por ejemplo Esquivel, 1999 realiza un estudio para el periodo 1940-1995 para las entidades federativas en el que encuentra convergencia β, es decir que las regiones más pobres crecen a tasas más altas que las ricas y convergencia σ referente a la dispersión de los datos, con la particularidad de que la velocidad de convergencia es mayor para el periodo anterior a 1960. De manera similar Cermeño 2001 muestra evidencia de convergencia condicional utilizando un panel dinámico sin regresores exógenos en el que se encuentra una desaceleración continua del crecimiento durante el periodo1970-1995, presentándose convergencia condicional bajo un proceso de decrecimiento del PIB estatal por persona. Por su parte Chiquiar (2005) realiza un estudio ampliando el periodo de estudio hasta el 2001 con la finalidad de observar el efecto que tuvo el Tratado de Libre comercio de América del Norte en el proceso de convergencia de las entidades federativas, encontrando que la ruptura en el proceso de convergencia situado en 1985 no se revierte y, más aún, se presenta un cambio estructural en el patrón de crecimiento al introducir nuevos flujos de inversión y comercio. Así mismo Rodriguez-Oreggia (2005) detecta un proceso de convergencia hasta 1985 y un proceso de divergencia a partir de ese año. Además, el autor realiza pruebas agregando factores para detectar convergencia condicional, concluyendo que factores como el capital humano o las rentas petroleras resultan significativos de forma alternada dependiendo del periodo en el que se encuentren. Rodriguez-Oreggia (2005) argumenta que bajo el periodo en el cual se que implementan medidas proteccionistas y/o populistas las rentas petroleras resultan positivamente significativas mientras que el capital humano no lo es, caso que se invierte durante el periodo de apertura comercial (1985-2000) en el que el capital humano toma relevancia. Estos resultados, de acuerdo al autor, pudieran deberse a que el talento dejó de consumirse de forma predominante en la burocracia y comenzó a ser productivo en el sector privado.

En síntesis, estos trabajos encuentran evidencia de la existencia de un proceso de convergencia hasta el año de 1985 seguido de un proceso de convergencia débil o divergencia a partir de ese año. Estos trabajos tienen en común que, debido a las metodologías utilizadas, operan bajo el supuesto homogeneidad de los parámetros de convergencia, por lo que se entiende que tras un choque con efectos diferenciados sobre las entidades se produzca esta pérdida de convergencia. Contrario a estos estudios, el trabajo de Carrion-I-Silvestre y German-Soto (2007) se puede ver como el primer intento que considera esta falta de homogeneidad a lo largo de la serie, ya que en este estudio utilizan pruebas de raíces unitarias con cambios estructurales (PIB per cápita estatal contra el PIB per cápita nacional) encontrando que se rechaza la hipótesis de raíces unitarias para varios estados al permitir el cambio estructural. De igual manera, estos autores encuentran que la componente determinística del PIB per cápita no es común entre todos los estados, poniendo de manifiesto la importancia de tomar en cuenta los cambios a lo largo de serie y mostrando además evidencia de comportamientos que son consistentes con la existencia de clubs de convergencia.

En cuanto a la búsqueda de clubs de convergencia Asuad y Quintana (2010) tratan de identificar, lo que ellos llaman núcleos de convergencia, al hacer un análisis visual de los grupos formados al clasificar a los estados de acuerdo a su tasa de crecimiento, entre sus resultados concluyen que los estados con más ingreso en 1970 lo siguen siendo en el 2008. En

el trabajo de Ruiz (2010) se aborda de forma interesante el tema de la heterogeneidad de los elementos económicos aplicando MCO para estimar la velocidad de convergencia de cada una de las entidades federativas, encontrando diferentes velocidades entre los estados, de tal forma que no solo habría la posibilidad de experimentar diversos estados estacionarios sino también diversas velocidades de convergencia, punto sobre todo importante al ser la homogeneidad de este parámetro un supuesto para la validez de los análisis de corte transversal y los basados en análisis de raíces unitarias. Por su parte Borrayo y Castañeda (2011) realizan un análisis exploratorio mediante técnicas de análisis de la dinámica de una distribución para la región del centro de México agregando efectos espaciales, con la finalidad de detectar regímenes espaciales (clubes) en el que se muestra evidencia de convergencia y alta movilidad, pero con la existencia de una estructura regional estratificada, es decir la formación de clubes, cabe mencionar que este estudio toma como unidad de estudio los municipios, a diferencia del resto que son a nivel de entidad federativa. En cuanto a los estudios que abordan el tema de la convergencia regional utilizando las propiedades estadísticas de las series de tiempo encontramos trabajos como el de Ayala, Chapa, Murgía (2011) y German-Soto, Salazar (2016); en su estudio utilizan la metodología de Nahar e Inder (2002) en la que se busca convergencia absoluta respecto a un *líder*. Para el primero tomaron como líder al promedio Nacional y a la Ciudad de México (DF), cabe señalar que este trabajo encontró que alrededor del 70 % de los estados presentan un patrón de convergencia, mientras que los segundos encontraron evidencia de convergencia de 34 % de los estados cuando se permite un cambio estructural en la serie. También Brida et al. (2013) encuentra un proceso de convergencia entre los estados clasificados como de alto desempeño, mientras que se presenta divergencia entre los estados clasificados como de bajo desempeño, estos resultados obtenidos bajo el esquema de una dinámica de dualismo estructural.

Respecto a los trabajos que utilizan la metodología de Phillips y Sul se encuentran los trabajos de Rodriguez, Lopez, Mendoza (2016) y Mendoza-Velázquez et al. (2019); y aunque ambos estudios son realizados a nivel estatal obtienen resultados muy diferentes. El primero encuentra seis clubs de convergencia al analizar el periodo de 1970 a 2012, mientras que los últimos encuentran solo dos clubs para el periodo 1940 a 2015 al aplicar la técnica sobre el ingreso per cápita.

Los estudios mencionados tienen la intención de mostrar las disparidades resultantes del proceso económico en México, pero además plantean la subsecuente necesidad de aplicar políticas diferenciadas a los diferentes territorios regionales y subnacionales para que estas políticas tengan efecto en el crecimiento de las entidades. Tomando eso en cuenta el aporte del presente trabajo es llevar el análisis a un nivel de detalle más fino, esto es, a un nivel de zona metropolitana, abonando en favor de este proceso de diferenciación. De esta forma se busca obtener una clasificación de las zonas metropolitanas que brinde información sobre el desarrollo histórico y (posiblemente) futuro de las Zonas Metropolitanas, y que además establezca un punto de partida para el estudio de los determinantes que definen su estado actual. De forma puntual se busca clasificar las Zonas Metropolitanas de México de forma endógena, a través del índice lumínico. Probando la hipótesis de si las zonas metropolitanas conforman clubs de convergencia estocástica respecto al crecimiento medido a través de la luminosidad asociada a las actividades económicas.

3.2. Metodología

Con la finalidad de conocer si las áreas metropolitanas se agrupan alrededor de diferentes estados estacionarios, este estudio sigue la metodología utilizada para el análisis de convergencia propuesta por Phillips y Sul (2009). Esta prueba es conocida como prueba *log t* de convergencia. Esta metodología tiene como objetivo estimar la convergencia absoluta o relativa entre un conjunto de entidades, pero además permite la conformación de subgrupos en los que se puede presentar convergencia entre los elementos del subgrupo generado. Además permite detectar entidades que presenten un comportamiento divergente a través de la aplicación cíclica de la prueba.

Un punto importante de esta metodología es que toma en cuenta la heterogeneidad que podrían presentar las entidades entre sí y en el tiempo, de tal forma que considera la posibilidad de un proceso de transición temporal de los elementos. Este proceso se puede observar iniciando con un modelo sencillo de un solo factor:

$$X_{it} = \delta_i \mu_t + \epsilon_{it}, \tag{3.1}$$

donde δ_i captura la distancia individual entre un factor común μ_t y la parte sistemática de X_{it}. De esta forma se modela la relación de X_{it} con respecto a μ_t a través de dos elementos idiosincráticos, el error ϵ_{it} y la parte sistemática δ_i, nótese que en este modelo la parte sistémica permanece fija en el tiempo, sin embargo Phillips y Sul (2007)[48] modifican este modelo y permiten que varíe en el tiempo convirtiendo δ_i en δ_{it}. Además permiten que δ_{it} contenga una componente aleatoria de tal manera que el modelo toma la forma:

$$X_{it} = \delta_{it} \mu_t \tag{3.2}$$

Esta ecuación se puede descomponer hacia la forma tradicionalmente usada para un panel de datos:

$$X_{it} = g_{it} + a_{it} \tag{3.3}$$

donde g_{it} incorpora los componentes sistemáticos, incluyendo los componentes permanentes que dan origen a la dependencia de corte transversal, mientras que a_{it} representa el componente transitorio. De tal forma que la ecuación 3.2 se relaciona con la ecuación 3.3 por medio de la transformación:

$$X_{it} = \left(\frac{g_{it} + a_{it}}{\mu_t} \right) \mu_t = \delta_{it} \mu_t \tag{3.4}$$

De esta manera X_{it} se divide en un componente compartido μ_t y uno idiosincrático δ_{it}, siendo ambos variantes en el tiempo y además permitiendo heterogeneidad temporal y entre individuos a través de δ_{it}. Es a través de esta sencilla especificación que se analiza el proceso de convergencia por medio de δ_{it}, de tal forma que la convergencia se da cuando:

$$\lim_{t \to \infty} \delta_{it} = \delta_i \tag{3.5}$$

Para la formulación de la prueba estadística que evalúa la condición de convergencia Phillips y Sul proponen un modelo semiparamétrico para el comportamiento de δ_{it} en el tiempo de la forma:

$$\delta_{it} = \delta_i + \sigma_i \xi_{it} L(t)^{-1} t^{-\alpha} \tag{3.6}$$

donde δ_i es un efecto que se mantiene en el tiempo, $\sigma_i > 0$, ξ_{it} es i.i.d. (0,1) en i pero débilmente dependiente en t, y $L(t)$ es una función que varía lentamente en el tiempo y que tiende a infinito cuando t tiende a infinito, esta función es la que Phillips y Sul (2007) asumen como $log(t)$. Como se puede observar, el parámetro α es el que define la velocidad a la que el componente transitorio decae a cero en el tiempo. Es bajo esta formulación que se asegura que δ_{it} converge a δ_i cuando $\alpha > 0$, convirtiéndose así en la hipótesis nula de interés, es decir:

$$H_0 : \delta_{it} = \delta \text{ y } \alpha \geq 0 \tag{3.7}$$

y la alternativa:

$$H_A : \delta_{it} = \delta \text{ para todo } i \text{ con } \alpha < 0 \tag{3.8}$$

o

$$H_A : \delta_{it} \neq \delta \tag{3.9}$$

La hipótesis alternativa incluye la posibilidad de divergencia, pero también puede considerar clubs de convergencia, por ejemplo si se tienen dos clubs la hipótesis alternativa sería de la forma:

$$H_A : \delta_{it} \rightarrow \begin{cases} \delta_1 \text{ y } \alpha \geq 0, \text{ si } i \in G_1 \\ \delta_2 \text{ y } \alpha \geq 0, \text{ si } i \in G_2 \end{cases} \tag{3.10}$$

donde G hace referencia a un club.

Para poder realizar la prueba con los datos observables Phillips y Sul (2007)[48] proponen medir la transición de δ_{it} por medio de la construcción de una medida relativa de los coeficientes de transición:

$$h_{it} = \frac{X_{it}}{N^{-1} \sum_{i=1}^{N} X_{it}} = \frac{\delta_{it}}{N^{-1} \sum_{i=1}^{N} \delta_{it}} \tag{3.11}$$

así h_{it} se convierte en una medida de transición relativa y funciona como proxy del proceso de convergencia (o divergencia) en el tiempo respecto al panel, midiendo la desviación del individuo i respecto a un estado estacionario común μ_t.

Como siguiente paso Phillips y Sul (2007) construyen la media cuadrada del diferencial de transición H_1/H_t donde:

$$H_t = \frac{1}{N} \sum_{i=1}^{N} (h_{it} - 1)^2 , \tag{3.12}$$

que mide la varianza de corte transversal respecto a un límite en común. Por último Phillips y Sul (2007)[48] muestran que la hipótesis formulada puede ser estadísticamente probada por medio del modelo de regresión:

$$\begin{aligned} &log(H_1/H_t) - 2\ log\ L(t) = \hat{c} + \hat{b}\ log\ t + u_t \\ &\forall \quad t = [rT], [rT]+1, \ldots, T \quad r \in (0,1) \end{aligned} \tag{3.13}$$

donde el parámetro r sirve para descartar la primera parte de la serie, siendo recomendable la utilización de $r = 0{,}3$ según los experimentos Monte Carlo reportados por Phillips y Sul (2007)[48] para muestras por abajo de T = 50. En esta regresión el parámetro clave para la prueba de convergencia es $\hat{b}$, que está relacionado con α, de hecho Phillips y Sul (2007)[48] muestran que se cumple la relación $\hat{b} = 2\hat{\alpha}$ donde $\hat{\alpha}$ es el valor estimado de α bajo la hipótesis nula. Así se puede probar la hipótesis nula de convergencia mediante una prueba t de una cola donde se puede rechazar la hipótesis nula de convergencia al 5 por ciento de significancia si $t_{\hat{b}} < -1{,}65$ según exponen los autores. En el caso que no se rechace la hipótesis nula, se tienen dos escenarios, el primero es cuando $\hat{b} < 2$ donde se tiene un proceso de convergencia condicional. Por otro lado, cuando $\hat{b} \geq 2$ se está en presencia de un proceso de convergencia absoluta (Phillips y Sul, 2009).

Así la metodología propuesta por Phillips y Sul (2007)[48] permite que la velocidad de convergencia sea diferente entre individuos, además de poder ser diferente en el tiempo para un mismo individuo por lo que este enfoque está preparado para tomar en cuenta cambios en la velocidad de convergencia de un país, estado o zona metropolitana en el tiempo; dándole mayor poder a la prueba sobre falsos negativos en la presencia de heterogeneidad temporal a diferencia de enfoques de raíz unitaria.

3.2.1. Algoritmo de convergencia

Para la aplicación de la prueba *log t* Phillips y Sul (2009)[49] desarrollaron un algoritmo para la búsqueda de clubs de convergencia, ya que un aspecto interesante del marco analítico propuesto es que el rechazo de la hipótesis nula del panel completo no implica la inexistencia de convergencia a nivel de subgrupos, ya que es posible evaluar si existen clubs de convergencia aplicándole de manera recursiva el algoritmo propuesto. El algoritmo tal y como lo presentan Phillips y Sul (2009)[49] pag. 1170, consta de los siguientes pasos:

1. Se ordenan los individuos de forma descendente respecto al promedio de la última tercera parte de la serie.

2. Se seleccionan los k individuos con los valores más altos, para formar el grupo base G_k con $2 \leq k < N$, al cuál se le aplica la regresión $log\, t$ y se selecciona el número k de elementos tal que se maximice el criterio:

$$k^* = arg\, max\{t_k\} \text{ sujeto a } t_k > -1{,}65$$

3. Se agrega uno a uno los elementos restantes al grupo base G_k y permanece en el grupo si el estadístico t es mayor que el criterio c^*.

4. Se aplican los pasos anteriores sobre los elementos que no fueron agregados al grupo, para crear un nuevo grupo. El procedimiento se repite hasta que todos lo elementos fueron asignados a un grupo. En caso que no sea posible crear el grupo base en alguna de las iteraciones se concluye que los elementos restantes tienen un comportamiento divergente.

Respecto al criterio de inclusión c^*, Phillips y Sul (2009)[49] explican que está asociado con qué tan estricto se desea ser en la inclusión de elementos al grupo. A mayor valor de c^* menor riesgo de incluir erróneamente un elemento al club. Así cuando c^* se aproxima a cero desde los negativos se vuelve más estricto. Como mencionan también en su artículo al tener un T pequeño se recomienda fijar $c^* = 0$, sin embargo al ser un criterio muy estricto recomiendan correr la prueba $log\, t$ sobre los clubs generados para descartar la posibilidad de estar generando un problema de sobre especificación, al generar un número mayor de clubs al que realmente existen.

3.3. Índice Lumínico per cápita por área metropolitana

Los datos lumínicos tienen la capacidad de reflejar la actividad económica (Sutton et al. 2007), sobre todo en regiones con instituciones estadísticas que no son de primer nivel o en aquellas con grados considerables de informalidad (Nordhaus y Chen 2015, Ghosh et al. 2009), pero además tienen la ventaja de ser versátiles en cuanto la disponibilidad por delimitación geográfica, como ya lo hace patente el estudio de Llamosas-Rosas, Rangel y Sandoval (2018). Otra ventaja que pensamos tiene el *IL* como proxy de la actividad económica es su carácter real, ya que en el caso de las zonas metropolitanas de la República Mexicana existen además dos distorsiones presentes en los datos de fuentes tradicionales de información, la informalidad y el componente de la renta petrolera para algunos estados, además de la falta de disponibilidad de las series en algunos casos. El *IL* no excluye la actividad económica informal y respecto a las distorsiones de los rentas petroleras estas pueden tener un efecto tan grande como el que provoquen de forma real en sus localidades. Por estos motivos, en este estudio se propone el uso del Indice Lumínico per cápitá (*ILpc*) para medir el crecimiento de largo plazo de las zonas metropolitanas de México.

Para estimar el índice se utilizaron los datos del Departamento de Defensa de Estados Unidos y procesados por la Oficina Nacional de Administración Oceánica y Atmosférica[3] (NOAA por sus siglas en inglés) y se siguió la metodología propuesta por Henderson et al.(2012)[28] promediando las observaciones por año, ya que existen periodos en los que dos satélites registran los datos para un mismo año. Los datos lumínicos generados tienen una frecuencia anual de 1992 a 2015 correspondientes a algún instante entre las 8:30 p.m. y las 10:00 p.m. (hora local). La unidad de observación es el pixel, que corresponde a un área aproximada de $0{,}88 Km^2$ de 1992 hasta el 2013, datos recabados mediante la tecnología OLS (*Operational Linescan System*). A partir de 2013 los datos fueron generados con la tecnología VIIRS (*Visible Infrared Imaging Radiometer Suite*) que permite mayor detalle, generando datos por pixel que corresponden a $0{,}44 Km^2$ aproximadamente. Cabe mencionar que la disponibilidad de datos bajo esta última tecnología es de periodicidad mensual para los años 2013 y 2014, por lo que se tomó el promedio anual a partir de los datos mensuales de forma similar a lo realizado por Chen y Nordhaus (2019)[12]. Otro punto en que se diferencian ambas tecnologías es respecto a los valores máximos que pueden tomar los pixeles, la tecnología más nueva se presenta en valores absolutos de $nano{-}Watts/(cm^2 \cdot sr)$ y no tiene un límite superior en cuanto a la medición, mientras que la anterior presenta valores relativos entre 0 y 63 del

[3] https://www.ngdc.noaa.gov/eog/dmsp/downloadV4composites.html

espectro visible de la luz. Cabe mencionar que para el año 2013 se cuenta con los datos bajo ambas tecnologías, de tal forma que se utilizó este año para validar la unión entre las series, en específico las observaciones bajo la tecnología OLS del satélite etiquetado como F18 y las observaciones bajo la tecnología VIIRS. Los datos bajo ambas tecnologías presentan una fuerte correlación, con un coeficiente de 0.9. En la gráfica 3.1 se puede observar la similitud entre los datos a nivel zona metropolitana bajo ambas tecnologías. Un punto que resalta al observar los datos es que no se presenta un valor consistentemente mayor por parte de la observación con el satélite VIIRS, por lo que no existe un sesgo sistémico relacionado a la tecnología, excepto en los casos en los que se satura la medición por OLS. Por lo anterior y siguiendo el procedimiento aplicado a las series con más de una observación por año, se calculó la serie promedio para el año 2013[4].

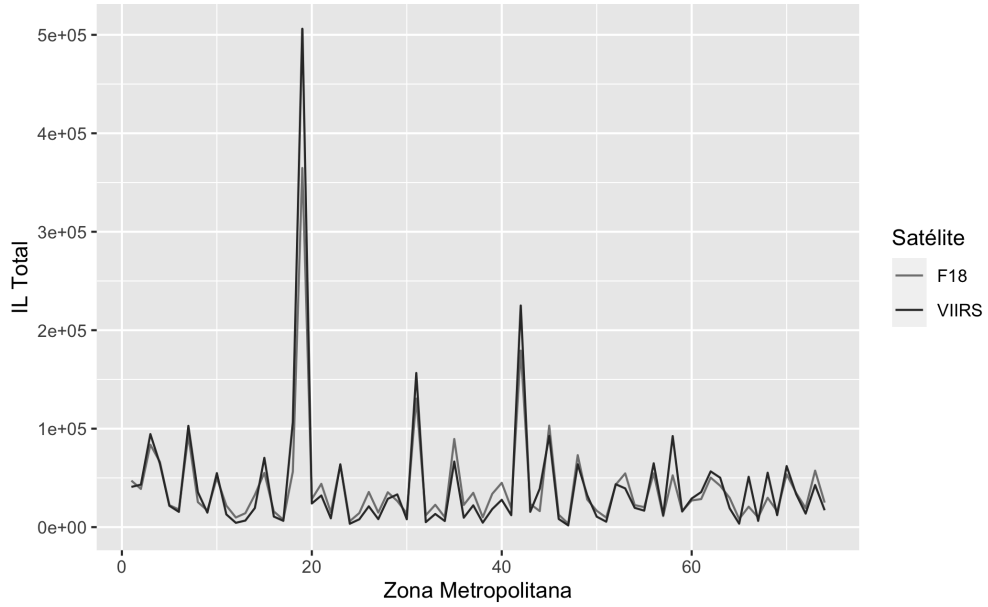


Figura 3.1: Luz total bajo ambas tecnologías para el año 2013 a nivel Zona Metropolitana. Fuente: Elaboración propia con datos del NOAA y CONAPO.

La fuerte correlación entre ambas series se puede explicar por el nivel de agregación territorial utilizado, a nivel zona metropolitana. Como demuestran Chen y Nordhaus (2015) [11] los datos obtenidos con la tecnología OLS en territorios con baja densidad poblacional tienen un pobre desempeño respecto a la estimación de la actividad económica. Es de hecho en territorios con estas características que se desempeñan de forma regular actividades como la agricultura, pesca, minería y explotación de recursos forestales. Sin embargo, este no es el caso de las zonas metropolitanas, donde se realizan actividades con uso intensivo de la luz, como producción de bienes, servicios, venta minorista, logística entre otros, es decir actividades secundarias y terciarias; por lo tanto la pérdida debido a la calidad de las imágenes

[4]En este sentido se busca conservar la variabilidad real de las series, ya que la metodología utilizada evalúa la distancia relativa contemporánea entre los elementos bajo estudio y no la consistencia del cambio en el tiempo como lo harían las pruebas de raíces unitarias.

se minimiza. En particular el trabajo de Chen y Nordhaus (2019)[12] establece que la relación entre el índice lumínico y el PIB es fuerte a nivel zona metropolitana,[5] de hecho mayor que a nivel estado, así también la relación entre la luz y el PIB en niveles es más fuerte que la relación entre el crecimiento del PIB y el crecimiento de la luminosidad, por lo menos para el caso de los Estados Unidos de América que cuenta con la información necesaria para el análisis.

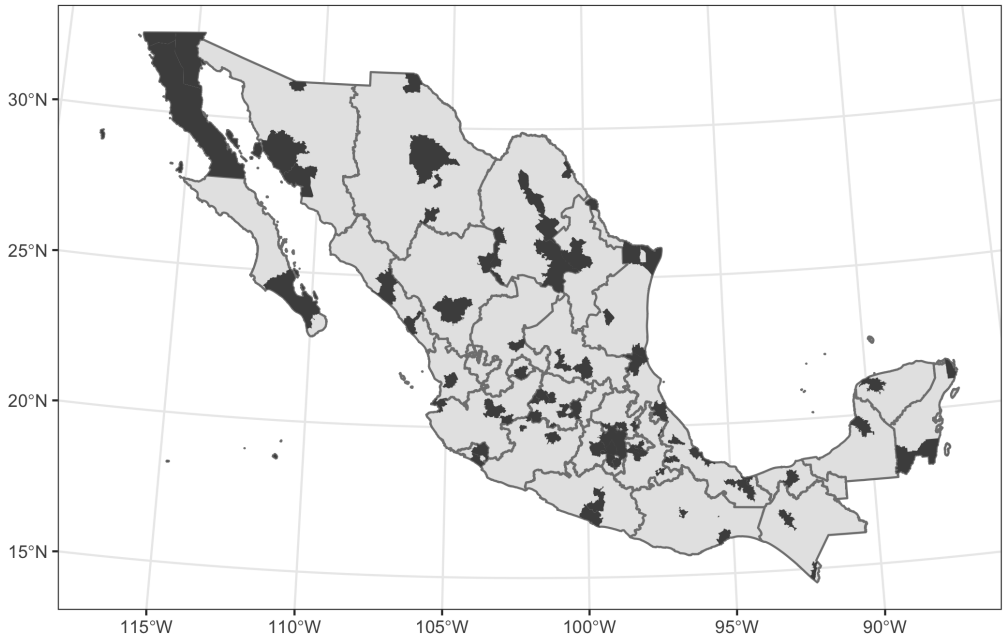
Para delimitar las áreas de las diferentes zonas metropolitanas se utilizó la delimitación definida conjuntamente por la Secretaría de Desarrollo Agrario, Territorial y Urbano (SEDATU), el Consejo Nacional de Población (CONAPO) y el Instituto Nacional de Estadística y Geografía (INEGI)[6]. Esta delimitación muestra que en México existen 74 zonas metropolitanas en el país, definidas en el documento como el *conjunto de dos o más municipios donde se localiza una ciudad de 100 mil o más habitantes, cuya área urbana, funciones y actividades rebasan los límites del municipio, incorporando dentro de su área de influencia directa a municipios vecinos, predominantemente urbanos, con los que mantiene un alto grado de integración socioeconómica*, adicionalmente se incluyen las capitales de los estados que no cumplan con el criterio anterior. Cabe mencionar que esta selección de municipios ha cambiado en el tiempo, agregándose nuevas municipalidades a medida que crecen las zonas urbanas, es por eso que este trabajo analiza los municipios de las zonas metropolitanas definidas en el año 2015, que es la más reciente al momento de la realización de este trabajo. Una vez obtenido el listado se obtuvo la georeferencia de los municipios por medio del mapa geoestadístico proporcionado por INEGI, el Mapa 1 muestra la ubicación geográfica obtenida de las zonas metropolitanas a lo largo del territorio nacional.

Como ya se ha probado en trabajos anteriores, iniciando con el estudio de Henderson et al. (2012), el Índice Lumínico (IL) está directamente relacionado con la actividad económica, específicamente para el caso de México Llamosas-Rosas et al. (2018) muestran que existe una relación log-lineal entre el IL y el Producto Interno Bruto (PIB), como es de esperarse esta relación no se pierde a nivel zona metropolitana, como se muestra en la Gráfica 3.2, en donde se compara el PIB nacional contra el IL de las zonas metropolitanas para los años 1992-2016. Esta relación log-lineal se mantiene, con un ajuste de $R^2 = 0{,}7$, en parte porque el 62,8 % de la población del país se concentra en estas poblaciones, según datos del documento de zonas metropolitanas antes mencionado. Además el IL no excluye la actividad informal, de esta forma, aún y cuando no se abarca la totalidad de la población este índice se mantiene cercano a las estimaciones oficiales.

Con la finalidad de tener una estimación de la producción per cápita para cada zona, el índice lumínico se ponderó utilizando los datos poblacionales de las diferentes áreas metropolitanas.Los datos relacionados a la población se tomaron de los documentos censales e intercensales de INEGI, para los años 1990, 1995, 2000, 2005, 2010 y 2015, interpolando linealmente estos datos para calcular los años faltantes. En la Figura 3.3 se muestra el Índice Lumínico per cápita (ILpc) total de las zonas metropolitanas de México para el periodo 1992-2015. Uno de los puntos que resaltan al observar este gráfico es el caso del periodo 1994-1995. Este puede parecer contraintuitivo ya que presenta un incremento en el índice lumínico siendo que se esperaría un decremento debido a la crisis económica de ese periodo. Sin embrago,

[5]Área estadística metropolitana o MSA por su siglas en inglés como lo define la Oficina de Administración y Presupuesto de los Estados Unidos de América.

[6]Del documento Delimitación de las zonas metropolitanas de México 2015[61].



Mapa 1: Zonas metropolitanas de México 2015. Fuente: Elaboración propia con datos de SEDATU-CONAPO-INEGI.

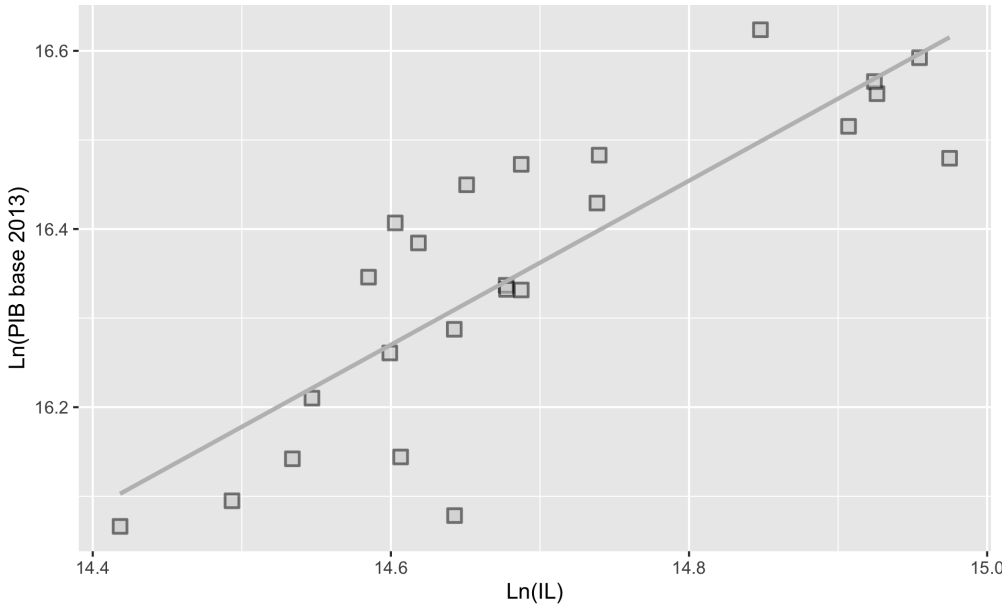


Figura 3.2: Relación Índice Lumínico contra PIB Nacional. Fuente: Elaboración propia con datos del NOAA e INEGI.

este comportamiento se explica desde dos vertientes, la primera es la referente a la fuente de los datos, ya que durante este periodo se tiene información de dos satélites, el F10 y el F12. Para 1994 se tienen datos del satélite F10 y F12, y para 1995 se tienen datos solo del satélite

F12. Esto provoca que se presente un brinco, al tener el satélite F12 registros con valores más altos debido a características técnicas del satélite[7]. Por otro lado, si observamos solo los datos del satélite F12 encontramos aún un incremento de 1994 a 1995, respecto a este punto se puede pensar en el efecto de la informalidad pasando de actividades formales intensivas en capital hacia la informalidad a nivel de calle.

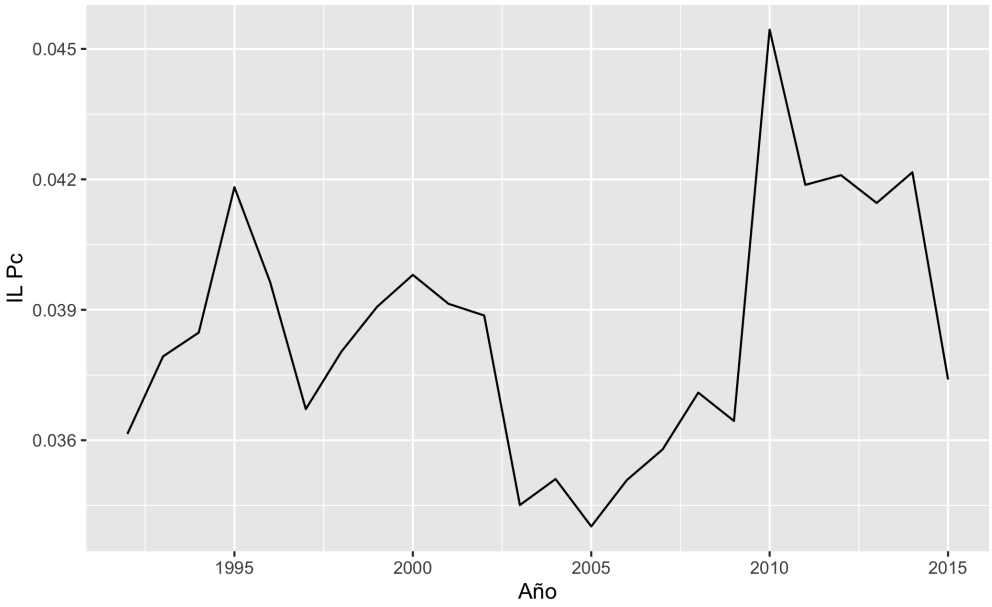


Figura 3.3: Índice Lumínico per cápita total de las zonas metropolitanas de México. Elaboración propia con datos del NOAA e INEGI.

Si observamos estos datos desagregados a nivel estatal (Figura 3.4) podemos observar que el comportamiento no es homogéneo entre los estados, como lo menciona Rodriguez (2016)[52] en el que identifica seis diferentes grupos al analizar el ingreso estatal. Si bien el trabajo anteriormente mencionado saca a relucir dichas diferencias, no son estas del todo obvias, como ejemplo de esto podemos tomar una agrupación tradicionalmente utilizada, la franja de los estados fronterizos del norte. Esta agrupación sale a relucir desde inspecciones meramente visuales como la de Asuad y Quintana (2010)[60] hasta análisis de desempeño como el de Brida et al. (2013)[7] que clasifica estos estados como integrantes del grupo denominado de alto desempeño. Si observamos los datos referentes al *ILpc* para las zonas metropolitanas pertenecientes a dichos estados podemos entender un poco el porqué de dichas conclusiones. Al observar la Figura 3.5, donde se muestra el *ILpc* de las zonas metropolitanas de los estados del norte agrupado por estado, podemos notar la similitud en el comportamiento del índice entre dichos estados, donde las diferencias son en su mayoría de nivel.

Sin embargo al observar a un nivel más detallado, es decir a nivel zona metropolitana, las similitudes ya no parecen ser tan obvias, además salen a relucir comportamientos diversos dentro los mismos estados, como se puede observar en la Figura 3.6 en la que se muestran

[7]Ver el trabajo de Guerrero y Mendoza (2018) en el que abordan las diferencias en las series de tiempo por tipo de satélite para el caso de la República Mexicana.

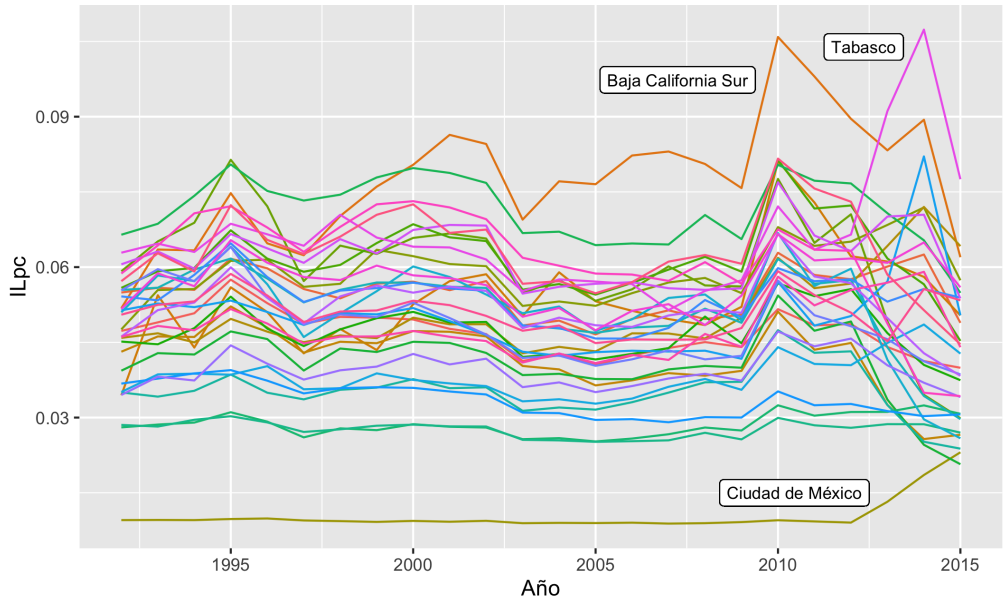


Figura 3.4: Índice Lumínico per cápita de las Zonas Metropolitanas agrupado por Estado. Elaboración propia con datos del NOAA e INEGI.

los datos a nivel zona metropolitana para Baja California y Tamaulipas, dos estados que en nivel agregado se comportaban de forma muy similar y que al observar el detalle a nivel zona metropolitana presentan diferencias en el comportamiento tanto al interior de los estados como entre ellos.

La intención de este trabajo es analizar esta dinámica a este nivel más detallado, que en nuestro conocimiento actual no ha sido abordado en la combinación de tipo de datos, detalle geográfico y metodología de análisis. Al identificar las distintas dinámicas de crecimiento lumínico se exhibe la heterogeneidad de los procesos presentes en las zonas metropolitanas, efecto que permea a los estados, a las regiones y a los países.

3.3.1. Estadísticos Descriptivos del Índice Lumínico

Al observar el *IL* por zona metropolitana nos muestra un escenario conocido en el que las tres áreas con mayor *IL* en promedio para el periodo 1992-2015 son en primer lugar el Valle de México en la Ciudad de México (anteriormente Distrito Federal), seguido de Monterrey en Nuevo Léon y Guadalajara en Jalisco, como se esperaría al tratarse de las principales zonas del país. Por el otro, los índices más bajos los presenta Teziutlán en Puebla, Moroleón-Uriangato en Guanajuato y Acayucan en Veracruz, que de forma interesante no contiene zonas metropolitanas de la franja fronteriza sur como se habría esperado. Por otra parte, la historia que cuenta el *IL* es muy diferente al controlarlo por habitante, dos de las tres zonas metropolitanas con mayor *IL* presentan el menor *ILpc* promedio, es decir, el Valle de México y Guadalajara; además de Xalapa en Veracruz. Por el contrario los mayores *ILpc* los presentan Piedras Negras en Coahuila, Tehuantepec en Oaxaca y Tula en Hidalgo. En

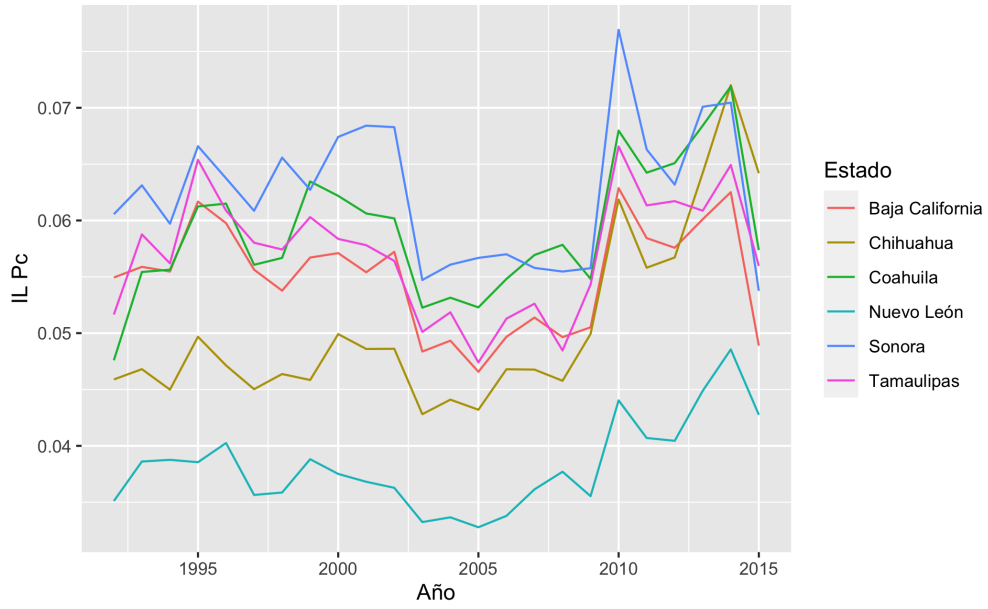


Figura 3.5: Índice Lumínico per cápita de las zonas metropolitanas para los estados de la frontera norte. Elaboración propia con datos del NOAA e INEGI.

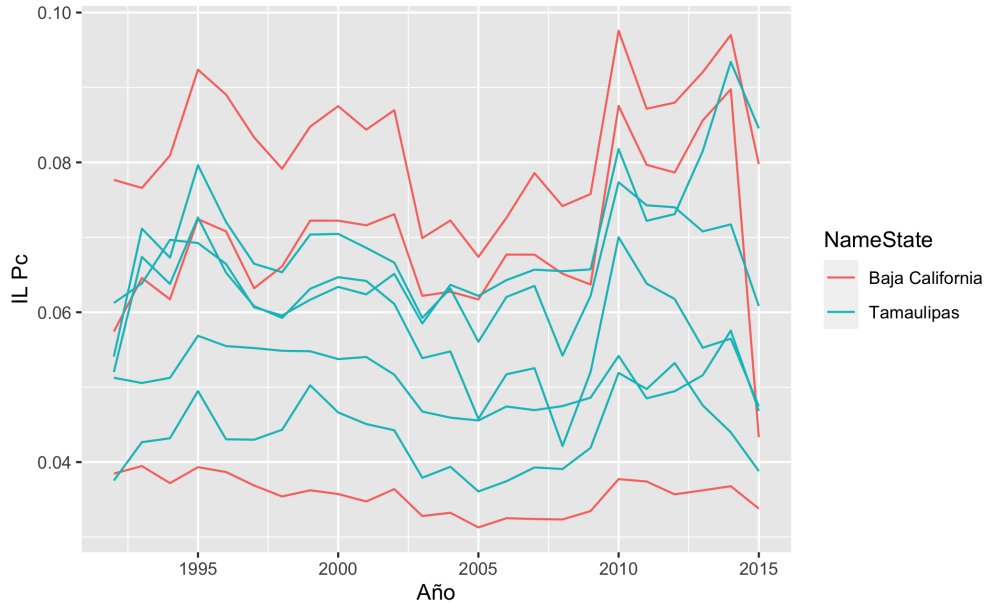


Figura 3.6: Índice Lumínico per cápita por Zona Metropolitana para Baja California y Tamaulipas.

Elaboración propia con datos del NOAA e INEGI.

general, llama la atención la presencia de Tehuantepec[8] ya que por lo regular las regiones del

[8]Para verificar la veracidad de los datos correspondientes a esta zona metropolitana se analizaron para los años 2013 y 2014 en los que se cuenta con datos mensuales, en los cuales se confirman los valores a nivel

sur del país han estado históricamente asociadas a bajos niveles de desarrollo, sin embargo al hacer el análisis más detallado es posible detectar comportamientos que no surgen en el agregado y que pueden suponer casos de éxito bajo el contexto particular de esas regiones. En las Figuras 3.7 y 3.8 se muestran las zonas metropolitanas con valores máximos respecto al *IL* y el *ILpc*, en las gráficas se muestran el grupo G_{maxIL} que lo conforman las tres zonas metropolitanas con mayor IL y el grupo $G_{maxILpc}$ que lo conforman las zonas metropolitanas con mayor ILpc. En la Figura 3.7 se muestran las zonas metropolitanas respecto al *IL*, en la que es evidente la diferencia tan amplia de la zona metropolitana del Valle de México, sin embargo al controlar por población el Valle de México pasa a la parte baja como se muestra en la Figura 3.8 y las zonas metropolitanas de Piedras Negras, Tehuantepec y Tula pasan a ser las destacadas en cuanto a la relación luz/población. Es interesante el análisis que se puede desprender de la interpretación de esta razón, tanto desde la perspectiva de bienestar como de una de productividad, y si bien no está dentro del alcance de este estudio si consideramos que podría dar pie a futuros trabajos.

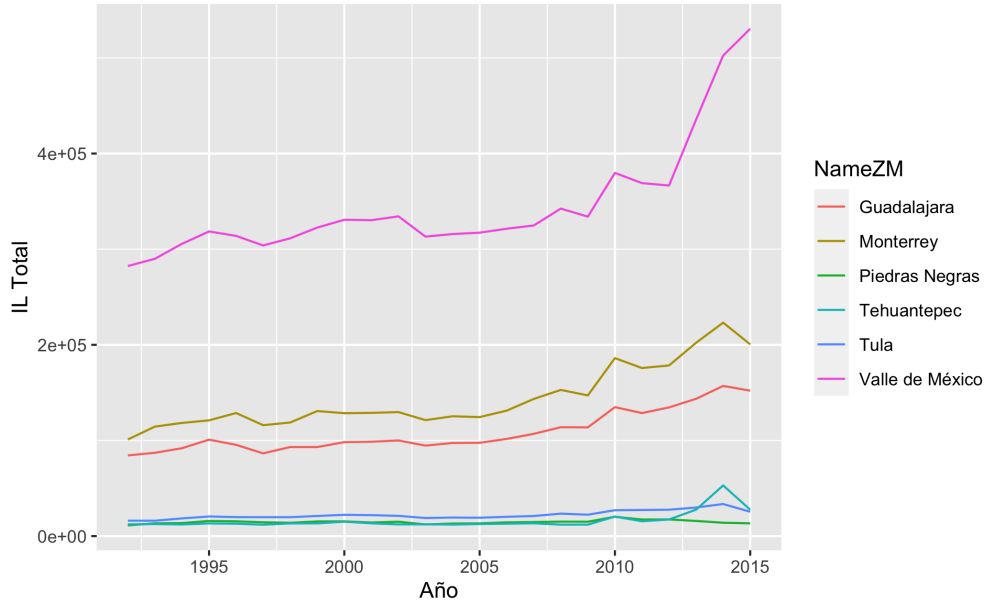


Figura 3.7: Grupos G_{maxIL} y $G_{maxILpc}$ graficados respecto al *IL*.

En cuanto a la distribución de los datos para las regiones analizadas podemos observar en la Figura 3.9 que los datos para los últimos tres años tienen valores máximos más altos que los años anteriores, esto se debe a que la tecnología utilizada en esos periodos permite lecturas más altas, sin embargo los datos en general para esos periodos presentan valores más bajos, por lo que el cambio en tecnología no imputa valores más altos a la serie de forma sistémica. De esta forma el análisis solo se vería afectado por esta situación en el caso de los elementos que presentaban valores en el límite superior bajo la tecnología inicial y que superen dicho límite para los periodos capturados con la tecnología más reciente.

anual.

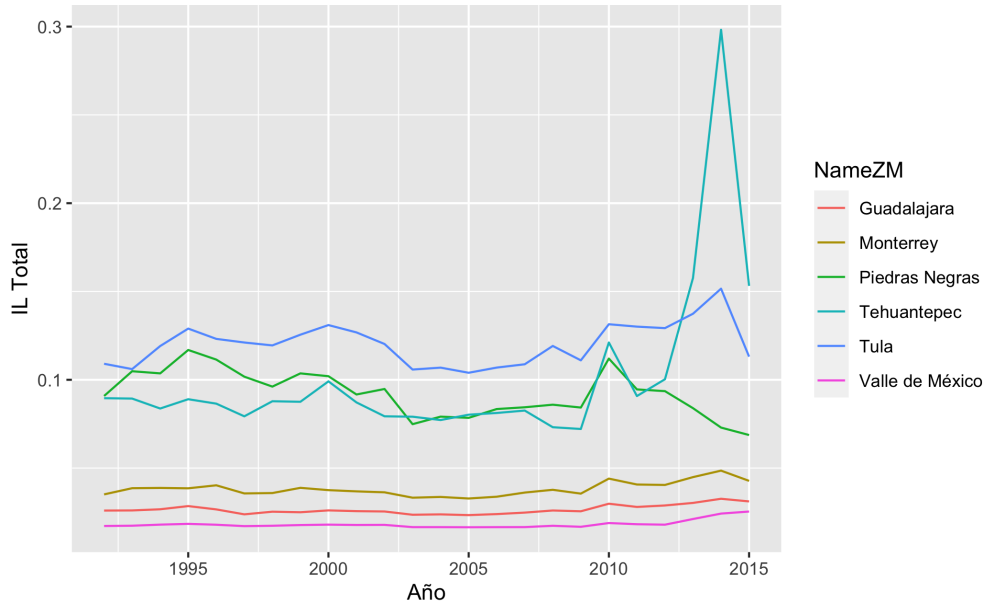


Figura 3.8: Grupos G_{maxIL} y $G_{maxILpc}$ graficados respecto al *ILpc*.

3.4. Resultados

Con la finalidad de identificar los clubs de convergencia en los que se dividen las sendas de crecimiento de las zonas metropolitanas de la República Mexicana se aplicó el algoritmo de Phillips y Sul (2009) a un total de 74 zonas metropolitanas, para el periodo 1992-2015, siendo la variable de estudio el Índice Lumínico per capita (*ILpc*). Los resultados de estas estimaciones se muestran en el Cuadro 3.2. La primer columna del cuadro muestra el número que identifica al grupo o club de convergencia al que pertenece la zona metropolitana (ZM), la segunda columna el número de zonas metropolitanas dentro del grupo, la tercer columna el coeficiente $\hat{b}$ del modelo definido en la ecuación (3.13), la cuarta el error estándar asociado al coeficiente y finalmente la última columna el estadístico t asociado a la hipótesis nula de convergencia (ecuación 7).

Del Cuadro 3.2 se observa que ningún club fue detectado como divergente ya que el estadístico t es mayor al valor crítico (-1.65) por lo que no se rechaza la hipótesis nula a un nivel de significancia de 0.05. Estos resultados contrastan con los resultados encontrados en la literatura. Por ejemplo, Ayala y Chapa (2011) encuentran que 9 entidades no convergen hacia el líder, mientras que al permitir la formación de clubs de convergencia se logra identificar dicho compartamiento como algo más que una divergencia. Además, respecto al número de clubs identificados, el trabajo de Rodríguez et al. (2016) encuentra la conformación de 6 clubs entre los estados mientras que al incrementar el nivel de detalle hacia zonas metropolitanas solo se encuentran 5 clubs, hecho que atribuimos a la utilización de un mayor nivel de detalle. Ya que se evita la generación de un clubes adicionales que capturan el comportamiento combinado de dos o más grupos base. De tal forma estos resultados sugieren la importancia de considerar una unidad de estudio más pequeña que la de entidades federativas. Aunque los trabajos antes mencionados abordan el tema de convergencia a nivel estatal resulta in-

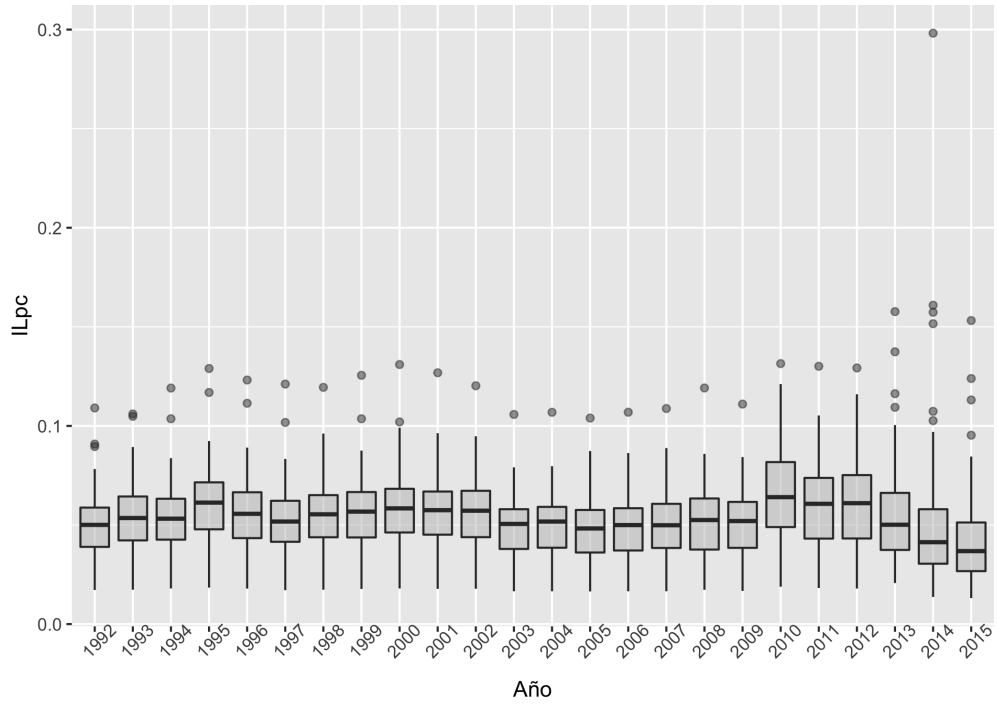


Figura 3.9: Distribución de los datos en el tiempo.

teresante analizar sus resultados a la luz de los resultados en nivel de zonas metropolitanas, en especial los referentes a los estados que presentaron divergencia o no convergencia en los estudios previos y observar de forma detallada el comportamiento de las zonas metropolitanas que los componen. Por ejemplo, en el estudio de Ayala y Chapa (2011), se encuentra que solamente cuatro estados no presentan convergencia con respecto a la Ciudad de México a un nivel de significancia del 1 %. De estas entidades, Baja California, Colima y Quintana Roo presentan clasificaciones mixtas bajo nuestro enfoque. Es decir, zonas metropolitanas en distintos clubes, todas con al menos un club diferente al que pertenece la Ciudad de México. El cuarto estado, Durango, tiene una zona metropolitana única que de hecho corresponde en su clasificación a la Ciudad de México en el ?Club Mesoamericano?, sin embargo es posible que este comportamiento no se haya hecho visible en el trabajo de Ayala y Chapa (2011) debido al efecto de la zona metropolitana de La Laguna, que comparte con Coahuila, efecto que en nuestro caso si fue aislado.

Finalmente, entre los resultados obtenidos resalta el hecho de que todos clubs están lejos de presentar un proceso de convergencia absoluta ya que para ningún club el estimador es positivo. Además, el hecho de que se presenten valores cercanos a cero indica un proceso de convergencia lento, que es consistente con algunos resultados en la literatura que dan cuenta de un proceso de convergencia débil (Esquivel, 1999; Cermeño, 2001; Chiquiar, 2005 y Rodriguez-Oregia, 2005).

Año	Media	Varianza	Q1	Q2	Q3
1992	0.0509	0.0002	0.0390	0.0501	0.0588
1993	0.0546	0.0003	0.0422	0.0535	0.0644
1994	0.0547	0.0003	0.0426	0.0532	0.0633
1995	0.0611	0.0004	0.0478	0.0613	0.0715
1996	0.0567	0.0003	0.0434	0.0556	0.0665
1997	0.0522	0.0003	0.0415	0.0517	0.0622
1998	0.0549	0.0003	0.0438	0.0554	0.0651
1999	0.0564	0.0003	0.0437	0.0568	0.0666
2000	0.0580	0.0004	0.0462	0.0584	0.0683
2001	0.0568	0.0003	0.0451	0.0575	0.0669
2002	0.0563	0.0003	0.0439	0.0572	0.0673
2003	0.0489	0.0002	0.0379	0.0506	0.0580
2004	0.0501	0.0003	0.0385	0.0517	0.0591
2005	0.0484	0.0003	0.0361	0.0483	0.0576
2006	0.0499	0.0003	0.0371	0.0500	0.0585
2007	0.0510	0.0003	0.0384	0.0499	0.0607
2008	0.0525	0.0003	0.0376	0.0525	0.0634
2009	0.0513	0.0003	0.0385	0.0520	0.0617
2010	0.0671	0.0006	0.0490	0.0641	0.0817
2011	0.0606	0.0005	0.0431	0.0607	0.0738
2012	0.0612	0.0005	0.0432	0.0610	0.0752
2013	0.0560	0.0007	0.0374	0.0501	0.0662
2014	0.0543	0.0018	0.0305	0.0413	0.0580
2015	0.0434	0.0006	0.0267	0.0368	0.0513

Cuadro 3.1: Estadísticos descriptivos del ILpc 1992-2015

3.4.1. Heterogeneidad interestatal.

Al estudiar la dinámica del crecimiento económico regional se pueden presentar diversos escenarios al tomar a las entidades federativas como unidad de estudio. Si tomamos como base que la dinámica económica de un estado es el resultado de la combinación de las dinámicas económicas de sus zonas metropolitanas podemos tener que 1) existe solo una zona metropolitana en el estado, 2) existen mas de una zona metropolitana en el estado pero todas las zonas metropolitanas exhiben la misma dinámica de crecimiento y 3) existen más de una zona metropolitana en el estado y por lo menos una de ellas presenta una dinámica distinta a las demás. Si se toman acciones de política pública basados en un estudio a nivel estado no se tendría problema en los casos 1 y 2, puesto que se habría etiquetado correctamente al estado. Sin embargo, esto no sería correcto para el tercer escenario ya que en el mejor de los casos se estaría etiquetando respecto a las características de la(s) zona(s) metropolitana(s) predominante(s). De este argumento radica la importancia de permitir la clasificación endógena de los elementos de forma independiente a la adscripción geopolítica y reconocer las condiciones particulares de cada zona económica.

Cuadro 3.2: Clubs de Convergencia de las Zonas Metropolitanas en México (1992-2015) .

Número de Club	**Número ZM**	**Coeficiente $\hat{b}$**	**Error estándar**	**Estadístico t**
1	2	-0.543	0.482	-1.127
2	6	-0.183	0.134	-1.364
3	13	-0.302	0.222	-1.361
4	29	-0.517	0.333	-1.554
5	24	-0.447	0.285	-1.568

En este sentido los resultados muestran que de los 32 estados solo 12 contienen una sola zona metropolitana. De los 20 estados restantes solo 4 tienen a todas sus zonas metropolitanas en un solo club. De tal manera que 16 de los 32 estados presentan heterogeneidad en los procesos de convergencia lumínica per cápita interestatal. Es decir, estas entidades caen en el tercer escenario, con al menos dos zonas metropolitanas en diferente club de convergencia. En el Cuadro 3.3 se muestran los estados clasificados bajo este criterio. La primer columna clasifica a los estados con una composición homogénea de sus zonas metropolitanas, ya sea porque solo cuentan con una zona o que todas pertenecen al mismo club mientras que la segunda columna muestra los estados con una composición heterogénea de las zonas metropolitanas que lo componen.

Cuadro 3.3: Estados clasificados respecto a su heterogeneidad intraestatal.

Tipo de Composición.	
Homogéneos	Heterogéneos
Aguascalientes	Baja California
Baja California Sur	Colima
Campeche	Chihuahua
Coahuila	Guanajuato
Chiapas	Hidalgo
Ciudad de Mexico	Jalisco
Durango	Mexico
Guerrero	Michoacán
Nayarit	Morelos
Nuevo León	Oaxaca
Queretaro	Puebla
Sinaloa	Quintana Roo
Tabasco	San Luis Potosi
Tlaxcala	Sonora
Yucatan	Tamaulipas
Zacatecas	Veracruz

Los estados que presentan una mayor heterogeneidad son Chihuahua, Hidalgo y Tamaulipas con tres clubs diferentes y Veracruz con cuatro clubs (Cuadro 3.4). Un resultado que llama la atención es el hecho que Veracruz contiene zonas metropolitanas de todos los clubs excepto el número 3. Si bien este resultado puede estar relacionado con que sea el estado con más zonas metropolitanas, contrasta con situaciones como la de Coahuila, que contiene cuatro zonas metropolitanas, todas pertenecientes al Club3.

Cuadro 3.4: Estados con mayor número de Clubs.

Estado	Núm. Zonas M.	Clubs
Chihuahua	4	2,3,4
Hidalgo	3	2,4,5
Tamaulipas	3	2,3,4
Veracruz	4	1,2,4,5

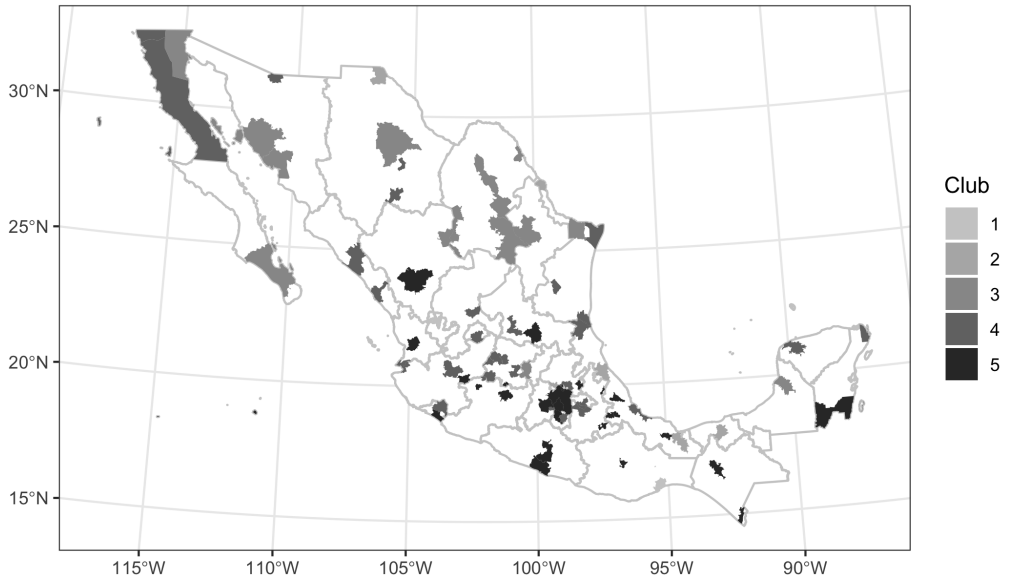
Estos resultados nos permiten comparar parcialmente contra los obtenidos en el trabajo de Rodríguez et al. (2016) realizado a entidades federativas, tomando aquellos estados que contienen zonas metropolitanas asignadas a un solo club. Es decir, los 16 estados con una composición homogénea de sus zonas metropolitanas. La forma de evaluar fue bajo el criterio utilizado en Rosenberg y Hirschberg (2007)[57] para la evaluación de procesos de agrupamiento, esta medida puede tomar valores entre cero y uno, donde uno es cuando todos los grupos contienen elementos de una sola clase. Bajo este criterio se obtuvo un valor de 0.37 entre la agrupación realizada y la que Rodríguez et al. reportan, que si bien no es un valor alto sí es mayor al valor que se obtendría al asignar de forma aleatoria uniforme entre los 5 clubs, es decir 0.20, por lo que se concluye que no están completamente desasociadas las clasificaciones.

3.4.2. Distribución geográfica de los Clubes.

Con respecto a la distribución geográfica de los clubes, éstos presentan resultados interesantes. Primero, la existencia de un club que está presente a lo largo de todo el país, literalmente de punta a punta del territorio nacional, el club 4, que no solo es él que tiene más elementos sino también el que está más uniformemente distribuido en el espacio. Por otro lado, existen dos clubs que dividen en dos al país. El club 3 que se presenta principalmente en el norte del país, con solo dos excepciones en Querétaro y Campeche. Y el club 5 que recuerda por su distribución a la región cultural mesoamericana (ver Mapa 2). En cuanto a los dos clubs restantes, que en conjunto agrupan solo el 10 % de las zonas metropolitanas, parecen agrupar a las zonas atípicas por decirlo de alguna manera, agrupando a pocos elementos pero con un comportamiento característico muy diferente a los demás.

3.4.3. Análisis de las sendas de transición relativas.

Como parte de la metodología propuesta en Phillips y Sul (2009) se encuentra el análisis gráfico de la trayectoria de convergencia (o divergencia) que sigue un elemento en el tiempo



Mapa 2: Distribución geográfica de los clubs de convergencia.

a través de la variable h_{it}, definida en la ecuación (3.11). La variable h_{it} traza la trayectoria individual para cada i respecto al promedio, así h_{it} se define como la *senda de transición relativa*. Además, h_{it} mide la distancia relativa de i respecto a la senda de crecimiento estacionaria μ_t, de tal forma que la senda de transición puede ser usada para evaluar la magnitud de un comportamiento divergente y evaluar si dicha divergencia es temporal o no (Phillips y Sul, 2009). Bajo un escenario de convergencia absoluta se esperaría que hit se aproxime a uno para todo i, mientras que bajo convergencia relativa se mantendrían los valores dentro de una banda. Si esta senda h_{it} se aproxima hacia la unidad desde valores mayores a uno, indica que las condiciones iniciales de i eran superiores al promedio, mientras que lo contrario aplica para valores h_{it} de menores a uno. Pro otro lado si se observa que h_{it} se aleja de la unidad, se entiende que hay un proceso de divergencia.

La Figura 3.10 muestra la senda de transición relativa promedio[9] de los cinco clubs de convergencia. Una vez más llama la atención el Club 4, que es el más numeroso y más extendido a lo largo del país, al presentar un comportamiento estable de convergencia, de hecho, este club es el único que toca en diversas ocaciones la unidad y solo se desvía al final de la serie por lo que 29 zonas metropolitanas presentan un proceso de convergencia con poca variabilidad entre sus elementos. Si bien, todos los clubes presentan un comportamiento estable al inicio de la serie es interesante que la totalidad de las sendas rompen su comportamiento estable en el mismo punto, pero no todas lo hacen en el mismo sentido. En el caso de los Clubes 1 y 2 se rompe el comportamiento estable hacia arriba de forma pronunciada, y de forma más suave por parte del Club 3. Esto implica que parte de sus elementos tuvieron un crecimiento considerable y que han comenzado un proceso de regreso a la convergencia,

[9]Para detalles de la implementación computacional consultar "ConvergenceClubs: A Package for Performing the Phillips and Sul's Club Convergence Clustering Procedure" de Roberto Sichera y Pietro Pizzuto (2019), ISSN 2073-4859.

recordemos que no se rechazó la hipótesis de convergencia para ningún club. Sin embargo, este escenario solo le aplica a menos del 30 % de las zonas metropolitanas. De hecho, las 53 zonas metropolitanas restantes tienen valores promedio de convergencia relativa por abajo de la unidad, que son las pertenecientes a los Clubes 4 y 5.

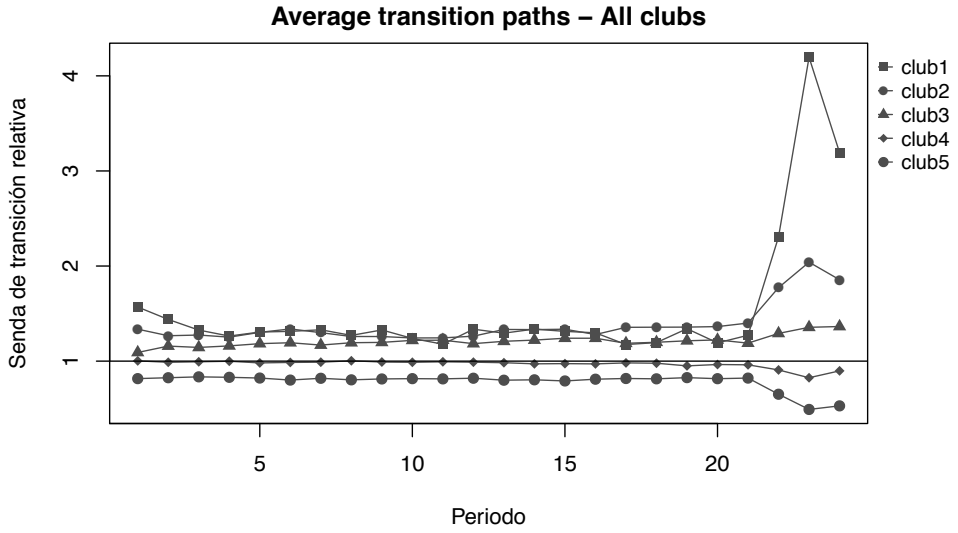


Figura 3.10: Senda de transición promedio de los Clubs.
Fuente: Elaboración propia.

En la Figura 3.10 se observa un cambio comportamiento a partir del periodo 22, que corresponde al año 2013, donde los datos fueron capturados con la tecnología más moderna. Con la finalidad de evaluar el efecto que tiene esa sección de la serie sobre los resultados, se aplicó el algoritmo sobre la serie truncada. Es decir, se eliminaron los datos de 2013 a 2015. Los resultados se muestran en el Cuadro 3.5 donde resaltan dos puntos: la existencia de un elemento divergente y el signo positivo de *beta* para tres de los cinco clubes. Con respecto al elemento divergente es aún más interesante al tratarse de la zona metropolitana del Valle de México, la cual ha sido utilizada como líder en estudios de convergencia (Ayala y Chapa 2011). Si bien fue utilizada desde una perspectiva de entidad federativa y no desde la perspectiva de una zona metropolitana, aún así no deja de llamar la atención este resultado. De entrada suena complicado que la totalidad de las entidades se acerquen al desempeño de la capital dadas las condiciones de infraestructura y el tipo de actividades económicas predominantes en la zona, razón por la cual es importante analizar las dinámicas de crecimiento de manera diferenciada.

En cuanto a los valores del estimador *beta* de los clubes 3, 4 y 5 resalta que aunque es positivo, su valor es cercano a cero, esto se puede entender dado el comportamiento estable que presenta la serie (Figura 3.11). Este resultado indica un proceso de convergencia lento para 44 de las 74 zonas metropolitanas, como ya se ha reportado en trabajos anteriores. Por otra parte, el club 1 y 2 de la serie presentan un valor negativo del estadístico t, que si bien está dentro del valor crítico $> -1{,}65$, lo cumple por poco, así que se concluye que *beta* no es

Cuadro 3.5: Convergence clubs results for truncated series (1992-2012).

Número de Club	**Número ZM**	**Coeficiente $\hat{b}$**	**Error estándar**	**Estadístico t**
1	21	-0.191	0.129	-1.483
2	8	-0.609	0.371	-1.643
3	20	0.031	0.17	0.185
4	12	0.138	0.253	0.546
5	12	0.1	0.215	0.463
Divergente	1	- -	- -	- -

significativamente diferente de cero, es decir se presenta un proceso de estancamiento en el proceso de convergencia de estas zonas metropolitanas.

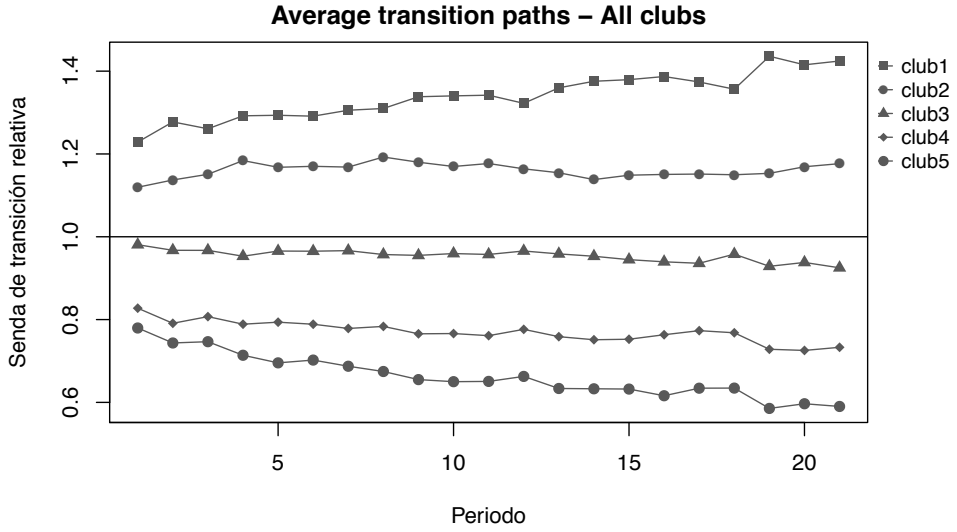


Figura 3.11: Senda de transición promedio los Clubs con serie truncada (1992-2012).

Para evaluar la similitud entre ambas clasificaciones se calculó la medida de homogeneidad descrita anteriormente, la homogeneidad entre la conformación de clubs con la serie completa y la serie truncada es de 0.16. Aunque el valor encontrado es bajo, llama la atención el hecho de que el 95 % de las zonas metropolitanas asignadas al Club 4 y Club 5 para la serie truncada, son asignadas al Club 4 y Club 5 de la serie completa, lo que se podría interpretar como continuidad en el agrupamiento. Este no es el caso para los Clubs 1 a 3, que se distribuyen de forma heterogénea al tomar la serie completa.

3.5. Conclusiones

Los procesos de convergencia económica se deberían dar de forma diferenciada en aquellas regiones en las cuales sus estructuras productivas sean parecidas. Es decir, las regiones deben de tener convergencia a diferentes estados estacionarios. En este trabajo se identifican los clubs de convergencia para las 74 zonas metropolitanas de México utilizando la metodología propuesta por Phillips y Sul (2009). Debido a la falta de información oficial a un nivel de área metropolitana se propone el el índice lumínico per cápita (ILpc) como variable proxy para medir la actividad económica, el cual ha sido utilizado por diversos autores (Ghosh et al., 2009 y Llamosas et al. , 2018). Tras extraer y trabajar con la información lumínica disponible públicamente a través del portal del NOAA se construyó el ILpc desde 1992 a 2015 para las 74 zonas metropolitanas de México.

Los resultados de este trabajo sugieren que las 74 zonas metropolitanas de Mexico se pueden agrupar en 5 clubes de convergencia respecto a la dinámica del Índice Lumínico per capita (*ILpc*). A diferencia de trabajos anteriores no se encontraron elementos divergentes; sin embargo, sí se detectó un club con solo dos elementos que describen un comportamiento notablemente diferente al de los demás.

Si bien la heterogeneidad a nivel estatal en México es un tema conocido, el presente estudio muestra que este fenómeno también se da a nivel intraestatal, ya que 50 % de los estados contienen zonas metropolitanas con más de una dinámica de crecimiento. Las razones que causan esta diversidad pueden provenir desde diferencias en la combinación de sectores productivos de la zona metropolitana en cuestión hasta un uso sesgado de los recursos públicos hacia algunas actividades económicas entre las diferentes zonas. Sin embargo, la identificación de los determinantes de dichas diferencias se encuentra fuera del alcance del presente, pero se vislumbra como un tema relevante de investigación.

De los cinco clubes identificados se encontró un club que podría identificarse como el comportamiento base o normal ya que se distribuye uniformemente en el territorio. Además de ser el más numeroso al concentrar casi el 40 % de las zonas metropolitanas. Por otra parte, dos clubs dividen al país en dos, uno posicionado principalmente en el norte y el otro en el resto del país con similitudes a la región mesoamericana. Los clubs restantes parecen detectar dinámicas muy específicas.

Además, se encontró que la parte final del ILpc, la cual es capturada con una tecnología que permite mayor precisión, muestra comportamientos que tienen un peso importante en el proceso de identificación de los clubs, ya que los agrupamientos al tomar o no en cuenta esta parte de la serie muestran algunas diferencias. Así, al analizar la serie para 1992-2012 se encontró una divergencia de la zona metropolitana del Valle de México, lo cuál resulta interesante ya que se suele utilizar esta entidad en el esquema seguidor-líder en otros estudios (Ayala y Chapa, 2011).

Finalmente, el presente estudio muestra evidencia de la diversidad que caracteriza en muchos aspectos a México, haciendo patente que el crecimiento económico no es la excepción. Dicha diversidad, en nuestra opinión, abona a una sana diversificación resultante de las particularidades de cada región. Sin embargo, para que dichas diferencias rindan frutos es necesario conocerlas, aprovecharlas y fortalecerlas de forma que beneficien a los habitantes de cada región, no desde una visión centralista y plana, sino desde una especializada y específica. Es decir, se requieren de políticas especificas de cada región para impulsar el crecimiento

económicos de éstas, argumento que otros estudios ya han mencionado (Hanson, 1998). Por lo tanto, resulta importante llevar el estudio de la dinámica económica por un proceso de refinamiento de la unidad de estudio. Empezando en un nivel más general, tal que capture las dinámicas predominantes (a nivel país), hasta uno más específico de tipo subnacional, regional o al nivel que permita identificar la dinámica de crecimiento a un nivel más especializado y desagregado.

Conclusiones finales

Este trabajo de tesis abona al estudio empírico de la dinámica macroeconómica al implementar metodologías y recursos provenientes de las ciencias de la información y la geografía. De esta forma se abordan aspectos clásicos como el crecimiento económico o la desigualdad espacial con herramientas pertenecientes al campo de la minería de datos, conceptos como el aprendizaje no supervisado o la utilización de fuentes alternas información. Cabe mencionar que el uso de estas fuentes alternas de información a su vez requieren el dominio de conceptos como el geo posicionamiento y el tratamiento de datos a gran escala. Así, por medio de los trabajos expuestos, se muestra la utilidad (o tal vez necesidad) de expandir los métodos de análisis utilizados para responder preguntas aún sin responder y así entender mejor el estado actual de las sociedades desde un enfoque económico.

En este sentido, el primer capítulo presenta un resultado importante al generar grupos de países por medio de un proceso de generación endógena no supervisado donde, por una parte, se detectan los grupos naturales como los de los países del África subsahariana o la zona euro, pero que también detecta casos especiales como los de Argentina y Uruguay donde presentan un mayor grado de similitud con el grupo de países predominantemente europeos, así como la inclusión de Turquía con países latinoamericanos o la asignación de Haití con países del continente africano. Con la importante característica de que estos resultados fueron generados por datos duros libres de cualquier asignación o manipulación previa. Además, la metodología presentada tiene la flexibilidad de incluir otro tipo de variables. Esto abre la posibilidad de agregar índices de gobernabilidad, corrupción, desigualdad, entre otros, de tal forma que es posible orientar la agrupación hacia temas de interés específico. En el segundo capítulo se analizaron 58 de las 74 zonas metropolitanas de la República Mexicana por medio del índice lumínico como proxy del ingreso, encontrando que no existe evidencia estadística de un proceso de U invertida en la desigualdad espacial de los municipios pertenecientes a las zonas metropolitanas. Sin embargo, se encontró que la variable proxy del ingreso presenta un signo negativo y significativo para todas las zonas metropolitanas, de tal forma que a mayor ingreso se observa menor desigualdad espacial metropolitana. Estos resultados generan nuevas interrogantes en base a la dinámica particular de las zonas metropolitanas, resultados que fueron posibles al utilizar el índice lumínico como proxy del ingreso.

En el tercer capítulo se encontró evidencia de la diversidad presente en las zonas metropolitanas de México, en cuanto a el crecimiento económico se refiere. En este capítulo se encontró que existen cinco clubs de convergencia. De tal forma que la división Norte-Sur no es suficiente para explicar las dinámicas de crecimiento al aumentar el nivel de detalle de entidades federativas a zonas metropolitanas. Estos resultados fueron fruto de la combinación de metodologías de análisis económico modernas, que toman en cuenta la heterogeneidad

de los individuos bajo estudio y el uso de técnicas de minería de datos para la extracción de información relevante. Por lo tanto, se concluye que el acceso a técnicas y metodologías multidisciplinarias permitirán a la ciencia económica responder viejas preguntas además de abrir la posibilidad a plantearse otras totalmente nuevas.

Bibliografía

[1] A. Alesina, S. Michalopoulos, and E. Papaioannou. Ethnic inequality. *Journal of Political Economy*, 124(2):428–488, 2016.

[2] S. Barrios and E. Strobl. The dynamics of regional inequalities. *Regional Science and Urban Economics*, 39(5):575–591, 2009.

[3] R. J. Barro. Economic growth in a cross section of countries. Technical report, National Bureau of Economic Research, 1989.

[4] R. J. Barro and J. W. Lee. A new data set of educational attainment in the world, 1950–2010. *Journal of development economics*, 104:184–198, 2013.

[5] R. J. Barro, X. Sala-i Martin, O. J. Blanchard, and R. E. Hall. Convergence across states and regions. *Brookings papers on economic activity*, pages 107–182, 1991.

[6] N. Basturk, R. Paap, and D. van Dijk. Structural differences in economic growth: An endogenous clustering approach. *Applied Economics*, 44(1):119–134, 2012.

[7] J. G. Brida, J. S. Pereyra, M. P. Anyul, and W. A. Risso. Regímenes de desempeño económico y dualismo estructural en la dinámica de las entidades federativas de méxico, 1970-2006. *Economía Mexicana. Nueva Época*, 22(1):101–149, 2013.

[8] R. M. Campos Vázquez and L. A. Monroy-Gómez-Franco. La relación entre crecimiento económico y pobreza en méxico. *Investigación económica*, 75(298):77–113, 2016.

[9] N. G. D. Center. Version 4 dmsp-ols nighttime lights time series.

[10] R. Cermeño. Decrecimiento y convergencia de los estados mexicanos: Un análisis de panel. *El Trimestre Económico*, pages 603–629, 2001.

[11] X. Chen and W. Nordhaus. A test of the new viirs lights data set: Population and economic output in africa. *Remote Sensing*, 7(4):4937–4947, 2015.

[12] X. Chen and W. D. Nordhaus. Viirs nighttime lights in the estimation of cross-sectional and time-series gdp. *Remote Sensing*, 11(9):1057, 2019.

[13] D. Chiquiar. Why mexico's regional income convergence broke down. *Journal of development Economics*, 77(1):257–275, 2005.

[14] I. N. de Estadística y Geografía. Marco geoestadístico.

[15] G. Doppelhofer, R. I. Miller, et al. Determinants of long-term growth: A bayesian averaging of classical estimates (bace) approach. *The American Economic Review*, 94(4):813–835, 2004.

[16] C. Duan and B. Chen. Analysis of global energy consumption inequality by using lorenz curve. *Energy Procedia*, 152:750–755, 2018.

[17] S. N. Durlauf and P. A. Johnson. Multiple regimes and cross-country growth behaviour. *Journal of applied econometrics*, 10(4):365–384, 1995.

[18] C. D. Elvidge, K. Baugh, M. Zhizhin, F. C. Hsu, and T. Ghosh. Viirs night-time lights. *International Journal of Remote Sensing*, 38(21):5860–5879, 2017.

[19] G. Esquivel. Convergencia regional en méxico, 1940-1995. *El trimestre económico*, pages 725–761, 1999.

[20] G. Esquivel and M. Messmacher. Sources of regional (non) convergence in mexico. *El Colegio de México y Banco de Mexico, Mexico City*, 2002.

[21] R. C. Feenstra, R. Inklaar, and M. P. Timmer. The next generation of the penn world table. *American economic review*, 105(10):3150–82, 2015.

[22] E. A. A. Gaytán, J. C. C. Cantú, and J. D. M. Hernández. Una reconsideración sobre la convergencia regional en méxico. *Estudios Económicos*, pages 217–247, 2011.

[23] V. German-Soto and N. S. Garza. Testing stochastic convergence among mexican states: A polynomial regression analysis. *Journal of Reviews on Global Economics*, 5:36–47, 2016.

[24] T. Ghosh, S. Anderson, R. L. Powell, P. C. Sutton, and C. D. Elvidge. Estimation of mexico's informal economy and remittances using nighttime imagery. *Remote Sensing*, 1(3):418–444, 2009.

[25] W. B. Group. *World development indicators 2014*. World Bank Publications, 2014.

[26] G. H. Hanson. Regional adjustment to trade liberalization. *Regional science and urban economics*, 28(4):419–444, 1998.

[27] J. A. Hartigan and M. A. Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108, 1979.

[28] J. V. Henderson, A. Storeygard, and D. N. Weil. Measuring economic growth from outer space. *American economic review*, 102(2):994–1028, 2012.

[29] S. S. Kortum. Research, patenting, and technological change. *Econometrica: Journal of the Econometric Society*, pages 1389–1419, 1997.

[30] S. Kuznets. Economic growth and income inequality. *The American economic review*, 45(1):1–28, 1955.

[31] S. Lawrence, Q. Liu, and V. M. Yakovenko. Global inequality in energy consumption from 1980 to 2010. *Entropy*, 15(12):5565–5579, 2013.

[32] C. Lessmann. Spatial inequality and development—is there an inverted-u relationship? *Journal of development economics*, 106:35–51, 2014.

[33] C. Lessmann and A. Seidel. Regional inequality, convergence, and its determinants–a view from outer space. *European Economic Review*, 92:110–132, 2017.

[34] R. Levine and D. Renelt. A sensitivity analysis of cross-country growth regressions. *The American economic review*, pages 942–963, 1992.

[35] I. Llamosas-Rosas, E. Rangel González, and M. Sandoval Bustos. Medición de la actividad económica en las principales zonas turísticas de playa en méxico a través de la luminosidad fotografiada desde el espacio. Technical report, Working Papers, 2018.

[36] J. Lluís Carrion-I-Silvestre and V. German-Soto. Stochastic convergence amongst mexican states. *Regional Studies*, 41(4):531–541, 2007.

[37] M. V. López. Desigualdad regional en el centro de méxico. una exploración espacial de la productividad en el nivel municipal durante el período 1988-2003. *Investigaciones Regionales-Journal of Regional Research*, (13):5–34, 2008.

[38] R. B. López and J. M. C. Arriaga. Análisis de transición dinámica: un enfoque no paramétrico aplicado a la región centro de méxico (1988-2003). *Problemas del Desarrollo*, pages 7–44, 2011.

[39] N. G. Mankiw, D. Romer, and D. N. Weil. A contribution to the empirics of economic growth. *The quarterly journal of economics*, 107(2):407–437, 1992.

[40] A. R. Márquez González and H. R. Ramírez Partida. Desigualdad y pobreza en los municipios del estado de nayarit. análisis desde el crecimiento económico (1990-2010). *Espiral (Guadalajara)*, 27(77):81–129, 2020.

[41] W. H. Masanjala and C. Papageorgiou. Rough and lonely road to prosperity: a reexamination of the sources of growth in africa using bayesian model averaging. *Journal of Applied Econometrics*, 23(5):671–682, 2008.

[42] A. Mendoza-Velázquez, V. German-Soto, M. Monfort, and J. Ordóñez. Club convergence and inter-regional inequality in mexico, 1940-2015. *Applied Economics*, 52(6):598–608, 2020.

[43] A. Mveyange. Night lights and regional income inequality in africa. Technical report, WIDER Working Paper, 2015.

[44] S. Nahar and B. Inder. Testing convergence in economic growth for oecd countries. *Applied Economics*, 34(16):2011–2022, 2002.

[45] W. Nordhaus and X. Chen. A sharper image? estimates of the precision of nighttime lights as a proxy for economic statistics. *Journal of Economic Geography*, 15(1):217–246, 2015.

[46] W. R. Ochoa. Convergencia económica interestatal en méxico, 1900-2004. *Análisis Económico*, 25(58):7–34, 2010.

[47] Y. Oswald, A. Owen, and J. K. Steinberger. Large inequality in international and intranational energy footprints between income groups and across consumption categories. *Nature Energy*, 5(3):231–239, 2020.

[48] P. C. Phillips and D. Sul. Transition modeling and econometric convergence tests. *Econometrica*, 75(6):1771–1855, 2007.

[49] P. C. Phillips and D. Sul. Economic transition and growth. *Journal of Applied Econometrics*, 24(7):1153–1185, 2009.

[50] G. Psacharopoulos. Returns to investment in education: A global update. *World development*, 22(9):1325–1343, 1994.

[51] E. Rangel-Gonzalez and I. Llamosas-Rosas. An alternative method to measure non-registered economic activity in mexico using satellite nightlights. 2019.

[52] D. Rodríguez Benavides, F. López Herrera, and M. Á. Mendoza González. Clubs de convergencia regional en méxico: un análisis a través de un modelo no lineal de un solo factor. *Investigaciones Regionales – Journal of Regional Research*, 34:7–22, 2016.

[53] L. I. Rodríguez-Gámez and J. A. Cabrera-Pereyra. Análisis espacial de las dinámicas de crecimiento económico en méxico (1999-2009). *Economía, sociedad y territorio*, 17(55):709–741, 2017.

[54] L. I. Rodríguez-Gámez and J. A. Cabrera-Pereyra. Convergencia económica entre municipios mexicanos: un enfoque de parámetros locales. *Ensayos. Revista de economía*, 39(2):143–186, 2020.

[55] E. Rodríguez-Oreggia. *Polarization of income under structural changes: winners and losers of regional growth in Mexico*. London School of Economics Department of Geography and Environment, 2002.

[56] E. Rodríguez-Oreggia. Regional disparities and determinants of growth in mexico. *The Annals of Regional Science*, 39(2):207–220, 2005.

[57] A. Rosenberg and J. Hirschberg. V-measure: A conditional entropy-based external cluster evaluation measure. In *Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL)*, pages 410–420, 2007.

[58] X. X. Sala-i Martin. I just ran two million regressions. *The American Economic Review*, pages 178–183, 1997.

[59] A. Sánchez Almanza. Crecimiento económico, desigualdad y pobreza: una reflexión a partir de kuznets. *Problemas del desarrollo*, 37(145):11–30, 2006.

[60] N. A. Sanén and L. Q. Romero. Crecimiento económico, convergencia y concentración económica espacial en las entidades federativas de méxico 1970-2008. *Investigaciones Regionales-Journal of Regional Research*, (18):83–106, 2010.

[61] C. Sedesol. Inegi, delimitación de las zonas metropolitanas de méxico 2015, 2018.

[62] P. S. Segerstrom. Endogenous growth without scale effects. *American Economic Review*, pages 1290–1310, 1998.

[63] R. Sichera and P. Pizzuto. Convergenceclubs: A package for performing the phillips and sul's club convergence clustering procedure. *The R Journal*, 11(2):142–151, 2019.

[64] R. M. Solow. A contribution to the theory of economic growth. *The quarterly journal of economics*, 70(1):65–94, 1956.

[65] R. M. Solow. Technical change and the aggregate production function. *The review of Economics and Statistics*, pages 312–320, 1957.

[66] P. C. Sutton, C. D. Elvidge, T. Ghosh, et al. Estimation of gross domestic product at sub-national scales using nighttime satellite imagery. *International Journal of Ecological Economics & Statistics*, 8(S07):5–21, 2007.

[67] T. W. Swan. Economic growth and capital accumulation. *Economic record*, 32(2):334–361, 1956.

[68] D. R. Valenzuela-Vega, E. M. Luna-Domínguez, and J. C. C. Cantú. Clubes de convergencia metropolitana en méxico: un análisis a través del índice lumínico. *Revista de Economía, Facultad de Economía, Universidad Autónoma de Yucatán*, 38(97):35–64, 2021.

[69] C. J. Vilalta et al. Una aplicación del análisis espacial al estudio de las diferencias regionales del ingreso en méxico. *Economía, sociedad y territorio*, 4(14):317–340, 2003.

[70] J. G. Williamson. Regional inequality and the process of national development: a description of the patterns. *Economic development and cultural change*, 13(4, Part 2):1–84, 1965.

Anexo

Listado de las Zonas Metropolitanas con el resultado del análisis de convergencia para el periodo 1992-2015 y 1992-2012.

Cuadro 3.6: Listado de Zonas Metropolitanas (parte 1)

n	Zona Metropolitana	Entidad	Club (1992-2015)	Club (1992-2012)
1	Tehuantepec	Oaxaca	1	1
2	Coatzacoalcos	Veracruz	1	3
3	Tula	Hidalgo	2	1
4	Minatitlán	Veracruz	2	1
5	Nuevo Laredo	Tamaulipas	2	1
6	Villahermosa	Tabasco	2	3
7	Juárez	Chihuahua	2	3
8	Poza Rica	Veracruz	2	1
9	Piedras Negras	Coahuila	3	1
10	Chihuahua	Chihuahua	3	1
11	Mexicali	Baja California	3	1
12	Monclova-Frontera	Coahuila	3	1
13	La Paz	Baja California Sur	3	1
14	La Laguna	Coahuila	3	2
15	Reynosa	Tamaulipas	3	2
16	Hermosillo	Sonora	3	3
17	Guaymas	Sonora	3	1
18	Querétaro	Querétaro	3	3
19	Saltillo	Coahuila	3	3
20	Campeche	Campeche	3	1
21	Monterrey	Nuevo León	3	4
22	Mérida	Yucatán	4	3
23	Matamoros	Tamaulipas	4	3
24	Tampico	Tamaulipas	4	4
25	Delicias	Chihuahua	4	1
26	Nogales	Sonora	4	5
27	Zacatecas-Guadalupe	Zacatecas	4	2
28	Pachuca	Hidalgo	4	2
29	Ensenada	Baja California	4	1
30	Celaya	Guanajuato	4	3
31	Guanajuato	Guanajuato	4	1
32	Cancún	Quintana Roo	4	5
33	Aguascalientes	Aguascalientes	4	4
34	Ciudad Victoria	Tamaulipas	4	3
35	Mazatlán	Sinaloa	4	4
36	Culiacán	Sinaloa	4	3
37	Veracruz	Veracruz	4	4

Cuadro 3.7: Listado de Zonas Metropolitanas (parte 2)

n	Zona Metropolitana	Entidad	Club (1992-2015)	Club (1992-2012)
38	Tlaxcala-Apizaco	Tlaxcala	4	3
39	Cuautla	Morelos	4	2
40	León	Guanajuato	4	5
41	San Luis Potosí	San Luis Potosí	4	4
42	San Francisco del Rincón	Guanajuato	4	2
43	Hidalgo del Parral	Chihuahua	4	2
44	Tijuana	Baja California	4	5
45	La Piedad-Pénjamo	Michoacán	4	1
46	Colima-Villa de Álvarez	Colima	4	2
47	Puebla-Tlaxcala	Puebla	4	5
48	Guadalajara	Jalisco	4	5
49	Puerto Vallarta	Jalisco	4	3
50	Tianguistenco	México	4	3
51	Tuxtla Gutiérrez	Chiapas	5	5
52	Chetumal	Quintana Roo	5	1
53	Durango	Durango	5	3
54	Tulancingo	Hidalgo	5	3
55	Toluca	México	5	4
56	Cuernavaca	Morelos	5	4
57	Orizaba	Veracruz	5	4
58	Valle de México	Ciudad de México	5	0
59	Morelia	Michoacán	5	3
60	Ocotlán	Jalisco	5	1
61	Moroleón-Uriangato	Guanajuato	5	3
62	Oaxaca	Oaxaca	5	5
63	Tepic	Nayarit	5	4
64	Tehuacán	Puebla	5	5
65	Tecomán	Colima	5	1
66	Rioverde	San Luis Potosí	5	1
67	Acapulco	Guerrero	5	3
68	Chilpancingo	Guerrero	5	3
69	Córdoba	Veracruz	5	5
70	Acayucan	Veracruz	5	1
71	Xalapa	Veracruz	5	5
72	Tapachula	Chiapas	5	4
73	Zamora	Michoacán	5	4
74	Teziutlán	Puebla	5	5