UNIVERSIDAD DE OVIEDO

Departamento de Economía Aplicada

The spatial scale and the empirics of regional economics: growth, convergence and agglomerations

Programa de Doctorado
"Economía y Sociología de la Globalización"
(Mención de Calidad)

Alberto Díaz Dapena

(2016)

Universidad de Oviedo
Departamento de Economía Aplicada

The spatial scale and the empirics of regional economics: growth, convergence and agglomerations

Memoria que, para la obtención del grado de doctor, presenta

Alberto Díaz Dapena

Programa de Doctorado
"Economía y Sociología de la Globalización"
(Mención de Calidad)

Programa de Formación del Profesorado Universitario
(Ref. FPU12/00236)

Bajo la dirección de los profesores

Esteban Fernández Vázquez
Fernando Rubiera Morollón

RESUMEN DEL CONTENIDO DE TESIS DOCTORAL

1.- Título de la Tesis	
Español/Otro Idioma:	Inglés: THE SPATIAL SCALE AND THE EMPIRICS OF REGIONAL ECONOMICS: GROWTH, CONVERGENCE AND AGGLOMERATIONS.

2.- Autor	
Nombre: ALBERTO DIAZ DAPENA	DNI/Pasaporte/NIE:
Programa de Doctorado: Economía Aplicada y Sociología de la Globalización	
Órgano responsable: ECONOMIA APLICADA	

RESUMEN (en español)

La elección de la escala espacial para llevar a cabo análisis empíricos debería ser un paso fundamental en la economía regional. Aunque los investigadores en economía regional han prestado atención durante décadas (Openshaw, 1984) al papel que juega la escala, que debe ser consistente con los supuestos y el marco teórico, la tradicional falta de información desagregada geográficamente ha obligado a los economistas regionales a usar datos agregados en grandes regiones administrativas para sus análisis empíricos. Las técnicas estadísticas y econométricas han experimentado mejoras importantes en la última década permitiendo la medición de fenómenos socioeconómicos cada vez más complejos (Islam, 2003). Esas mejoras nos permiten tratar temas como la dependencia espacial, las relaciones no lineales o la heterogeneidad, y hacer inferencias en relaciones económicas e impactos de una manera mucho más precisa. La gran variedad de técnicas econométricas disponibles hace que la elección de una estrategia de estimación específica sea una decisión relevante que debe ser adecuadamente justificada en toda investigación empírica. Las bases de datos también han mejorado significativamente en los últimos años, pasando a tener información más precisa y con mayor desagregación espacial para la mayoría de las variables. Sin embargo, no se presta tanta atención a que la escala espacial en el análisis empírico debe depender de la cuestión que se quiera investigar y del marco teórico. Esta decisión puede afectar a los resultados incluso más que la elección de una versión concreta de un estimador.

Un motivo que explique esta tendencia de ignorar la importancia de la unidad espacial de la investigación y la escala de los análisis podría ser la influencia de la Economía Neoclásica. Este marco teórico básicamente ignora la importancia de la escala en el análisis económico. Sus modelos más conocidos se basan en rendimientos decrecientes en los factores (véase Solow, 1956; Mankiw et al., 1992), la movilidad de factores y la difusión del conocimiento como se explica en Barro et al. (1991), Barro and Sala-I-Martin (1992) o Sala-I-Martin (1994). Dichos modelos macroeconómicos están construidos sobre la agregación de agentes de la economía representativos, independientes y homogéneos. Por lo tanto, los modelos teóricos operan sin importar la escala o el concepto de región usado en el análisis.

Pero esta conclusión no es robusta cuando los modelos básicos de crecimiento incluyen un proceso con interacciones locales. Por ejemplo, Lucas (1988) enfatiza la acumulación de capital humano a través de la educación y del “aprender haciendo”. Además, Lucas (2001) desarrolla la teoría de que las zonas centrales acumulan capital humano, generando externalidades positivas, lo que refuerza la idea de un incentivo en las áreas urbanas hacia la acumulación del conocimiento y los efectos spillover. Un proceso con estas características sería imposible de distinguir en una escala agregada. En consecuencia, la escala sí afecta a los resultados y las conclusiones. Este tipo de procesos en la escala local son resaltados con modelos posteriores que introducen mecánicas acumulativas en las regiones -véase Romer (1990 y 1994), Myrdal (1957), Hirschman (1958), Kaldor (1957), Dixon y Thirlwall (1975)-.

Hay un debate importante sobre la naturaleza de las externalidades positivas creadas en las áreas urbanas. La literatura teórica más aceptada podría ser la propuesta por Marshal (1890) sobre las economías de aglomeración. Este fenómeno puede dividirse en dos: las economías de localización y las economías de urbanización. Las economías de localización describen las

externalidades causadas por la interacción entre actividades del mismo sector, que atrae trabajadores especializados, proveedores y acceso al conocimiento. Las economías de urbanización explican que la concentración de actividades desencadena las relaciones entre individuos -conocidas como capital social- e infraestructuras -por ejemplo, ferrocarriles, centros de innovación y hospitales-. Parr (2002) resume esta idea con una delimitación moderna y clara del concepto de economías de aglomeración.

La literatura sobre economías de aglomeración ha evolucionado desde la década de 1950 con contribuciones que explican los procesos desarrollados en las áreas urbanas, como por ejemplo Isard (1956), Ziprf (1949), Jacobs (1969) y Porter (1990). Estas ideas se resumen en el modelo centro-periferia (véase Krugman, 1991; Krugman y Venables, 1995; Fujita et al., 2001). Una nueva literatura empírica y teórica ha surgido de ese modelo, conocida como Nueva Geografía Económica (NGE, en adelante). Según Krugman (1998), la NGE explica la economía usando modelos dinámicos con un equilibrio general, que se obtiene a través de la competencia entre las fuerzas de dispersión y de concentración con economías de escala. De acuerdo con la literatura de la NGE: (i) hay incentivos a concentrar fuertemente la producción en áreas centrales, y (ii) los procesos de especialización y comercio intra-regionales y entre países refuerzan los procesos de concentración y, en consecuencia, de divergencia (véase también Baldwin y Forslid, 2003; Ottaviano y Thisse, 2004 o Behrens y Thisse, 2007).

En resumen, la escala geográfica no es relevante según la Economía Neoclásica. La falta de interacción entre los agentes representativos en este modelo genera homogeneidad en todas las escalas. Sin embargo, la NGE se centra en las fuerzas centrífugas, que crean concentración de las actividades y heterogeneidad. Desde este punto de vista, el concepto de región y la agregación no son neutrales. Diferentes clasificaciones del territorio pueden llevar a la eliminación de información relevante en las relaciones entre las zonas centrales y periféricas.

El supuesto de información homogénea en grandes áreas puede ser extremadamente arbitrario según el criterio de agregación. Las bases de datos gubernamentales han estado tradicionalmente limitadas por unidades administrativas agregadas debido a la falta de información detallada. Sin embargo, estas regiones han sido frecuentemente diseñadas por razones no económicas sino históricas o políticas. Así que los datos agregados que basan este tipo de clasificación es una combinación de diferentes unidades económicas. Esta ausencia de información desagregada a nivel local implica que los análisis regionales no tenían más opción que usar estos datos, a pesar de los problemas de agregación. Pero los análisis de la NGE requieren una especial atención a las dinámicas del nivel local más que de las áreas nacionales. El análisis de dichas dinámicas utilizando información agregada que no distingue entre áreas urbanas y rurales puede carecer de robustez.

Aun así, la disponibilidad de datos ha aumentado en los últimos años, existiendo un número creciente de bases de datos con información desagregada -o incluso con datos individuales-. Hoy en día es posible adaptar los datos al nivel de agregación o al concepto de región económica más apropiado a nuestra investigación. La elección de una escala adecuada en el análisis económico y sus consecuencias en los resultados podría volverse tan importante como la de un estimador correcto.

El objetivo principal de esta tesis es explorar el papel jugado por la escala espacial en los análisis empíricos de la economía regional. Estudia cómo un nivel geográfico no consistente con los supuestos puede afectar a las conclusiones finales y llevar a unos resultados sin sentido -o, al menos, no tan claros como deberían ser cuando la escala es elegida correctamente-. Aunque esta idea puede ser aplicada a todos los análisis espaciales, es en el campo del crecimiento económico y las diferencias económicas entre territorios donde puede darse una mayor infraestimación de la importancia de la escala. Por lo tanto, esta tesis presta especial atención a la importancia de la elección del nivel espacial en los estudios de crecimiento y convergencia, así como en análisis de productividad.

La tesis empieza poniendo el foco en el fenómeno de desigualdad entre territorios usando el conocido análisis de β-convergencia. El análisis de β-convergencia es particularmente interesante para el objeto de esta tesis, pues es un campo de la literatura está directamente conectado tanto con el marco Neoclásico como con la NGE. La NGE sugiere que las desigualdades regionales en PIB per cápita surgen debido a las diferencias entre las áreas rurales y urbanas en términos de capital humano y externalidades de la actividad, mientras que las teorías neoclásicas predicen homogeneidad en los niveles de PIB per cápita entre regiones. Esta primera parte de la tesis estudia el problema de agregación espacial de los datos en la estimación de ecuaciones de β-convergencia. Se basa en estudios previos que ya han prestado atención al efecto de la agregación, como en el trabajo de Theil (1954) para el caso general con modelos de regresión lineal o, más recientemente, de Arbia y Petrarca (2011) para el caso de datos dependientes espacialmente. Además, se introduce explícitamente en el análisis la naturaleza jerárquica de los datos económicos en lo que respecta a unidades espaciales y analiza la importancia de cada nivel en el proceso mediante un enfoque econométrico de análisis multinivel.

La segunda parte de esta tesis estudia las aglomeraciones urbanas y cómo las dinámicas entre los territorios rurales y urbanos pueden afectar al resultado, y después continua con el análisis de la productividad y sus relaciones con la densidad de población. Sigue la literatura reciente, que ha prestado atención a cuantificar el impacto de las economías de aglomeración sobre la productividad -véase, por ejemplo, Rosenthal y Strange (2001), para un análisis más extenso o Ciccone y Hall (1996); Combes (2000), Combes et al. (2008), o Artis et al. (2012)-. Más recientemente, Combes y Gobillon (2015) han revisado las contribuciones más relevantes de las economías de aglomeración, que cubren tanto los intentos de estimarlos en base a datos regionales agregados como las estrategias más recientes que utilizan datos individuales. Mientras que esta última opción puede considerarse preferible cuando hay datos disponibles, en ocasiones la falta de información observable a nivel individual hace necesaria la estimación utilizando alguna media a la escala espacial dada. Si ese es el caso, utilizar datos que promedian unidades geográficas altamente desagregadas permite considerar la escala espacial adecuada para medir las economías de aglomeración, en tanto que los datos agregados espacialmente implican asumir un alto nivel de homogeneidad intra-regional.

RESUMEN (en Inglés)

The choice of the spatial scale for conducting the empirical analysis should be a fundamental initial step in regional economics. Even when scholars in regional economics have paid attention for decades (Openshaw, 1984) to the role played by this scale, which must be consistent with the assumptions and the theoretical framework, the traditional lack of geographically disaggregated data has forced regional economists to use information aggregated to a large scale and use administrative large regions in their empirical analysis. Statistical and econometric techniques have experienced important improvements in the last decades for the measurement of increasingly complex socio-economic phenomena (Islam, 2003). These improvements allow us to deal with issues as spatial dependence, nonlinearities or heterogeneity and making inferences on economic relationships and impacts in a much more accurate way. The large variety on the available econometric techniques in the regional economist’s toolkit makes the choice of the specific estimation strategy a relevant decision that should be conveniently justified in every empirical research. Databases also have improved significantly in the last decades, having more precise and more spatially disaggregated information for most of the variables. However, the possibilities of considering different spatial scales for the empirical analysis that should depend on the research question or the theoretical framework are still not generalized, even when this decision can influence the results more than the selection of a specific version of an estimator.

A possible reason that explains this tendency to ignore the relevance of the spatial unit of investigation and the scale of the analysis could be the influence of Neoclassical Economics. This theoretical framework basically neglected the importance of the scale in economic analysis. Their well-known models are based on decreasing returns to scale in factors (see Solow, 1956; Mankiw et al., 1992), the mobility of factors, and the spread of knowledge as explained in Barro et al. (1991), Barro and Sala-I-Martin (1992) or Sala-I-Martin (1994). These macroeconomic models are built on the aggregation of representative, independent and homogeneous agents of the economy. So, theoretical models should operate no matter the scale or the concept of region used in the analysis.

Nevertheless, this conclusion is not robust when basic growth models include a process with local interactions. For example, Lucas (1988) emphasizes human capital accumulation through

schooling and learning-by-doing. In addition, Lucas (2001) develops the theory that central zones accumulate human capital, generating positive externalities. It remarks the idea of a positive incentive in the urban areas towards accumulation of knowledge and spillovers. A process with these characteristics would be impossible to distinguish within an aggregated scale. Consequently, the scale does have an influence on the results and the conclusions. This type of processes in the local scale are highlighted with later models that introduce accumulative mechanics in the regions -see Romer (1990 y 1994), Myrdal (1957), Hirschman (1958), Kaldor (1957), Dixon and Thirlwall (1975).

There is an important discussion about the nature of the positive externalities created within the urban areas. The most accepted theoretical literature could be the proposal of Marshal (1890) about the agglomeration economies. This phenomenon can be divided into two: the location economies and the urbanization economies. The location economies describe the externalities caused by the interaction of activities of the same sector. This interaction attracts specialized workers, suppliers and access to knowledge. The urbanization economies explain that the concentration of the activity triggers relationships between individuals –known as social capital– and infrastructures –e.g. railways, innovation centers or hospitals–. Parr (2002) summarizes this idea with a modern and clear delimitation of the concept of agglomeration economies.

The literature of the agglomeration economies has evolved from the 1950's with more contributions that explain the processes developed in the urban areas, such as Isard (1956), Ziprf (1949), Jacobs (1969) and Porter (1990). These ideas are summarized in the core-periphery model (see, Krugman, 1991; Krugman and Venables, 1995; Fujita et al., 2001). A new empirical and theoretical literature emerged from this model, known as New Economic Geography (NEG hereinafter). According to Krugman (1998), NEG explains the economy by using dynamic models with a general equilibrium. The equilibrium is obtained through a competition between forces of dispersion and concentration with scale economies. According to NEG literature: (i) there are incentives to largely concentrate the production in the central areas, and (ii); the intra-regional and inter-country processes of specialization and trade reinforce the processes of concentration and, in consequence, of divergence (see also Baldwin and Forslid, 2003; Ottaviano and Thisse, 2004 or Behrens and Thisse, 2007).

To sum up, geographical scale is not relevant according to the Neoclassical Economics. The lack of interaction between the representative agents in this models generates homogeneity in all the scales. However, the NEG focuses on the centrifugal forces, which create concentration of the activities and heterogeneity. From this point of view, the concept of region and the aggregation are not neutral. Different classifications of the territory could lead to the elimination of valuable information on the relationship between central and peripheral locations.

The assumption of homogeneous data for wide areas can be extremely arbitrary depending on the aggregation criteria. Databases from governments have been traditionally limited to aggregated administrative units due to the lack of detailed information. However, these regions have been usually designed by not economic but historical or political reasons. So the aggregated data that base on this type of classification is mix different economic units. Depending on the research question, it can undermine the economic analysis. This absence of disaggregated data in a local scale implies that the regional analysis had no option but to use this information, despite the problems of aggregation. However, NEG analysis requires an especial attention to the dynamics of the local level rather than national areas. The analysis of these dynamics using aggregated information that do not distinguish between urban and rural areas may lack of robustness.

Nevertheless, the availability of data has grown in the last years. There is an increasingly amount of databases with disaggregated information –or even individual data–. Nowadays, it is possible to adapt the data to the level of aggregation or to the concept of economic region more appropriate to our research analysis. The choice of a suitable scale in the economic analysis and its consequences in the results could become as important as the choice of the correct estimator.

The central aim of this thesis is to explore the role played by the spatial scale in the empirics of regional economics. It studies how a geographical scale not consistent with the assumptions could affect the final conclusions and lead to obtain meaningless results –or, at least, not as clear as they could be when the scale is properly chosen-. Although this idea could be applied to all spatial analysis is in the field of economic growth and territorial economic differences where most relevant underestimations of the relevance of the scale could be happening. So, this thesis particularly pays attention to the relevance of the election of the spatial scale in growth and convergence studies as well as in productivity analysis.

The thesis starts focusing on the phenomenon of inequalities between territories using the well-known β-convergence analysis. β-convergence analysis is particularly interesting for the aim of the thesis, since it is a field of the literature directly connected with both the Neoclassical framework and the NEG. NEG suggests that regional inequalities in GDPpc emerges due to the differences between rural and urban areas in terms of human capital and externalities of the activity, while neoclassical theories predict homogeneity of the levels of GDPpc across regions. This first part of this thesis studies the problem of spatial aggregation of data when estimating β-convergence equations. It bases on previous studies that have already called the attention to the effect of the aggregation, like in the work by Theil (1954) for the general case on linear regression models or, more recently, by Arbia and Petrarca (2011) for the case of spatially dependent data. Additionally, it explicitly introduces in the analysis the hierarchical nature of economic data when referring to spatial units and analyzes the importance of each level in the process by using econometric approach of multilevel analysis.

The second part of this dissertation studies urban agglomerations and how the dynamics between rural and urban territories can affect the results and then it continues with the analysis of the productivity and its relationship with population density. It follows the recent literature, which has paid attention to quantify the impact of agglomeration economies on productivity –see, for example, Rosenthal and Strange (2001), for an extensive review or Ciccone and Hall (1996); Combes (2000), Combes et al. (2008), or Artis et al. (2012)–. More recently, Combes and Gobillon (2015) have reviewed the most relevant contributions to the empirics of agglomeration economies, which covers both the attempts to estimate them basing on aggregated regional data to the more recent strategies that use individual data. While this last option is arguably preferable when data are available, sometimes lack of observable information at individual level makes necessary the estimation basing on some average at a given spatial scale. If this is the case, using data that average highly disaggregated geographical units allows for considering an appropriate spatial scale to measure agglomeration economies, since spatially aggregated data imply assuming a high level of intra-regional homogeneity.

A mi familia,

Resumen en español

La elección de la escala espacial para llevar a cabo análisis empíricos debería ser un paso fundamental en la economía regional. Aunque los investigadores en economía regional han prestado atención durante décadas (Openshaw, 1984) al papel que juega la escala, que debe ser consistente con los supuestos y el marco teórico, la tradicional falta de información desagregada geográficamente ha obligado a los economistas regionales a usar datos agregados en grandes regiones administrativas para sus análisis empíricos. Las técnicas estadísticas y econométricas han experimentado mejoras importantes en la última década permitiendo la medición de fenómenos socioeconómicos cada vez más complejos (Islam, 2003). Esas mejoras nos permiten tratar temas como la dependencia espacial, las relaciones no lineales o la heterogeneidad, y hacer inferencias en relaciones económicas e impactos de una manera mucho más precisa. La gran variedad de técnicas econométricas disponibles hace que la elección de una estrategia de estimación específica sea una decisión relevante que debe ser adecuadamente justificada en toda investigación empírica. Las bases de datos también han mejorado significativamente en los últimos años, pasando a tener información más precisa y con mayor desagregación espacial para la mayoría de las variables. Sin embargo, no se presta tanta atención a que la escala espacial en el análisis empírico debe depender de la cuestión que se quiera investigar y del marco teórico. Esta decisión puede afectar a los resultados incluso más que la elección de una versión concreta de un estimador.

Un motivo que explique esta tendencia de ignorar la importancia de la unidad espacial de la investigación y la escala de los análisis podría ser la influencia de la Economía Neoclásica. Este marco teórico básicamente ignora la importancia de la escala en el análisis económico. Sus modelos más conocidos se basan en rendimientos decrecientes en los factores (véase Solow, 1956; Mankiw et al., 1992), la movilidad de factores y la difusión del conocimiento como se explica en Barro et al. (1991), Barro and Sala-I-Martin (1992) o Sala-I-Martin (1994). Dichos modelos macroeconómicos están construidos sobre la agregación de agentes de la economía representativos, independientes y homogéneos. Por lo tanto, los modelos teóricos operan sin importar la escala o el concepto de región usado en el análisis.

Sin embargo, esta conclusión no es robusta cuando los modelos básicos de crecimiento incluyen un proceso con interacciones locales. Por ejemplo, Lucas (1988) enfatiza la acumulación de capital humano a través de la educación y del “aprender haciendo". Además, Lucas (2001) desarrolla la teoría de que las zonas centrales acumulan capital humano, generando externalidades positivas, lo que refuerza la idea de un incentivo positivo en las áreas urbanas hacia la acumulación del conocimiento y los efectos spillover. Un proceso con estas características sería imposible de distinguir en una escala agregada. En consecuencia, la escala sí afecta a los resultados y las conclusiones.

Este tipo de procesos en la escala local son resaltados con modelos posteriores que introducen mecánicas acumulativas en las regiones. Romer (1990 y 1994) explica que las empresas en áreas urbanas generan progreso técnico endógeno que atrae capital humano y genera externalidades positivas. Myrdal (1957a) y Hirschman (1958) también explican que los territorios siguen un proceso acumulativo. De acuerdo a su investigación, ubicaciones exitosas activan economías a escala internas y externas. Como resultado, atraen factores de regiones subdesarrolladas e incrementan el proceso de economías de escala, que es formalizado en Kaldor (1957), Dixon y Thirlwall (1975). Su investigación explica que, debido a la ley de Verdoom, una región puede desarrollar procesos acumulativos. La producción en este modelo estimula la productividad. En un entorno competitivo, ese crecimiento de la productividad reduce los precios y, en consecuencia, aumenta la demanda. El resultado es un nuevo crecimiento de la producción. Tal y como explican, la interacción en el nivel local genera esas economías de escala. La escala geográfica es, por lo tanto, un elemento clave para analizar este tipo de mecanismos.

Hay un debate importante sobre la naturaleza de las externalidades positivas creadas en las áreas urbanas. La literatura teórica más aceptada podría ser la propuesta por Marshal (1890) sobre las economías de aglomeración. Este fenómeno puede dividirse en dos: las economías de localización y las economías de urbanización. Las economías de localización describen las externalidades causadas por la interacción entre actividades del mismo sector, que atrae trabajadores especializados, proveedores y acceso al conocimiento. Las economías de urbanización explican que la concentración de actividades desencadena las relaciones entre individuos -conocidas como capital social- e infraestructuras -por ejemplo, ferrocarriles, centros de innovación y hospitales-. Parr (2002) resume esta idea con una delimitación moderna y clara del concepto de economías de aglomeración.

La literatura sobre economías de aglomeración ha evolucionado desde la década de 1950 con contribuciones que explican los procesos desarrollados en las áreas urbanas. Isard (1956) modifica el modelo de Christaller (1933) y muestra cómo la jerarquía de los centros urbanos (véase Zipf, 1949) crece para proveer de bienes. De acuerdo con su teoría, un bien se convierte en central cuando hay gente suficiente en un radio. Jacobs (1969) explica que las ciudades generan innovación debido a la interacción de personas de distintos sectores, y Porter (1990) indica que las empresas pueden mejorar su ventaja competitiva formando parte de una red de empresas e instituciones. Con esta estructura, la red consigue proveedores y trabajadores especializados (véase Duranton y Puga, 2000; Glaeser, 1998 o Glaeser, 1994), condiciones gubernamentales adecuadas, elevada competencia local y acceso al conocimiento (véase Hall, 2000; Castells, 1996 o Desmet y Fafchamps, 2005).

Estas ideas se resumen en el modelo centro-periferia (véase Krugman, 1991; Krugman y Venables, 1995; Fujita et al., 2001). Una nueva literatura empírica y teórica ha surgido de ese modelo, conocida como Nueva Geografía Económica (NGE, de aquí en adelante). Según Krugman (1998), la NGE explica la economía usando modelos dinámicos con un equilibrio general, que se obtiene a través de la competencia entre las fuerzas de dispersión y de concentración con economías de escala. De acuerdo con la literatura de la NGE: (i) hay incentivos a concentrar fuertemente la producción en áreas centrales, y (ii) los procesos de especialización y comercio intra-regionales y entre países refuerzan los procesos de concentración y, en consecuencia, de divergencia (véase también Baldwin y Forslid, 2003; Ottaviano y Thisse, 2004 o Behrens y Thisse, 2007).

De acuerdo a este modelo, el centro tiende a concentrar la actividad de todas las regiones de su alrededor cuando los beneficios de los vínculos hacia delante y hacia atrás son mayores que los costes de transporte de concentrar la actividad en una ubicación central.

En resumen, la escala geográfica no es relevante según la Economía Neoclásica. La falta de interacción entre los agentes representativos en este modelo genera homogeneidad en todas las escalas. Sin embargo, la NGE se centra en las fuerzas centrífugas, que crean concentración de las actividades y heterogeneidad. Desde este punto de vista, el concepto de región y la agregación no son neutrales. Diferentes clasificaciones del territorio pueden llevar a la eliminación de información relevante en las relaciones entre las zonas centrales y periféricas.

El supuesto de información homogénea en grandes áreas puede ser extremadamente arbitrario según el criterio de agregación. Las bases de datos gubernamentales han estado tradicionalmente limitadas por unidades administrativas agregadas debido a la falta de información detallada. Sin embargo, estas regiones han sido frecuentemente diseñadas por razones no económicas sino históricas o políticas. Así que los datos agregados que basan este tipo de clasificación es una combinación de diferentes unidades económicas. Esta ausencia de información desagregada a nivel local implica que los análisis regionales no tenían más opción que usar estos datos, a pesar de los problemas de agregación. Pero los análisis de la NGE requieren una especial atención a las dinámicas del nivel local más que de las áreas nacionales. El análisis de dichas dinámicas utilizando información agregada que no distingue entre áreas urbanas y rurales puede carecer de robustez.

Aun así, la disponibilidad de datos ha aumentado en los últimos años, existiendo un número creciente de bases de datos con información desagregada -o incluso con datos individuales-. Hoy en día es posible adaptar los datos al nivel de agregación o al concepto de región económica más apropiado a nuestra investigación. La elección de una escala adecuada en el análisis económico y sus consecuencias en los resultados podría volverse tan importante como la de un estimador correcto.

El objetivo principal de esta tesis es explorar el papel jugado por la escala espacial en los análisis empíricos de la economía regional. Estudia cómo un nivel geográfico no consistente con los supuestos puede afectar a las conclusiones finales y llevar a unos resultados sin sentido -o, al menos, no tan claros como deberían ser cuando la escala es elegida correctamente-. Aunque esta idea puede ser aplicada a todos los análisis espaciales, es en el campo del crecimiento económico y las diferencias económicas entre territorios donde puede darse una mayor infraestimación de la importancia de la escala. Por lo tanto, esta tesis presta especial atención a la importancia de la elección del nivel espacial en los estudios de crecimiento y convergencia, así como en análisis de productividad.

La tesis está estructurada como sigue. La Sección 1 tiene una triple contribución: (i) explora los problemas teóricos de la identificación de procesos locales utilizando estimaciones agregadas; (ii) presenta una ilustración empírica basada en datos altamente desagregado que permiten cuantificar el efecto de la ubicación específica en el crecimiento económico, y, (iii) propone una metodología multinivel para explicar las diferentes escalas en el análisis y cuantificar su importancia. Después, la Sección 2 se centra en el proceso acumulativo si se estudia a una escala espacialmente detallada. La primera parte de este capítulo estima economías de aglomeración para regiones funcionales y estudia la heterogeneidad a través de la distribución de unidades espaciales. La segunda parte de esta sección mide la influencia de esas aglomeraciones en la economía nacional en base al concepto de “granos”. El Capítulo VII termina con algunas conclusiones.

Más en detalle, la tesis empieza poniendo el foco en el fenómeno de desigualdad entre territorios usando el conocido análisis de β-convergencia. El análisis de β-convergencia es particularmente interesante para el objeto de esta tesis, pues es un campo de la literatura está directamente conectado tanto con el marco Neoclásico como con la NGE. La NGE sugiere que las desigualdades regionales en PIB per cápita surgen debido a las diferencias entre las áreas rurales y urbanas en términos de capital humano y externalidades de la actividad, mientras que las teorías neoclásicas predicen homogeneidad en los niveles de PIB per cápita entre regiones. Esta primera parte de la tesis estudia el problema de agregación espacial de los datos en la estimación de ecuaciones de β-convergencia. Se basa en estudios previos que ya han prestado atención al efecto de la agregación, como en el trabajo de Theil (1954) para el caso general con modelos de regresión lineal o, más recientemente, de Arbia y Petrarca (2011) para el caso de datos dependientes espacialmente. Además, se introduce explícitamente en el análisis la naturaleza jerárquica de los datos económicos en lo que respecta a unidades espaciales y analiza la importancia de cada nivel en el proceso mediante un enfoque econométrico de análisis multinivel.

La segunda parte de esta tesis estudia las aglomeraciones urbanas y cómo las dinámicas entre los territorios rurales y urbanos pueden afectar al resultado, y después continua con el análisis de la productividad y sus relaciones con la densidad de población. Sigue la literatura reciente, que ha prestado atención a cuantificar el impacto de las economías de aglomeración sobre la productividad -véase, por ejemplo, Rosenthal y Strange (2001), para un análisis más extenso o Ciccone y Hall (1996); Combes (2000), Combes et al. (2008), o Artis et al. (2012)-. Más recientemente, Combes y Gobillon (2015) han revisado las contribuciones más relevantes de las economías de aglomeración, que cubren tanto los intentos de estimarlos en base a datos regionales agregados como las estrategias más recientes que utilizan datos individuales. Mientras que esta última opción puede considerarse preferible cuando hay datos disponibles, en ocasiones la falta de información observable a nivel individual hace necesaria la estimación utilizando alguna media a la escala espacial dada. Si ese es el caso, utilizar datos que promedian unidades geográficas altamente desagregadas permite considerar la escala espacial adecuada para medir las economías de aglomeración, en tanto que los datos agregados espacialmente implican asumir un alto nivel de homogeneidad intra-regional.

Table of contents

I. INTRODUCTION

The choice of the spatial scale for conducting the empirical analysis should be a fundamental initial step in regional economics. Even when scholars in regional economics have paid attention for decades (Openshaw, 1984) to the role played by this scale, which must be consistent with the assumptions and the theoretical framework, the traditional lack of geographically disaggregated data has forced regional economists to use information aggregated to a large scale and use administrative large regions in their empirical analysis. Statistical and econometric techniques have experienced important improvements in the last decades for the measurement of increasingly complex socio-economic phenomena (Islam, 2003). These improvements allow us to deal with issues as spatial dependence, nonlinearities or heterogeneity and making inferences on economic relationships and impacts in a much more accurate way. The large variety on the available econometric techniques in the regional economist's toolkit makes the choice of the specific estimation strategy a relevant decision that should be conveniently justified in every empirical research. Databases also have improved significantly in the last decades, having more precise and more spatially disaggregated information for most of the variables. However, the possibilities of considering different spatial scales for the empirical analysis that should depend on the research question or the theoretical framework are still not generalized, even when this decision can influence the results more than the selection of a specific version of an estimator.

A possible reason that explains this tendency to ignore the relevance of the spatial unit of investigation and the scale of the analysis could be the influence of Neoclassical Economics. This theoretical framework basically neglected the importance of the scale in economic analysis. Their well-known models are based on decreasing returns to scale in factors (see Solow, 1956; Mankiw et al., 1992), the mobility of factors, and the spread of knowledge as explained in Barro et al. (1991), Barro and Sala-I-Martin (1992) or Sala-I-Martin (1994). These macroeconomic models are built on the aggregation of representative, independent and homogeneous agents of the economy. So, theoretical models should operate no matter the scale or the concept of region used in the analysis.

Nevertheless, this conclusion is not robust when basic growth models include a process with local interactions. For example, Lucas (1988) emphasizes human capital accumulation through schooling and learning-by-doing. In addition, Lucas (2001) develops the theory that central zones accumulate human capital, generating positive externalities. It remarks the idea of a positive incentive in the urban areas towards accumulation of knowledge and spillovers. A process with these characteristics would be impossible to distinguish within an aggregated scale. Consequently, the scale does have an influence on the results and the conclusions.

This type of processes in the local scale are highlighted with later models that introduce accumulative mechanics in the regions. Romer (1990 and 1994) explains that the firms in urban areas generate endogenous technical progress which attracts human capital and creates positive externalities. Myrdal (1957) and Hirschman (1958) also explain that territories follow an accumulative process. According to their research, successful locations activate internal and external scale economies. As a result, they attract factors from under-developed regions and increase the process of scale economies, which is formalized in Kaldor (1957), Dixon and Thirlwall (1975). Their research explains that, due to the Verdoon law, á region can develop accumulative processes. The production in this model boosts productivity. In a competitive framework, this productivity growth reduces prices and, therefore, increases the demand. The result is a new production growth. As they explained, the interaction at the local scale generates these scale economies. The geographical scale is therefore a key element in order to analyze this type of mechanisms.

There is an important discussion about the nature of the positive externalities created within the urban areas. The most accepted theoretical literature could be the proposal of Marshal (1890) about the agglomeration economies. This phenomenon can be divided into two: the location economies and the urbanization economies. The location economies describe the externalities caused by the interaction of activities of the same sector. This interaction attracts specialized workers, suppliers and access to knowledge. The urbanization economies explain that the concentration of the activity triggers relationships between individuals –known as social capital– and infrastructures –e.g. railways, innovation centers or hospitals–. Parr (2002) summarizes this idea with a modern and clear delimitation of the concept of agglomeration economies.

The literature of the agglomeration economies has evolved from the 1950's with more contributions that explain the processes developed in the urban areas. Isard (1956) modified the model of Christaller (1933) and showed how a hierarchy of urban centers (see Zipf, 1949) grows to provide central goods. Jacobs (1969) explains that the cities generate innovation due to the interaction of people from different sectors, and Porter (1990) explains that the firms can improve their competitive advantage belonging to a network of business and institutions. With this structure, the network obtains specialized suppliers and workers (see Duranton and Puga, 2000; Glaeser, 1998 or Glaeser, 1994), suitable government conditions, a high local competition and access to knowledge (see Hall, 2000; Castells, 1996 or Desmet and Fafchamps, 2005).

These ideas are summarized in the core-periphery model (see, Krugman, 1991; Krugman and Venables, 1995; Fujita et al., 2001). A new empirical and theoretical literature emerged from this model, known as New Economic Geography (NEG hereinafter). According to Krugman (1998), NEG explains the economy by using dynamic models with a general equilibrium. The equilibrium is obtained through a competition between forces of dispersion and concentration with scale economies. According to NEG literature: (i) there are incentives to largely concentrate the production in the central areas, and (ii); the intra-regional and inter-country processes of specialization and trade reinforce the processes of concentration and, in consequence, of divergence (see also Baldwin and Forslid, 2003; Ottaviano and Thisse, 2004 or Behrens and Thisse, 2007).

This model defines a large metropolis in which scale and agglomeration economies are strong, in opposition with the small size places located far away from this core. The core-periphery model explains an economic system in the special case of two regions. Some assumptions need to be made in order to simplify this problem. In this economy we have two sectors: on the one hand there is a competitive agricultural sector with an exogenous part of the population; on the other hand there is a monopolistically competitive manufacturing sector with a labor force that moves to the region with the highest wage. This model explains that the core benefits from forward links –lower price due to concentration of the industry– and backward link –higher wages due to higher income–.

According to this model, the center tends to concentrate the activity of all the surrounding regions when the benefits from the forward and backward link are bigger than the transport costs of concentrating the activity in a central location.

To sum up, geographical scale is not relevant according to the Neoclassical Economics. The lack of interaction between the representative agents in this models generates homogeneity in all the scales. However, the NEG focuses on the centrifugal forces, which create concentration of the activities and heterogeneity. From this point of view, the concept of region and the aggregation are not neutral. Different classifications of the territory could lead to the elimination of valuable information on the relationship between central and peripheral locations.

The assumption of homogeneous data for wide areas can be extremely arbitrary depending on the aggregation criteria. Databases from governments have been traditionally limited to aggregated administrative units due to the lack of detailed information. However, these regions have been usually designed by not economic but historical or political reasons. So the aggregated data that base on this type of classification is mix different economic units. Depending on the research question, it can undermine the economic analysis. This absence of disaggregated data in a local scale implies that the regional analysis had no option but to use this information, despite the problems of aggregation. However, NEG analysis requires an especial attention to the dynamics of the local level rather than national areas. The analysis of these dynamics using aggregated information that do not distinguish between urban and rural areas may lack of robustness.

Nevertheless, the availability of data has grown in the last years. There is an increasingly amount of databases with disaggregated information –or even individual data–. Nowadays, it is possible to adapt the data to the level of aggregation or to the concept of economic region more appropriate to our research analysis. The choice of a suitable scale in the economic analysis and its consequences in the results could become as important as the choice of the correct estimator.

The central aim of this thesis is to explore the role played by the spatial scale in the empirics of regional economics. It studies how a geographical scale not consistent with the assumptions could affect the final conclusions and lead to meaningless results –or, at least, not as clear as they could be when the scale is properly chosen-. Although this idea could be applied to all spatial analysis is in the field of economic growth and territorial economic differences where most relevant underestimations of the relevance of the scale could be happening. So, this thesis particularly pays attention to the relevance of the election of the spatial scale in growth and convergence studies as well as in productivity analysis.

The thesis is structured as follows. Section 1 has a threefold contribution: (i) explores the theoretical problems of identification of local processes using aggregated estimations; (ii) presents an empirical illustration basing on highly disaggregated data that allows for quantifying the effect of specific location on economic growth, and, (iii) proposes a multilevel methodology in order to account for different scales in the analysis and quantify their importance. Next, Section 2 focuses on the accumulative processes if it is studied at a spatially detailed scale. The first part of this chapter estimates agglomeration economies for functional regions and study the heterogeneity across the distribution of spatial units. The second part of this section measures the influence of these agglomerations on the national economy basing on the concept of "grains". Chapter VII ends the dissertation with some conclusions and discussion.

More in detail, the thesis starts focusing on the phenomenon of inequalities between territories using the well-known β-convergence analysis. β-convergence analysis is particularly interesting for the aim of the thesis, since it is a field of the literature directly connected with both the Neoclassical framework and the NEG. NEG suggests that regional inequalities in Gross Domestic Product per capita (GDPpc) emerges due to the differences between rural and urban areas in terms of human capital and externalities of the activity, while neoclassical theories predict homogeneity of the levels of GDPpc across regions. This first part of this thesis studies the problem of spatial aggregation of data when estimating β-convergence equations. It bases on previous studies that have already called the attention to the effect of the aggregation, like in the work by Theil (1954) for the general case on linear regression models or, more recently, by Arbia and Petrarca (2011) for the case of spatially dependent data. Additionally, it explicitly introduces in the analysis the hierarchical nature of economic data when referring to spatial units and analyzes the importance of each level in the process by using econometric approach of multilevel analysis.

The second part of this dissertation studies urban agglomerations through the dynamics between productivity and its relationship with population density. It follows the recent literature, which has paid attention to quantify the impact of agglomeration economies on productivity –see, for example, Rosenthal and Strange (2001), for an extensive review or Ciccone and Hall (1996); Combes (2000), Combes et al. (2008), or Artis et al. (2012)–. More recently, Combes and Gobillon (2015) have reviewed the most relevant contributions to the empirics of agglomeration economies, which covers both the attempts to estimate them basing on aggregated regional data to the more recent strategies that use individual data. While this last option is arguably preferable when data are available, sometimes lack of observable information at individual level makes necessary the estimation basing on some average at a given spatial scale. If this is the case, using data that average highly disaggregated geographical units allows for considering an appropriate spatial scale to measure agglomeration economies, since spatially aggregated data imply assuming a high level of intra-regional homogeneity.

Additionally, a methodology based on the granular hypothesis of Gabaix (2011) is proposed in order to measure the importance of the concentration in the national outcome. This hypothesis means that the idiosyncratic behavior of the large units should be capable of explaining a significant part of the aggregate shocks. Under the usual scenario of concentration, we should presumably find some kind of granular hypothesis.

We analyze whether this behavior is present in the case of the urban concentrations of the US, as can be expected taking into account the degree of spatial concentration in the economy. We use data from the Bureau of Economic Analysis on personal income, which is disaggregated at a local level, US counties, from 1969 to 2011. The granular residual of the largest cities is calculated and used to explain the US aggregate economic evolution. The overall results provide support for the granular hypothesis: the idiosyncratic shocks to the top counties can explain a significant fraction of the volatility of US aggregate data.

Section 1

II. SPATIAL SCALE AND AGGREGATION EFFECTS

II.1. Territorial disparities, economic growth and convergence analysis

It is very relevant to observe that theories such as the regional economics approach or the NEG models have a more local perspective than their neo-classical counterparts, which mainly focus on national or large regions analysis and less focussed in the spatial aggregation. Under their perspective, cities and metropolis (local areas) are located in the centre of the analysis. They draw the attention to cities as the missing link between the macroeconomic theories of growth and the spatial empirical analysis[1].

The role of spatial concentration and convergence has been widely documented in regional economics. Some explanations can be found in the literature, such as the endogenous growth framework. It underlines the effect of agglomeration effects, essentially positive externalities due to location. When activities are together, they tend to increase competitiveness, spillovers and specialized factors. The New Economic Geography (see Fujita et al., 2001) explains that there is a tension between the core and the periphery that depends on increasing returns, transport costs, and centripetal and centrifugal forces.

The most important methodologies to test the hypothesis of the different theories are the analysis of the economic growth and productivity. This section focuses on the evolution of the economic growth and disparities of the territories. Convergence analysis is the most suitable and extended methodology in this field to test the theoretical implications in both, Neoclassical Economics and NEG.

The convergence hypothesis stablishes that all the economies will tend to the same GDP per capita in the long run. As explained above, NEG and Neoclassical Economics obtain different conclusions due to the assumptions of their models. There

[1] See, for example, the empirical analysis of Ciccone and Hall (1996) who found a positive relation between density and productivity.

are many ways of studying convergence among territories. Nevertheless, sigma (σ), stochastic and, especially, beta (β) convergence are the most commonly applied instruments.

σ-convergence is perhaps the simplest approach. It consists in quantifying the dispersion or variability of income per capita or a similar variable in logarithms along different moments in time: if the standard deviation of the variable of interest decreases along time, this is considered as an indication of convergence. This kind of analysis is usually conducted as an exploratory or preliminary analysis in the study of convergence.

β-convergence analysis measures whether poor territories grow faster than rich territories. According to this measure, evidence of β-convergence in the sample would indicate that the gap between rich and poor territories is diminishing over time. This estimation is made through econometric estimation of a linear regression between the GDP pc growth and the initial GDP pc in the period.

Stochastic convergence is based on a time series test for unit roots, which make possible to test for persistent differences in the series of income or total production. This methodology can be directly linked with the empirical approach of time-series analysis.

However, the σ and, specifically, β-convergence analysis are the most commonly applied approaches in empirical studies on regional convergence. The advantage of β-convergence is that based in the neoclassical models of economic growth and allows to obtain a direct contrast of their assumptions (see Sala-I-Martin, 1996 or De la Fuente, 2002). In addition, it allows to include information of other relevant factors as well as it has been improved with several econometric techniques. This is the reason why in the subsequent chapters of this section we will limit our discussion to traditional β-convergence analysis.

β-convergence was introduced by Baumol (1986). He used a simple Ordinary Least Squares regression of the income per capita or similar variable growth rate in a territory on the initial level of that variable. So, this model is based in the following equation:

$$\Delta \ln y_i = \alpha + \beta \ln y_{0i} + u_i \qquad (2.1)$$

Where Δlny_i is the growth rate of income per capita during a period of time of the i spatial unit and y_{0i} is the income per capita in the initial moment of the period. When no other regressor is considered, we talk about an analysis of unconditional β-convergence, whereas if other explanatory variable is included we conduct a conditional β-convergence analysis. With this estimation framework we can see if poorer areas tend to grow faster or not than the rich ones as Solow (1956) and posterior models based on it predicts. If the parameter β is estimated with negative sign, this indicates that lower levels of income per capita produce higher growth rates, leading to a process of convergence in the long run. A positive estimate of β would reveal a process of divergence[2].

In his empirical study Baumol (1986) estimated an equation like (2.1) for a dataset of industrial countries and using the output per worker as indicator of growth. He regressed productivity growth from 1870 to 1979 on labour productivity in 1870. He obtained a β parameter of 0.75 that indicates a very weak and slow process of convergence between different countries.

This result of Baumol seminar paper was reproduced for regions in later research. There are several examples of international researches with cross-section information which estimate the rate of convergence between regions of the same country. In all of them is usual to find an estimation of unconditional beta-convergence, because it can be assumed that regions within the same country are defined by similar relevant factors. As a consequence, the diminishing returns will lead all regions to the same steady state. Sala-I-Martin (1994) found a 3% of convergence for the U.K. (1950-1990), 1.7% for the U.S.A. (1880-1990), 1.6% for France (1950-1990), 1.4% for Germany (1950-1990) and 1% for Italy (1950-1990). In addition to this, other authors also find similar results for other countries: Coulombe and Lee (1993) found a 2.4% for Canada (1961-1991), Shioji (1992) estimated a 1.9% for Japan (1955-1990), De la Fuente (2002) also found a 2.95% rate of unconditional convergence for Spain. In other words, all they conclude that generally poor regions grow faster than the rich ones, so, finding a positive rate of convergence is almost

[2] There are other ways to measure convergence than the ones depicted in this section. For example, the γ-convergence focuses on relative rankings of the income per capita of the territories. The stochastic convergence conducts a time series test for unit roots, which make possible to check if there are persistent differences in the series of income or total production. However, the σ and, specifically, the β-convergence analysis are the most commonly applied approaches in empirical studies on regional convergence. This is the reason why in the subsequent sections of this chapter we will limit our discussion to traditional β-convergence analysis.

regularity. Moreover, there is some consensus on the idea that a 2% is almost the 'magic number' in rates of convergence between regions.

However, recent researches indicate that other results can also be found. For example, it seems that the US regions are not converging for a recent period. Tsionas (2000) pointed the distribution of regional income in logarithms between 1977 and 1996 only had little variations, and the beta-convergence estimation indicates a pattern of divergence. On the other hand, Koo et al. (1998) concluded that developing countries could obtain a bigger rate of convergence than the developed ones.

A conditional convergence analysis is also possible if we include information of the relevant factors, which can explain the steady-state of an economy (see - Barro et al., 1991). According to a Solow model, these factors could be the percentage of savings, the population growth or the technologic growth. When we do not have data of these relevant factors, we can overcome this lack of information with panel data: we can use a constant term for each region in a model with fixed effects to calculate the influence of these relevant factors. A result of conditional convergence has a different conclusion for regional policy. In this case, there is convergence when we take into account the relevant factors of the different economies, so the regional policy should change them in poor regions. In this kind of analysis, a bigger rate of convergence is observed, with a 12.73% obtained by De la Fuente for Spain or a 6.14% for Korea indicated by Koo. This type of results could show that regions within the same country don't necessary have the same steady state.

All this empirical convergence models are based in the neoclassical framework and, consequently, do not put too much attention to the spatial unit. But, as have been seen, other models of economic growth show how important are the local factors. Under these perspectives agglomeration economies or other centripetal forces have an effect at the local level, so the aggregation into packages of regional data, which is normal in convergence analysis, can hide all this intraregional information. The possibility of this error cast doubts on the empirical evidence found in convergence studies. This is the main motivation in this chapter to investigate the effect of aggregating data

This problem has been partially explored in Miller and Genc (2005) and Resende (2011). They argue the importance of the level of disaggregation over the convergence analysis. This is the main motivation for their analysis, which quantifies the convergence between several possible spatial divisions for the US and Brazil. For the US case there was not found evidence of a significant effect of the scale effect on the results, while the opposite happened in the case of Brazil. Both studies are limited to specific cases for a particular time period. It would be interesting, in consequence, to extend the analysis to a more general framework.

II.2. The beta convergence model

Equation (2.1) could be considered as the first β-convergence estimation. However, this equation does not include any other relevant factor which could be relevant for economic growth.When no other regressor is considered, we talk about an analysis of unconditional β-convergence, whereas if other explanatory variables (X_i) are included we conduct a conditional β-convergence analysis – see equation (2.2).

$$\Delta \ln y_i = c + g \ln y_{i0} + \theta \ln X_i + \Delta \ln \varepsilon_i$$
$$\beta = \left(1 - e^{-\lambda t}\right) \tag{2.2}$$

In order to choose the control variables of equation (2.2) we depart from a typical Cobb-Douglas production function of per capita income in the period t and region – municipality- i:

$$y_i = A_i K_i^{\alpha} L_i^{\beta} \varepsilon_i \tag{2.3}$$

taking differences on the log-linear form of (2.3)

$$\Delta \ln y_i = \Delta \ln A_i + \alpha \Delta \ln k_i + (\beta + \alpha - 1) \Delta \ln L_i + \Delta \ln \varepsilon_i \tag{2.4}$$

where, as usual, y is the income per capita and A is the level of technology, while L and K stand for the stocks of labor and capital in the economy. In this equation, income growth is a function comprising four components: technology growth, changes in capital per capita, population growth and changes due to exogenous shocks. The definition of these terms would provide us a hint about the proper variables for the estimation.

Equation (2.2) has been widely applied in the literature. However, the assumption of a unique speed of convergence seems to be very strong when we address not fully integrated economies. Several authors have already noted that each region could have its own steady state (see Canova and Marcet, 1995; Islam, 1995; Islam. 2003).

Despite the enormous literature in order to estimate the coefficient of convergence without problems of bias, there is almost no information about the appropriate scale that should be applied to measure this process. That is why this research proposes to understand the process of convergence as a group of forces operating at different levels. In order to test this type of movements this analysis includes a process of convergence for each group of regions as in equation (2.5).

$$\Delta \ln y_{ij} = c - \beta \ln(y_0)_{ij} + \theta \ln X_{ij} - \beta_j \ln(y_0)_{ij} + c_j, \qquad \forall i \in [1, N], \forall j \in [1, M] \tag{2.5}$$

One of the most suitable methodologies in order to test this type of movements is called the multilevel methodology. The main goal of this focus is the measurement of the importance of the different levels of the hierarchy. One example of how this methodology works can be found in Ballas and Tranmer (2012). They use the UK census to identify the relative importance of area, household and individual characteristics in variations in happiness. In this research we adapt the hierarchy concept to space, assuming the possibility of having different processes of convergence across spatial levels. The general equation of conditional convergence (2.2) can be extended in a new particular multilevel conditional convergence equation (2.5) that allows each of the M countries to have its own process of convergence, with a general process for N regions.

It can be observed that the variability in the growth of a region is divided into two components, allowing interactions between and within countries. Finally, following Rey and Montouri (1999), equation (2.5) can be augmented with spatial effects. Equation (2.6) introduces spatial interactions through the diffusion effect of the idiosyncratic behaviours. These spatial interactions with the neighbours should decrease as the distance increases to avoid explosive effects.

$$\Delta \ln y_{ij} = c - \beta \ln(y_0)_{ij} + \theta \ln X_{ij} - \beta_j \ln(y_0)_{ij} + c_j + \varepsilon_{ij} + \sum \xi_{ik}, \quad \forall ik \in [1, N], \forall j \in [1, M] \tag{2.6}$$

II.3. The role of spatial scale and aggregation in convergence analysis

An accurate evaluation of the aggregation effect in the convergence analysis requires different steps. The first step must be a systematical evaluation of the difficulties to extract conclusions for the sub-regional reality from convergence analysis with aggregate data. More specifically, we aim at quantifying the effect of neglecting small-scale processes derived from estimating β-convergence equations based on spatially aggregated data. Our research bases on previous studies that have already called the attention to the effect of the aggregation, like in the work of Theil (1954) for the general case on linear regression models or, more recently, by Arbia and Petrarca (2011) for the case of spatially dependent data studying the so-called Modifiable Areal Unit Problem (MAUP) issue.

According to the seminal work by Openshaw (1984), the object of the analysis in empirical studies conducted at some spatial scale should be described before anyone tries to measure its characteristics. However, the situation with real data is hard to achieve: the region exists only after the data are collected. As a consequence, the definition of the object is arbitrary and it can be changed. Behrens and Thisse (2007) also explain that grouping location is a problem really similar to the poorly representative called "representative consumer". Openshaw studied the relationship between the percentage of Republican and elderly voters with different territory aggregations. Surprisingly, the range of the possible correlation coefficient goes from -0.99 to 0.99.

The MAUP can be divided in two parts. Firstly, we have the scale effect, which is the most important part of the problem. This aggregation bias appears if we aggregate our data into larger units, for example cities to regions. Secondly, there is a zoning effect. In this case, we have a problem with the shape of the units because a different form can also change our results.

We have seen that some authors found convergence in their regional databases. The main problem is that they find convergence with information which has been aggregated. As a consequence, the new aggregated variable does not have to maintain the characteristics of the original variable, especially when we do not aggregate following some economic but purely administrative division. If a MAUP issue were present, this would imply that our results of convergence would not be the same if our data were group into a different group of regions or directly observable at a smaller scale.

Theoretical evaluation is made in this dissertation with the introduction of the aggregation structure in the convergence equation. The non-linear aggregation leads us to a different outcome Arbia and Petrarca (2011). So, our analysis tries to understand this possible problem in terms of bias and efficiency. Measurement of these deviations are made through a Montecarlo simulation with a random generation of the initial situation of GDPpc in different scenarios.

However, these deviations are not the only problem of aggregation in convergence equations. Estimation of local processes with aggregate data can be unreliable. A measurement of location effects in the case of Mexico using the convergence equation is shown as an example. In this analysis, the size of the states of Mexico would make a measurement based on distances extremely ambiguous with aggregate data.

Suitability of econometric estimations which are not affected by the scale is discussed with an estimation for the well-known EU. This measurement tries to relax the assumption of a generated process is only created at the local level. In addition, it can help us to understand the importance of each level in the process. Multilevel methodology assumes that the sub-regions are connected realities with different levels of aggregation. In this scenario, the economic model recognises that there may be forces operating at both, the region and the sub-regions. In other words, the region may also have an important role in the generation process of growth and convergence.

II.4. The effect of the aggregation on the OLS estimation of β–convergence

In this chapter we analyze the theoretical consequences of scale misspecification. The advantage of this approach is that the results are easy to generalize. Meanwhile, the empirical analysis of this phenomenon is always conditioned by the sample database. This characteristic makes it more suitable to start with.

While panel data estimators are the type of estimation strategy most commonly followed by far in the context of analyzing country data, in the context of regional analysis is not uncommon to base the estimation of β-convergence equations on cross-sectional data due to information availability (see, for example, Azzoni, 2001, for Brazil; Rodríguez-Pose and Sánchez-Reaza, 2005, for Mexico; Cuadrado-Roura, 2001, for Europe; or Raiser, 1998, for China). This chapter studies the properties of a traditional ordinary least squares (OLS) estimator of β-convergence equations based on a cross-section of data.

Let us assume an economy that is divided into different spatial units that are created according to several criteria for geographical aggregation. More specifically, suppose that the economy is divided into $\mathrm{i} = 1, \dots, \mathrm{n}$ basic spatial units –municipalities or cities– that are aggregated into $\mathrm{j} = 1, \dots, \mathrm{m}$ ($\mathrm{m} < n$) groups –regions–. In line with the ideas of New Economic Geography and endogenous growth theories, we assume that the process of income generation takes place at the basic spatial scale of n units. This chapter studies the effects on the conclusions of convergence analysis depending on the scale at which the outcome data are observable: directly observable at the original scale (n local places) or at the aggregated scale (m regions). If the conclusions about the coefficient depend on the level of aggregation, this will be a signal that a potential MAUP is somehow 'contaminating' our analysis.

Our starting point will be the formulation developed in Arbia and Petrarca (2011) for the case of cross-sectional data in a linear regression model that are generated at a given spatial level, but then observed at a more aggregate scale. The following equation describes the model to be estimated at a disaggregated scale with n spatial units:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2.7}$$

where $\mathbf{y}$ is the $(\mathrm{n} \times 1)$ vector with the dependent variable, $\mathbf{X}$ is a $(\mathrm{n} \times \mathrm{K})$ matrix with the K regressors considered in the equation, $\boldsymbol{\beta}$ is the $(\mathrm{K} \times 1)$ vector with the parameters to be estimated and $\boldsymbol{\varepsilon}$ is the typical $(\mathrm{n} \times 1)$ disturbance, which is assumed to distribute normally around zero with a constant variance σ^2. If the data of the n units are aggregated at a higher geographical scale with m locations, the new dataset is defined by:

$$\mathbf{y}^* = \mathbf{G}\mathbf{y} \tag{2.8}$$

$$\mathbf{X}^* = \mathbf{G}\mathbf{X} \tag{2.9}$$

$$\boldsymbol{\varepsilon}^* = \mathbf{G}\boldsymbol{\varepsilon} \tag{2.10}$$

Being $\mathbf{G}$ the aggregation matrix with dimensions (m × n), including elements like:

$$\mathbf{G} = \begin{bmatrix} g_{11} & \cdots & g_{1r_1} & & \cdots & & 0 & \cdots & 0 \\ 0 & \vdots & 0 & g_{21} & \cdots & g_{2r_2} & 0 & \vdots & 0 \\ & \cdots & & & \cdots & & & \cdots & \\ 0 & \cdots & 0 & & \cdots & & g_{m1} & \cdots & g_{mr_m} \end{bmatrix} \tag{2.11}$$

where each row indicates that the original data is aggregated -grouped- into m different locations, being the number of original spatial units differently aggregated in each case $(r_1, r_2, \ldots, r_m)$.

In this context, the aggregated equation is defined as:

$$\mathbf{y}^* = \mathbf{X}^*\boldsymbol{\beta}^* + \boldsymbol{\varepsilon}^* \tag{2.12}$$

where:

$$\mathrm{E}(\boldsymbol{\varepsilon}^*) = \mathrm{E}(\mathbf{G}\boldsymbol{\varepsilon}) = \mathbf{0} \tag{2.13}$$

$$\mathrm{Var}(\boldsymbol{\varepsilon}^*) = \mathrm{E}(\boldsymbol{\varepsilon}^*\boldsymbol{\varepsilon}^{*\prime}) = \mathrm{E}(\mathbf{G}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{G}') = \mathbf{G}\mathbf{G}'\sigma^2 \tag{2.14}$$

In their paper, Arbia and Petrarca (2011) deal with the specific case of perfect aggregation where the elements of this aggregation matrix $\mathbf{G}$ are unitary values:

$$\mathbf{G} = \begin{bmatrix} 1 & \cdots & 1 & & \cdots & & 0 & \cdots & 0 \\ 0 & \vdots & 0 & 1 & \vdots & 1 & 0 & \vdots & 0 \\ & \cdots & & & \cdots & & & \cdots & \\ 0 & \cdots & 0 & & \cdots & & 1 & \cdots & 1 \end{bmatrix} \tag{2.15}$$

Being the number of ones in every row always equal to $r = m/n$. They show how the OLS estimator of $\boldsymbol{\beta}^*$ $(\widehat{\boldsymbol{\beta}}^*)$ of equation (2.12) is an unbiased estimator of $\boldsymbol{\beta}$ in the original equation (2.7), being the variance of the OLS estimator in the aggregated equation (2.12) bigger than the original variance of the OLS estimator in (2.7):

$$\begin{aligned} \mathrm{E}(\widehat{\boldsymbol{\beta}}^*) = \mathrm{E}\left([\mathbf{X}'^*\mathbf{X}^*]^{-1}\mathbf{X}'^*\mathbf{y}^*\right) = \mathrm{E}\left([\mathbf{X}'\mathbf{G}'\mathbf{G}\mathbf{X}]^{-1}\mathbf{X}'\mathbf{G}'\mathbf{G}\mathbf{y}\right) \\ = \mathrm{E}\left([\mathbf{X}'\mathbf{X}]^{-1}\mathbf{X}'\mathbf{y}\right) = \mathrm{E}(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\beta} \end{aligned} \tag{2.16}$$

$$\mathrm{Var}(\widehat{\boldsymbol{\beta}}^*) = \mathbf{G}\mathbf{G}'\sigma^2[\mathbf{X}'\mathbf{G}'\mathbf{G}\mathbf{X}]^{-1} > Var(\widehat{\boldsymbol{\beta}}) \tag{2.17}$$

In other words, the scale effect does not represent a problem of bias, although it generates an efficiency problem.

The β-convergence equations, however, are not characterized by this same response to the scale effect, due to some particularities in the aggregation scheme of the dependent and the independent variables and the logarithmic form of the equation. In order to justify this claim, let us state the typical absolute β-convergence equations estimated for a cross-section of n spatial units as:[3]

$$\Delta \ln y_i = \alpha + \beta \ln(y_{i0}) + \varepsilon_{it} + \varepsilon_{it}; \text{ or}$$
$$\ln(y_{it}) = \alpha + (1+\beta)\ln(y_{i0}) + \varepsilon_i \qquad (2.18)$$

Where the growth in an economic indicator y as GDP or income, value added, etc. per capita between periods 0 and t in location i regressed on the logs of the initial variable per capita (y_{i0}) on the same location. One problem with aggregated data for estimating equations like (2.18) is that the non-linearities in the dependent and explanatory variables are not compatible with the equivalences between the aggregated and disaggregated equation. More specifically, the aggregate version of the absolute β-convergence equations equation will be:

$$\Delta \ln y_j^* = \alpha^* + \beta^* \ln\left(y_{j0}^*\right) + \varepsilon_{jt}^* \text{ or}$$
$$\ln\left(y_{jt}^*\right) = \alpha^* + (1+\beta^*)\ln\left(y_{j0}^*\right) + \varepsilon_{jt}^* \qquad (2.19)$$

Being:

$$\mathbf{y_0^*} = \mathbf{G}\mathbf{y_0} \qquad (2.20)$$

$$\boldsymbol{\varepsilon}^* = \mathbf{G}\boldsymbol{\varepsilon} \qquad (2.21)$$

Matrix $\mathbf{G}$ represents the aggregation scheme for the initial values per capita, with a typical element g_{ij} indicating the population share of the basic spatial unit i on the aggregated location j measured in the initial period. In contrast to the type of equations aggregated as in (2.8), the dependent variable of the equation estimated with aggregate data is given by:

$$\ln(\mathbf{y_t^*}) = \ln(\mathbf{H}\mathbf{y_t}) \neq \mathbf{G}\mathbf{y_t} \qquad (2.22)$$

[3] A similar exercise could be done for conditional β-convergence equations just by adding more regressors to this basic equation. We have opted for working with this simple case for the sake of simplicity but the main conclusions in terms of the effects of aggregation on its estimation, however, would hold.

where $\mathbf{H}$ is the aggregation matrix where a typical element h_{ij} indicates the population share of the spatial unit i on region j measured in the final period. In general, this matrix is not necessarily equal to $\mathbf{G}$, given that the elements of $\mathbf{H}$ are the population shares in the final period and the populations in each period can be different.

Note that equation (2.16) states that the expected value of the OLS estimator with aggregated data is given by $\mathrm{E}\big([\mathbf{X}'\mathbf{G}'\mathbf{G}\mathbf{X}]^{-1}\mathbf{X}'\mathbf{G}'\mathbf{G}\mathbf{y}\big)$ and it is equal to β, while a different aggregation scheme would modify the form of the estimator being its expected value $\mathrm{E}\big([\mathbf{X}'\mathbf{G}'\mathbf{G}\mathbf{X}]^{-1}\mathbf{X}'\mathbf{G}'\mathbf{H}\mathbf{y}\big)$. When the elements of matrix $\mathbf{H}$ are larger than the elements of $\mathbf{G}$, the estimator will present a positive bias, while a negative bias will be the consequence of the elements of $\mathbf{H}$ being smaller than those in $\mathbf{G}$. The comparison between these two matrices can be made in terms of the Euclidean norms of their row-vectors, comparing $\sqrt{\mathbf{h}'_j\mathbf{h}_j}$ with $\sqrt{\mathbf{g}'_j\mathbf{g}_j}$. These norms would account for the concentration of population shares on each region j –they can be interpreted as a Herfindahl index for the distribution of population in region j-. If population in the final period is more unequally distributed than in the initial period and, in general, $\sqrt{\mathbf{h}'_j\mathbf{h}_j} \geq \sqrt{\mathbf{g}'_j\mathbf{g}_j}$ this would lead to a positive bias in the estimation of β. The opposite situation will happen when the population in the final period is more evenly distributed within regions than in the initial period.

Even if the aggregation criterion reflected in $\mathbf{H}$ was the same as the aggregation scheme present in matrix $\mathbf{G}$, an additional problem derived for the non-linear nature of the β-convergence equation will be present, affecting the properties of the OLS estimation from aggregated data. Assuming a case where $\mathbf{G} = \mathbf{H}$, note that $\ln(\mathbf{y}_t^*) = \ln(\mathbf{G}\mathbf{y}_t) \neq \mathbf{G}\mathbf{y}_t$. This problem is the same with the matrix of explanatory variables $\mathbf{X}^*$(which in the case of absolute β-convergence equations corresponds to the log of the initial levels $\mathbf{y}_0^*$) given that $\ln(\mathbf{y}_0^*) = \ln(\mathbf{G}\mathbf{y}_0) \neq \mathbf{G}\ln(\mathbf{y}_0)$.[4] Specifically, we could argue that $\ln(\mathbf{y}_t^*) \leq \mathbf{H}\ln(\mathbf{y}_t)$ and $\ln(\mathbf{y}_0^*) \leq \mathbf{G}\ln(\mathbf{y}_0)$ basing on Jensen's inequality. These inequalities imply that equations (2.16) and (2.17) do not hold, affecting the expected value and the variance of the OLS estimator of an aggregate equation as (2.12). The dependent variable $\mathbf{y}_t^*$ in the case of β-convergence equations with

[4] For the sake of clarity in the exposition, in the remaining of this section we refer to the matrix of potential regressors $\boldsymbol{X}$ included as explanatory variables in the specification of a general β-convergence equation. Absolute β-convergence equation only considers initial values $\boldsymbol{y}_0$ in matrix $\boldsymbol{X}$.

aggregated data is $\ln(\mathbf{Hy_t})$, being the matrix of regressors $\mathbf{X}^*$ given by $\ln(\mathbf{GX})$. The expected value and the variance of the OLS estimator for this aggregated equation are respectively:

$$\begin{aligned}\mathrm{E}\big(\hat{\boldsymbol{\beta}}^*\big) &= \mathrm{E}([\mathbf{X}'^*\mathbf{X}^*]^{-1}\mathbf{X}'^*\mathbf{y}_\mathbf{t}^*) = \mathrm{E}\big([\ln(\mathbf{GX})'\ln(\mathbf{GX})]^{-1}\ln(\mathbf{GX})'\ln(\mathbf{Hy_t})\big)\\ &\neq \mathrm{E}([\mathbf{X}'\mathbf{G}'\mathbf{GX}]^{-1}\mathbf{X}'\mathbf{G}'\mathbf{Gy_t}) \neq \boldsymbol{\beta}\end{aligned} \tag{2.23}$$

$$\begin{aligned}\mathrm{Var}\big(\hat{\boldsymbol{\beta}}^*\big) &= \mathrm{Var}\big([\mathbf{X}'^*\mathbf{X}^*]^{-1}\mathbf{X}'^*\mathbf{y}_\mathbf{t}^*\big) == \mathrm{Var}\Big([\ln(\mathbf{GX})'\ln(\mathbf{GX})]^{-1}\ln(\mathbf{GX})'\ln(\mathbf{Hy_t})\Big)\\ &\geq \sigma^2[\mathbf{X}'\mathbf{G}'\mathbf{GX}]^{-1} \geq \mathrm{Var}(\hat{\boldsymbol{\beta}})\end{aligned} \tag{2.24}$$

The result in (2.24) is equivalent to (2.17), indicating the augmenting effect of the aggregation on the variance of the estimator. However, equation (2.23) shows how a problem of bias emerges now as well, in contrast to the result in (2.16). The positive or negative sign of the bias. It depends on the aggregation schemes represented on matrices $\mathbf{G}$ and $\mathbf{H}$ -because of the per capita nature of the dependent and explanatory variables- and it is not straightforward, since their elements are influenced by the population dynamics of the spatial units aggregated into larger regions. The issue of the logarithmic transformation adds more complexity to the study of the bias.

Equation (2.23) shows how an OLS estimation of β-convergence based on aggregated data can be affected by a problem of bias. This problem is caused by the differences in the aggregation matrices $\mathbf{G}$ and $\mathbf{H}$, which respectively affect the values of the explanatory and dependent variables, and for the non-linear nature of the β-convergence equations. We will show this basing on the basic formulation:

$$\ln(y_{it}) = \alpha + (1+\beta)\ln(y_{i0}) + u_i \tag{2.25}$$

Considering vector y_0, which contains the initial values included as regressor in the β-convergence equation, Jensens's inequality states that and $\ln(Gy_0) \leq G\ln(y_0)$. Note that it is possible to re-write this inequality as:

$$\ln(Gy_0) = \hat{c}_0 G\ln(y_0) \tag{2.26}$$

where $\hat{\mathbf{c}}_\mathbf{0}$ is a diagonal $(m \times m)$ matrix with a typical element $\hat{c}_{0j}$ defined as:

$$\hat{c}_{0j} = \frac{\ln(\mathbf{g'}_\mathbf{j}\mathbf{y_{jo}})}{\mathbf{g'}_\mathbf{j}\ln(\mathbf{y_{jo}})} \leq 1 \tag{2.27}$$

In (2.27), $\mathbf{g'}_{\mathrm{j}}$ refers to the (row) vector of matrix $\mathbf{G}$ that aggregates the initial values of $\mathbf{y_o}$ that belong to the aggregated region j ($\mathbf{y_{jo}}$). Similarly, concerning the aggregation of the dependent variable, we can write:

$$\ln(\mathbf{Hy_t}) = \hat{\mathbf{c}}_{\mathbf{t}}\mathbf{H}\ln(\mathbf{y_t}) \tag{2.28}$$

where the elements of the diagonal matrix $\hat{\mathbf{c}}_{\mathbf{t}}$ are given by the expression:

$$\hat{\mathrm{c}}_{\mathrm{tj}} = \frac{\ln(\mathbf{h'}_{\mathbf{j}}\mathbf{y_{jt}})}{\mathbf{h'}_{\mathbf{j}}\ln(\mathbf{y_{jt}})} \leq 1 \tag{2.29}$$

Equation (2.23) can be consequently rewritten as:

$$\begin{aligned} \mathrm{E}(\widehat{\boldsymbol{\beta}}^{*}) &= \mathrm{E}([\mathbf{y_0^{*}}'\mathbf{y_0^{*}}]^{-1}\mathbf{y_0^{*}}'\mathbf{y_t^{*}}) = \mathrm{E}\big([\ln(\mathbf{Gy_0})'\ln(\mathbf{Gy_0})]^{-1}\ln(\mathbf{Gy_0})'\ln(\mathbf{Hy_t})\big) \\ &= \mathrm{E}([\ln(\mathbf{y_0})'\mathbf{G}'\hat{\mathbf{c}}'_{\mathbf{0}}\hat{\mathbf{c}}_{\mathbf{0}}\mathbf{G}\ln(\mathbf{y_0})]^{-1}\ln(\mathbf{y_0})'\mathbf{G}'\hat{\mathbf{c}}'_{\mathbf{0}}\hat{\mathbf{c}}_{\mathbf{t}}\mathbf{H}\ln(\mathbf{y_t})) \end{aligned} \tag{2.30}$$

In a situation as the described in Arbia and Petrarca (2011), where the equation is linear ($\hat{\mathrm{c}}_{\mathrm{tj}} = \hat{\mathrm{c}}_{0\mathrm{j}} = 1$; $\mathrm{j} = 1, \ldots, \mathrm{m}$) and the aggregation scheme is the simple sum of spatial units ($\mathbf{G} = \mathbf{H}$) makes (2.30) to be equal to equation (2.16) and the OLS estimator is unbiased. $\widehat{\boldsymbol{\beta}}^{*}$ will be biased, however, in situations that depart from that baseline. The specification of a β-convergence equation as depicted in (2.25), with non-linear relations and different aggregation schemes in the dependent variable and the regressor makes the OLS estimation biased, depending the sign of the bias on the relationship between the matrices $\mathbf{G}$, $\mathbf{H}$, $\hat{\mathbf{c}}_{\mathbf{0}}$ and $\hat{\mathbf{c}}_{\mathbf{t}}$.

The scale effect in the estimation of the β-convergence equations leads, in summary, to estimates that can be biased and with higher variance than in the original disaggregated equations. The next part of this chapter explores by means of a numerical simulation the empirical implications of this problem.

II.5. Simulation with spatially disaggregated and aggregated data

Once the effect of the aggregation level on the OLS estimator has been studied, it is important to quantify its consequences when applied to the empirical analysis of β-convergence. A numerical experiment is conducted in this part with this purpose in mind. Our experiment assumes that the data are generated at the level of $i = 1, \dots, n$ basic spatial units by the following equation that determines the growth in the relevant variable as:

$$\Delta \ln y_{ij} = \alpha + \beta \ln(y_0)_i + \varepsilon_{it} \text{ ; or } \ln(y_i) = \alpha + (1+\beta)\ln(y_0)_i + \varepsilon_{it} \qquad (2.31)$$

being y_{i0} the value of the relevant variable at the starting period and y_i its final value. In the experiment we have arbitrarily set the value of the intercept α at 1.1, and $\boldsymbol{\varepsilon} \sim N(0{,}0.5)$. The idea is to compare the OLS estimates of parameter β, which is the key element in the analysis of β-convergence, in two situations that vary on the spatial scale on which the data are observed:

i. the reference situation or benchmark, that assumes that we have data observable at the same scale at which they are generated, i.e., for the $i = 1, \dots, n$ basic spatial units
ii. a case where the data are only observable at an aggregated spatial scale into $j = 1, \dots, m$ units. In this second scenario, we assume that we only have data on y_j^* and y_{j0}^* and from them we estimate the parameters of the equation:

$$\ln(y_t^*)_j = \alpha + (1+\beta)\ln(y_0^*)_j + \varepsilon_{jt}^* \qquad (2.32)$$

In order to have a numerical experiment as realistic as possible, we have taken as reference for simulating possible structures of aggregation of spatial data the real sub-regional and regional divisions in three different countries: namely the U.S., Germany and Chile. These three countries are taken as examples of developed economies, each of them presenting a particular configuration in their regional divisions. For example, the basic spatial units for the case of Chile are the comunas ($n = 100$) that form the total of $m = 13$ administrative regions. Similarly, in Germany we can find the basic spatial units defined by the concept of kreise ($n = 393$) that are aggregated into $m = 14$ länders. Finally, the U.S is divided into n=3,088 counties that are aggregated forming the m=50 states.

In order to provide with sensible values to the growth equation depicted in (2.32), we have taken real data for the initial value of the variable of interest. In the cases of the U.S and Chile, we have defined y_{i0} as the income per capita, while in the case of Germany –due to data availability at the desired spatial scale- it is defined as GDP per capita. The time span on which we estimate (2.32) is also different for each country and conditioned by data limitations: for the U.S. there is a series of income at county level from 1969 to 2011 published by the Bureau of Economic Analysis; in Chile we have data on income for the comunas between 1996 and 2006 available in the Casen Survey of the Ministry of Planning; and for Germany the Destatis Statistisches Bundesamt contains estimates of GDP for the kreise between 2000 and 2011. Additionally, data on population are required to have indicators of income or GDP per capita. We have opted for using real data on population as well. Note that data of population in the initial and the final periods are required in order to aggregate spatially the per capita values of the variable of interest. The values per capita in the initial and final periods -the explanatory and dependent variable in (2.32), respectively- are aggregated by weighting the values in levels at the scale of basic spatial units by their population shares on these periods. Summary statistics of all these variables can be found in Table 2.1.

Table 2.1. $\boldsymbol{\beta}$ –convergence equation for the EU-28 (2000-2011) for different spatial scales.

	NUT1	NUT2	NUT3
β	-0.31***	-0.26***	-0.23***
Constant	3.36***	2.85***	2.52***
λ	3.37%	2.74%	2.38%
R^2	57.72	44.82	36.44
N	98	272	1305

Source: Eurostat REGIO database, ESA-1995. The speed of convergence ($\boldsymbol{\lambda}$) is obtained from the following expression: $\boldsymbol{\lambda} = \frac{-\boldsymbol{ln}(1+\boldsymbol{\beta})}{\boldsymbol{T}}100$, being T the number of years. *** represents estimates significantly different from zero at 1%.

All these pieces of information have been used for the data generating process described in equation (2.31). The key element on this equation is the parameter β, whose value determines if we have a process of convergence –if negative- or divergence –if positive-. In the experiment, different scenarios have been considered depending on the value of parameter β, setting its values ranging between -0.3 and

0.3. For each value of the parameter and for each country we have simulated 5,000 trials and we have estimated the parameter by applying OLS in scenarios i) and ii).

Table 2.2 summarizes the results obtained on each case, reporting the true value of the parameter together with the average OLS estimate, the empirical variability of the estimates –standard deviation– and a measure of deviation –mean squared error- between the true values and the OLS estimates.

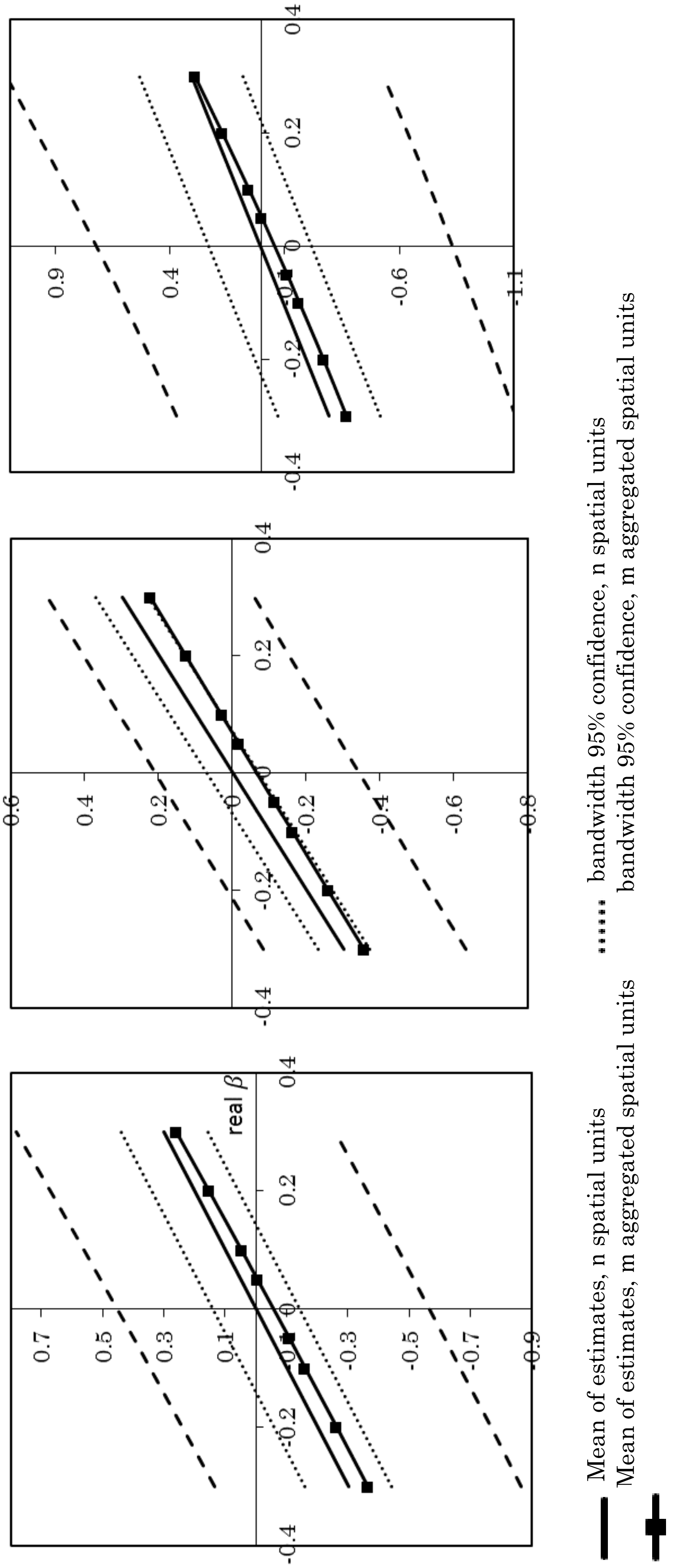
Table 2.2 Results of an OLS estimation with different spatial configurations. 1,000 trials.

	Germany (2000-2011)		U.S. (1969-2001)		Chile (1996-2006)	
True β	n = 393	m = 14	n = 3,088	m = 50	n = 100	m = 13
	-0.301	-0.364	-0.302	-0.357	-0.293	-0.367
-0.30	(0.072)	(0.255)	(0.036)	(0.140)	(0.115)	(0.378)
	[0.005]	[0.069]	[0.001]	[0.023]	[0.013]	[0.147]
	-0.201	-0.263	-0.202	-0.260	-0.193	-0.267
-0.20	(0.072)	(0.257)	(0.036)	(0.140)	(0.115)	(0.384)
	[0.005]	[0.070]	[0.001]	[0.023]	[0.013]	[0.152]
	-0.101	-0.160	-0.102	-0.164	-0.093	-0.162
-0.10	(0.072)	(0.258)	(0.036)	(0.141)	(0.115)	(0.39)
	[0.005]	[0.070]	[0.001]	[0.024]	[0.013]	[0.156]
	-0.051	-0.109	-0.052	-0.115	-0.043	-0.109
-0.05	(0.072)	(0.259)	(0.036)	(0.141)	(0.115)	(0.393)
	[0.005]	[0.071]	[0.001]	[0.024]	[0.013]	[0.158]
	0.049	-0.004	0.048	-0.019	0.057	0.0007
0.050	(0.072)	(0.261)	(0.036)	(0.141)	(0.115)	(0.401)
	[0.005]	[0.071]	[0.001]	[0.025]	[0.013]	[0.163]
	0.099	0.048	0.098	0.029	0.107	0.057
0.1	(0.072)	(0.262)	(0.036)	(0.142)	(0.115)	(0.405)
	[0.005]	[0.071]	[0.001]	[0.025]	[0.013]	[0.166]
	0.199	0.154	0.198	0.125	0.207	0.173
0.2	(0.072)	(0.264)	(0.036)	(0.142)	(0.115)	(0.413)
	[0.005]	[0.072]	[0.001]	[0.026]	[0.013]	[0.171]
	0.299	0.261	0.298	0.221	0.307	0.292
0.3	(0.072)	(0.266)	(0.036)	(0.143)	(0.115)	(0.42)
	[0.005]	[0.073]	[0.001]	[0.027]	[0.013]	[0.178]

Average estimates are reported for each true value of parameter β. Empirical standard deviations are shown in parentheses. Mean squared errors between true values and estimates are shown in brackets.

Additionally, Figure 2.1 visually illustrates the results of the simulations reported in Table 2.2. In these plots the x-axis represents the true value of the β parameter considered in equation (2.32). For each value of β, the mean estimate obtained in the 5,000 trials using disaggregated or aggregated data is represented in the y-axis. If the results were not biased, we would expect a 45º line crossing the origin of the two axes with the true values and the estimates. 95% confidence bandwidths are also plotted, based on the normal distribution of the estimates.

Figure 2.1. OLS estimator with local and aggregate data, 1000 replications.



As expected, the empirical variability of the OLS estimates are substantially lower when estimated from the n basic data points than in the case of the m aggregated spatial units, since the sample size are smaller when working with aggregate data. Not surprisingly, these differences are more remarkable for the case of the U.S. when compared with the other two countries in the experiment, given that the ratio r = m/n is much smaller for the U.S. The loss of efficiency derived from estimating equation (2.32) with m aggregated regions instead of estimating (2.31) with n spatial units is not entirely produced, however, by this inflation of the variance. One substantial part can be attributed to the bias as stated in equation (2.23). The estimates based on aggregated data present a negative bias underestimating the true value of the β parameter. The negative bias is partially a consequence of populations generally more uniformly distributed within each type of aggregated region (U.S. states, German länders or Chilean regiones) in the final period (2011 for the U.S. and Germany and 2006 for Chile) than in the initial one (1969 for the U.S., 2000 for Germany and 1996 for Chile).

Although the simulations have been made for countries with different characteristics and spatial configurations, the results seem to be robust. As expected, the mean of the OLS estimates with n data points are practically equal to the true coefficient. In contrast, for each value of the true parameter, the regression based on aggregate regions tends on average to estimates smaller than the real coefficient. The mean bias of the eight values set for parameter β in the simulation is -0.051 for Germany, -0.061 for the United States and -0.082 for Chile. In summary, the effect produced by the aggregation of the spatial units in our experiments negatively biases the conclusions drawn from the OLS estimation of β-convergence equations.

II.6. Discussion of the results

The study of convergence is one of the more prolific research lines in the literature on regional economics. Empirical analysis is fundamental to evaluate the different theoretical paradigms in economic growth that have opposite conclusions about the persistence of the differences among territories along time. Conclusions derived from convergence analysis provide the support to maintain, reduce or increment expensive policies, such as the Regional Cohesion Policy in the EU. Different improvements have been proposed in the estimation techniques applied to quantify empirically the speed of convergence or divergence among territories. However, most of this empirical literature does not pay attention to how relevant could be the geographical scale in which the convergence is measured, although one of the most important differences among neoclassical theoretical equations and other alternative approaches is the spatial scale in which economic growth is studied.

The objective is to provide an evaluation of the empirical consequences on changes in the spatial scale in the most commonly used approach for convergence analysis: the estimation of equations of β-convergence. The characteristics of an OLS estimator applied to cross-sectional data -which is a relatively common situation in empirical studies-, are derived. We found that geographical aggregation produces estimators with higher variance –part of it produced by the reduction in the sample size-, but also biased if compared with the OLS estimator based on the original disaggregated spatial units.

To provide quantitative evidence about the effect of the spatial scale in β-convergence analysis we conduct numerical simulations with different spatial configurations of real countries: Germany, U.S. and Chile. The results in the simulation confirm the loss of efficiency caused by the aggregation of spatial data, some of which is due to differences in sample size, but the negative bias generated is also significant. One important implication derived from our results is that the estimation of β-convergence equations based on aggregated data should take into account that an important part of the information, related with intra-regional dynamics, could be missing.

Our results, however, do not necessarily indicate that estimates of β-convergence equations with aggregated data are misleading or not useful: in some situations the availability of spatially disaggregated data is very limited and some type of aggregation is required. In addition, in economies where aggregate regions are characterized by low levels of intraregional heterogeneity, aggregation of spatial data could be not a real issue when dealing with convergence analysis. Our results, however, suggest that the spatial scale on which data are taken for estimating β-convergence equations should be carefully defined, since this specification can be partially affecting the conclusions of the analysis.

There are relevant issues not studied here that would require further research. For instance, this chapter studied the MAUP effect on a simple OLS estimator with cross-sectional data. The proliferation of time series with regional data has made possible, however, applying estimators based on a structure of panel data. The consequences of spatial aggregation in the context of estimators applied to dynamic panels are an important issue that should be included in the research agenda on the estimation of β-convergence equations.

III. CONVERGENCE ANALYSIS OF SPATIAL SCALE: FROM THEORY TO EMPIRICS

III.1. Measuring the location effect

Following the results of part II.4, aggregation in regions could modify the results even in the best scenario. In addition, local estimation of the convergence equation allows us to measure idiosyncratic processes of the local level. These processes introduce an important heterogeneity within the regions which can also affect the prosperity of a territory. One of these processes is the location of the activity with-in the regions. The location structures the activity of within the territory depending on the distance towards the market. So, estimations with large aggregates could introduce an artificial homogeneity, avoiding to observe this kind of movements.

Location of the activity is one of the most important topics of the regional economics. The interaction between territories and the distance to cores of production and demand is one of the key elements in regional models. However, how does the distance to the main market configure a country? A favorable position could generate growth and affect the process of convergence while it could damage others. Nevertheless, firms could prefer to go to a close dynamic territory with a big market and possible agglomeration economies. When this process dependence on the space, we could end in a country with important problems of integration.

Mexico is probably the best scenario to study the location effects over convergence both in time and space. The physical proximity to the U.S. market of the regions in the north of Mexico implies a geographic advantage. These territories could benefit from this proximity to the U.S. – main destination of the Mexican exports – with a higher creation of firms, trade, and mobility. In addition, in 1994, this country entered in the North American Free Trade Agreement (NAFTA). Nevertheless, due to the size of its states, it becomes clear that this type of estimation has to be done with local data. The size of the states – the mean area is 61,265 km^2 – generates a homogeneity between states that would mask a strong heterogeneity within the states.

The processes of continental economic integration, such as the North American Free Trade Agreement (NAFTA) or the European Union have an obvious impact on the balance of international trade as well as on the specialization and economic structure of the member countries. Likewise, these processes also affect the spatial distribution of economic activity, as well as the evolution of inequalities between regions.

In Paelinck and Polèse (1999), a theoretical approach is proposed to explain how a process of economic integration can affect regional development axes and alter existing spatial economic dynamics. Their proposal integrates the theoretical framework of the new economic geography (Krugman, 1991; Krugman and Venables, 1995 and Fujita et al., 2001) into classical approaches to urban and regional science. Assuming that both a continental economic core and a national economic core are identifiable and that each country has dynamic –central- and peripheral regions, the authors differentiate between core and peripheral locations according to their position with respect to important production and demand centers at any given time. The key to their approach lies in that these centers can undergo major changes when the process of continental economic integration starts. The continental economic core can alter the balance between dynamic and peripheral regions with varying intensity depending on where the national economic core is located before integration. There are several possibilities that attempt to simplify two basic scenarios: (i) cases where the continental core reinforces the dynamics of the national core and (ii) cases where tension between the continental and the national core is generated due to occupying different positions. Canada is an example of the first case since its major cities (Montreal, Toronto, Vancouver, etc.) are located on the U.S. border, whereby continental dynamics will strengthen national dynamics. Mexico, however, is a good example of the second scenario, where Mexico City, in the center of the country, counterbalances the strong effect of the Northern border with the United States. The authors predict that, in the case of Mexico, the intense development of NAFTA regulations will generate a growing tension between the center and the North which may reconfigure the most dynamic regions towards intermediate territories. The authors also conclude that a process of specialization can be expected to occur, with the center specializing in services, while the North concentrates on manufacturing, as well as a growing gap developing between these areas and the country's Southern states.

The work of Combes et al. (2008) helps us determine what the intra-territorial dynamics generated in this tension between the North (bordering the U.S.) and the center (Mexico City) may be like. According to Combes et al. (2008), companies will establish themselves in regions with a high market potential and leave regions with poor access to markets. The market potential depends on the position of each region with respect to the country's major cities, but also on its accessibility to the North, which is especially important in the manufacturing and export sectors. This means that not only is increasing impoverishment predictable in the South with respect to the North, but that within the Northern area, between Mexico City and the U.S. border, there will be major interstate differences depending on the specific position of each location and its accessibility to both the country's and the continent's major cities and transportation networks.

The aim of this chapter is to examine the position effect of the municipalities through a β-convergence analysis of the integration process of the case of Mexico

Among the many papers on convergence applied to the case of Mexico, the following stand out for being the most recent and for their use of more advanced estimation techniques: Gómez-Zaldívar and Ventosa-Santaulària (2009), Carrion-i-Silvestre and German-Soto (2009), Villarreal and Tykhonenko (2007), Rodríguez-Pose and Sánchez-Reaza (2005), Aroca et al. (2005) and Sánchez-Reaza and Rodríguez-Pose (2002). Particularly, in Gómez-Zaldívar and Ventosa-Santaulària (2009), Villarreal and Tykhonenko (2007), Sánchez-Reaza and Rodríguez-Pose (2002) and Rodríguez-Pose and Sánchez-Reaza (2005) is found that Mexican States doing more trade with the U.S. grew faster than others, but that there was no significant change in this pattern after NAFTA. They do find evidence that the economic pull of Mexico City lessened after NAFTA, lending support to the hypothesis that trade has decreased agglomeration in Mexico. Gómez-Zaldívar and Ventosa-Santaulària (2009) underline that trade reforms negatively affected Mexico City and the poorest States in Mexico, while Carrion-i-Silvestre and German-Soto (2009) find convergence, but mainly during the eighties, which is to say that while a convergence process continued after NAFTA, it was less intense. They also find that Northern States converged faster than the rest of the country, widening the disparity between the Northern States and the rest of the country. In contrast, Aroca et al. (2005) do not find that NAFTA substantially changed growth patterns in Mexico, and instead argue that agglomeration has emerged in the form of several income clusters.

The main contribution of this analysis to the aforementioned empirical literature is to propose a convergence analysis that allows us to identify how North (U.S. border) - Center (Mexico City) tensions are affecting regional disparities in Mexico, introducing a conventional conditional β-convergence model to which we incorporate the effects of location with respect to the North border. To do so, we will propose a modification in the basic convergence equation (2.2) in line with the methodological proposal described by González Rivas (2007).

All previously mentioned empirical studies use State-level data, which masks the spatial distribution of economic activity and severely restricts the number of their observations. This analysis, in contrast, applies this approach to local data – municipalities– in order to observe the intra-State differences that may be occurring. For example, agglomeration economies are positive externalities that appear due to the spatial concentration of economic activity. Urban economic theory expects companies to obtain productive advantages by locating themselves in close proximity to other firms, and that these benefits can explain the formation and growth of cities and industrial locations (Marshall, 1890). The main sources of agglomeration externalities arise from improved opportunities for labor market pooling, knowledge interactions, specialization, the sharing of inputs and outputs, and from the existence of public goods. As the scale and density of urban and industrial agglomerations grow, the external benefits available to companies are also expected to increase (Graham, 2006).

Finally, we propose applying this analysis to a broad time frame, 1980-2010, which is possible through data from the Mexican economic census. The availability of a period spanning four decades allows us to distinguish between the pre-NAFTA and post-NAFTA phases in order to observe the effects that the trade agreement has had.

III.2. Estimation strategy

This part explains the specification of equation (2.2) and (2.4) for the particular case of a panel estimation for Mexico. We base our work on the approach by González

Rivas (2007), which we adapt in this analysis. As in equation (2.4), the control variables are introduced as follows.

The increase in capital per capita is expressed in Equation (3.1). The effect of the geographical position of each place with regards to a specific relevant point, as could be the distance from each municipality to the U.S.-Mexico border, is introduced in this component as D_i. The coefficient of this variable will indicate the effect of the U.S. economy on Mexican growth. The interaction of $\ln D_i$ with $\ln y_{it-1}$ represents the effect on the convergence of the region. We can further incorporate other factors of economic growth, such as specialization in industrial activities, by means of a location quotient:

$$\Delta \ln k_{it} = p \ln k_{it-1} + \beta \ln y_{it-1} + r \ln D_i + \vartheta \ln y_{it-1} \ln D_i + \rho \ln LQ_{it-1} \quad (3.1)$$

Equation (3.1) includes (i) the capital stock in each territory, which is represented by k_{it-1}; (ii) the convergence effect described by Barro et al. (1991), $\ln y_{it-1}$; (iii) the effect of position by means of distance with respect to some specific point, $r \ln D_i$; (iv) the interaction of this distance with the convergence effect $\ln y_{it-1} \ln D_i$; and, finally, (v) a component which takes into account cumulative processes with an index of specialization in the manufacturing industry $\ln LQ_{it-1}$, which controls for processes of agglomeration highlighted in urban economics models (for an example, see Fujita et al., 2001).

Furthermore, we explain innovation as the sum of three components, as represented in equation (3.2):

$$\Delta \ln A_{it} = g \ln h_{it-1} + s \ln D_i + n \ln D_i \ln h_{it-1} + \delta \ln E_i + \sum_{t=1}^{T} \gamma_t T_t \quad (3.2)$$

which are: (i) human capital h_{it-1}, a component of technological change in line with models of endogenous technological change as in Romer (1990); (ii) the effect of the position D_i; (iii) the interaction with human capital $\ln D_i \ln h_{it-1}$; (iv) an exogenous component as the influence from unpredictable shocks E_i; and, finally, (v) a sum of dummies T_t for each period, in order to consider the homogeneous and neutral technological change from the point of view of Hicks, where the first period is taken as a reference.

Combining (3.1) and(3.2), our final model to estimate with the control variables is:

$$\begin{aligned}\Delta \ln y_{it} = {} & g \ln h_{it-1} + s \ln D_i + n \ln D_i \ln h_{it-1} + \delta \ln E_i + \sum_{t=1}^{T} \gamma_t T_t \\ & + \alpha(p \ln k_i + \rho \ln LQ_{it-1} + \beta \ln y_{it-1} + r \ln D_i \\ & + \vartheta \ln y_{it-1} \ln D_i) + (\beta + \alpha - 1)\Delta \ln L_{it} + \Delta \ln \varepsilon_{it}\end{aligned} \tag{3.3}$$

The expression of the convergence effect in equation (3.3) would be:

$$\frac{\partial \Delta \ln y_{it}}{\partial \ln y_{it-1}} = \alpha(\beta + \vartheta \ln D_i) \tag{3.4}$$

The derivative has the usual β-coefficient and the effect of the proximity to the U.S. border. These components will be the key elements in our research.

Equation (3.3) shows a problem in terms of estimation. The use of classical panel estimators – fixed effects and random effects models - is not possible to estimate equation (3.3). Distance D_i represents a time-invariant effect, so its effect cannot be estimated with the approach of a fixed effects model as in González Rivas (2007), Sánchez-Reaza and Rodríguez-Pose (2002) or Cuadrado-Roura (2001) – see Wooldridge (2011).

As a result, a linear estimation with pooled data and including cross-sectional heteroscedasticity covariance structure (see Greene, 2012) and time dummies was chosen. Given the structure of panel in the dataset, this model allows for including a variance term (σ_i^2) for each spatial unit. It specifies a heteroscedastic group variance covariance matrix (V) as in equation (3.5):

$$V = \begin{bmatrix} \sigma_1^2 I & 0 & \dots & 0 \\ 0 & \sigma_2^2 I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 I \end{bmatrix} \tag{3.5}$$

The Feasible Generalized Least Squares (FGLS) estimator that includes this covariance structure is represented in equation (3.6):

$$\hat{\beta} = [X'V^{-1}X]^{-1}[X'V^{-1}Y] \tag{3.6}$$

where Y and X are the matrices for the dependent and independent variables. The estimator for σ_i^2 is obtained through its estimator - $\hat{\sigma}_i^2$ - in equation (3.7):

$$\widehat{\sigma}_i^2 = \frac{1}{T}\sum_{t=1}^{T}\widehat{\epsilon}_{it}^2 \qquad (3.7)$$

The residuals $\widehat{\epsilon}_{it}$ for $\widehat{\sigma}_i^2$ are initially obtained from the Ordinary Least Squares (OLS) estimator. After the first stage, these are the residuals obtained from the previous iteration. This model can be obtained by iterating until equations (3.6) and (3.7) converge.

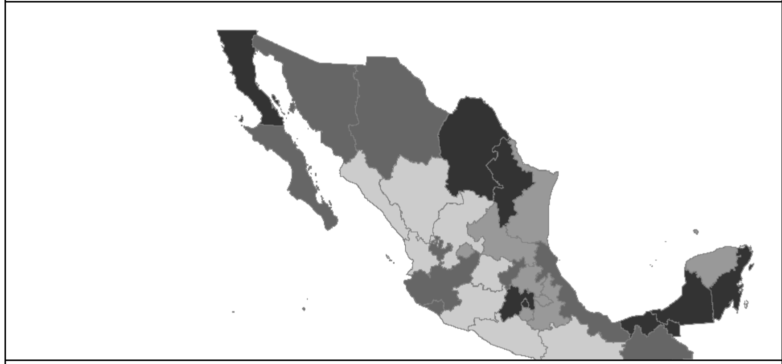
III.3. Database

The database for studying the case of Mexico comes from Mexico's economic census. Every five years the Mexican National Institute of Statistics (INEGI) elaborates this database at different spatial levels (states and municipalities). It provides information on the geographical distribution of the population and economic activity for the period from 1980 to 2010 with 5-year gaps.

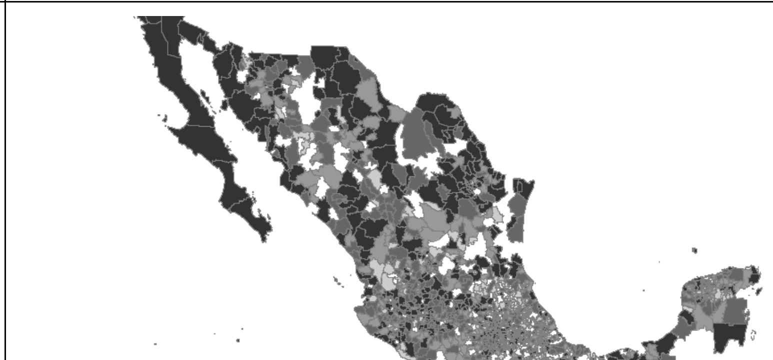
Mexico is divided into 31 states and Mexico City, for a total of 2,377 municipalities. Figure 3.1 represents the Gross Value Added per capita (GVApc) for the states (a and c) and municipalities (b and d), in 1980 (a and b) and 2010 (c and d). This figure also allows us to observe the existence of a pattern in the GVApc distribution The territories of the North normally present higher levels of GVApc and, at the municipal scale, we can see how the higher development extend from North to Center following the main communication corridors of the country. The internal heterogeneity within states can be observed through the Theil Index presented in Table 3.1 This index can be decomposed into two parts: within-state and between-state variability. The results point to an important variability within the states, which is at least 70% of the total variability in the GVA per capita for each year.

Figure 3.1. Gross Value Added per capita of Mexican States and Municipalities, 1980 and 2010.

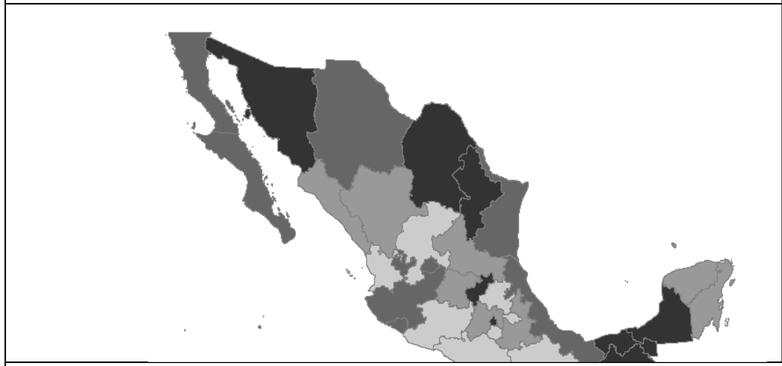
(a) GVApc by States, 1980



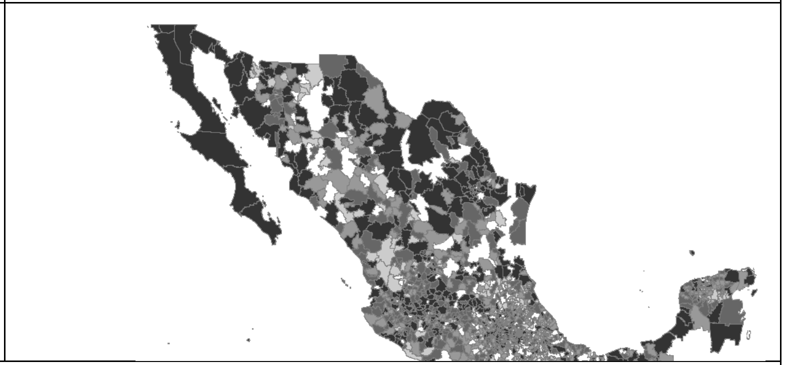
(b) GVApc by Municipalities, 1980



(c) GVApc by States, 2010



(d) GVApc by Municipalities, 2010



Quantiles: Q1 Q2 Q3 Q4

Source: Mexico´s Economic Census. Mexican National Institute of Statistics (INEGI).

Table 3.1. Theil Index decomposition, Mexico 1980-2010.

Theil index	1980	1985	1990	1995	2000	2005	2010
Within	1.779	1.298	1.360	1.153	1.175	1.310	1.495
Between	0.287	0.288	0.518	0.444	0.394	0.412	0.640
Total	2.065	1.586	1.878	1.597	1.569	1.722	2.135

Source: Own elaboration from data in the Mexican Economic Census, several years. Primary sector not included

Focusing the analysis only on between-state variability would neglect the huge differences within the states. Additionally, since the location of each region i is included in the model (D_i), taking states as the spatial unit of the analysis would imply averaging out the same position for all the municipalities within a given state[5], which can be a very unrealistic assumption. Taking into consideration these two issues, the spatial units taken as reference for the empirical exercise in this chapter have been the Mexican municipalities.

In the Mexican Economic Census, GVA is reported in constant pesos of 2004 for all the industries, but data for the primary sector are not available. Excluding the primary sector makes the GVA negative in some small and rural areas. Therefore, municipalities with a negative GVA in any year were eliminated and the final number of municipalities considered in this research was 1,902. However, some municipalities do not have information in all the control variables. From GVA and population data, the dependent variable in (2.2) can be constructed as the GVA per capita growth taking 5-years lags.

Position effects for each municipality i are introduced by variable D_i, which is defined as the distance to the U.S. border. This variable measures the propensity of a territory to be influenced in different aspects by its geographical position. In the first place, there is an important effect on the location of industry due to the trade flows between the two countries. So, it is assumed that the areas closer to the U.S. border have an intrinsic advantage due to their competitive position. In addition, there are also higher foreign investments, greater migration flows to the U.S., and potential spillovers and cultural influence between the two countries. We attempt to use distance as a summary of the different effects of integration with the neighbor country. The distance to the U.S. border, (D_i), is a continuous variable that reflects

[5] This problem is extremely important in Mexico due to the size of the states. For example, the state of Sonora has 10 municipalities in the border with the United States while the furthest municipality is at a distance of 657 kilometers.

the minimum road distance (in thousands of kilometers) from municipality i to the closest U.S. border crossing point.[6]

The estimation of human capital (h_{it}) is introduced by means of a proxy for education, as in Mankiw et al. (1992). This proxy is the percentage of the population with a college degree. Capital stocks (k_i) can be difficult to measure, especially at a local level. Additionally, there are not official estimates of capital stocks at state or municipal scale for Mexico. As a result, we introduce a dummy variable for each state to take into account this component of the steady state.

Specialization is considered through a location quotient (LQ_{it}) for the industrial sector of each territory.[7] This coefficient compares industrialization in a municipality with the rest of the country. Therefore, it is not surprising that the mean of this index is near to one, but it shows substantial variability, which indicates a possible polarization of the economy. Finally, labor force growth (ΔlnL_{it}) is introduced with the common measure of population growth in that municipality. A summary of the variables included in the model are reported in Table 3.2. [8]

Table 3.2. Statistical description of the variables (1980-2010).

Concept	Variable	Mean	Std. Dev.
GVApc	y_{it}	6.704	41.641
Human capital	h_{it}	0.031	0.038
Location Quotient	LQ_{it}	0.965	5.453
Population growth	ΔlnL_{it}	0.055	0.124
Distance to the U.S. border	D_i	1.066	0.443

[6] To create this variable, we first obtained the name of the municipality (INEGI, 2008). Second, we calculated the road distance from each of the municipalities to the different U.S. border crossing points, by entering in the points of origin and destination on the webpage "Traza tu Ruta" (Route Planner) provided by the Secretaría de Comunicaciones y Transportes (2008). Finally, we chose the shortest distance for each municipality from the different distances provided by each U.S. border crossing point.

[7] We use the standard formula $\frac{e_i^{ind}/e_i}{E^{ind}/E}$. This coefficient will be greater than one if the percentage of employment in the industry is higher than in the rest of the country. It measures the relative concentration of industry in the region.

[8] The variables in levels, such as human capital and the specialization of each territory, have been lagged following the specification in (3.3).

III.4. Results: effects of continental integration on convergence dynamics in Mexico

Equation (3.3) is estimated by FGLS in two different periods: the first period covers from 1980 to 1995, while the second spans from 1995 to 2010, which allows for a comparison of the results. This equation includes the traditional factors available in our datasets as well as time and state dummies. Table 3.3 summarizes the results.

Table 3.3. Conditional β-convergence model with position effects for Mexico.

	1980-1995	1995-2010
$\ln y_{it-5}$	-0.381***	-0.228***
$\Delta \ln L_{it}$	0.158***	-0.212***
$\ln D_i \ln h_{it-5}$	-0.053***	-0.030***
$\ln h_{it-5}$	0.398***	0.227***
$\ln LQ_{it-5}$	0.110***	0.081***
$\ln D_i$	-0.271***	-0.108**
$\ln y_{it-5} \ln D_i$	0.031***	0.005
Time dummies		
1990	-0.671***	
1995	-0.276***	
2005		-0.042***
2010		-0.250***
Constant	2.116***	1.433***
Municipalities	1854	1893
Periods of time	3	3
Wald χ^2_{40}	9241.79	4611.77

Note: *, ** and *** represent estimates significantly different from zero at 10%, 5% and 1%, respectively.

The estimates show the existence of a clear convergence process in both the two periods studied, which is consistent with the findings of the most recent convergence studies for Mexico such as Sánchez-Reaza and Rodríguez-Pose (2002) and Rodríguez-Pose and Sánchez-Reaza (2005). We can also observe a notable reduction in the β coefficient in the second period, falling from -0.381 to -0.228. This slowing down of convergence after entering NAFTA is also consistent with the recent work of Díaz-Dapena et al. (2016) and Villarreal and Tykhonenko (2007). This result is robust to the inclusion of the standard determinants of the steady state. The estimates indicate that there is a substantial decline in the process of convergence in the last period not caused by the standard determinants of the steady state, which suggests a reduction in convergence right after the NAFTA process starts. However, since all Mexican municipalities are affected by economic integration, it has to be taken into account that this change cannot be assigned only to the integration process, as other common factors could also affect convergence in the same period.

Regarding the control variables, the human capital coefficient is positive in both periods. This coefficient is coherent with the literature that explains productivity growth through human capital, such as in Mankiw et al. (1992). This variable tries to measure that labor productivity is not homogeneous. So, according to this theory, workers with a higher level of education tend to generate a higher value added. In macroeconomic terms, a better-educated population increases the municipality's productivity. The interaction term of this variable with distance would indicate an increase in the influence of human capital created by the influence of the U.S. economy. This variable has a negative and significant coefficient in both periods. This result is expected and consistent with the idea of spillovers. Regions near the U.S. border benefit more from human capital. So, it seems that they tend to have an economy where human capital is more qualified than in the regions in the South.

The coefficient of specialization also has the expected positive sign. This coefficient could indicate the process of location economies (see Rosenthal and Strange, 2001; Beardsell and Henderson, 1999; Porter, 1990). However, it could also be seen as a variable of the division of labor with the advantages in terms of productivity pointed to by classical economics since Adam Smith.

Population growth has the expected sign in the 1980-1995 period with a negative coefficient between 0 and -1. However, in the second period, this coefficient is positive. This result is contrary to the classical hypothesis of a steady state as in Mankiw, Romer and Weil, where population growth has a negative effect. This result could indicate that population growth could also have a positive effect through the other components of equation (3.3).

One important contribution to the previous literature is the inclusion in the estimation of the distance to the Northern border, by which the effect of Mexico's process of continental integration on the spatial distribution of economic activity can be evaluated. Our estimates show that the distance has a negative and significant coefficient in both periods, indicating that municipalities located closer to the U.S.

border, as expected, are more dynamic than those in the rest of the country and that they tend to grow faster. The estimate of the coefficient of the distance, however, reduces considerably in the post-NAFTA period. This may seem counterintuitive at first, as integration can be expected to reinforce the importance of being closer to the Northern border, but it is perfectly consistent with what the Paelinck and Polèse (1999) model predicts. Integration supports the tension between the continental core and the national core, pulling on the dynamic regions all along the entire strip of states above Mexico City, as is predicted in their paper. This diminishes the importance of being close to the border and reduces the coefficient of the D_i variable, although it is still significant and negative, indicating that being located in the North of the country continues to be a growth factor. The estimates of the coefficients of the interaction term of distance and per capita income reinforce these previous conclusions. In the first period, 1980-1995, municipalities far from the border tend to converge more slowly and have a positive and significant coefficient of the interaction term. Consequently, the convergence effect can be seen as decreasing with distance.

The marginal effects are useful to evaluate the total effect of the independent variables, including the interaction terms, in order to have a more accurate estimation of the effect of the explanatory variables including in (3.3). Marginal effects estimated in the mean for the variables with interaction terms can be seen in Table 3.4. This analysis confirms the decline in the process of convergence in the second period, while the effect of the distance to the border is reduced to a non-significant effect on growth in the second period when it is evaluated in the mean.

Table 3.4. Calculation of the marginal effects in the mean for the case of Mexico.

	1980-1995	1995-2010
$\ln y_{it-5}$	-0.385***	-0.228***
$\ln D_i$	-0.054***	-0.002
$\ln h_{it-5}$	0.398***	0.227***

Note: *, ** and *** represent estimates significantly different from zero at 10%, 5% and 1%, respectively.

A visual representation of the marginal effect of the variable $\ln y_{it-5}$ interacting with the distance to the U.S. border in the periods 1980-1995 and 1995-2010 can be seen in Figure 3.2.

Figure 3.2. β-convergence effect

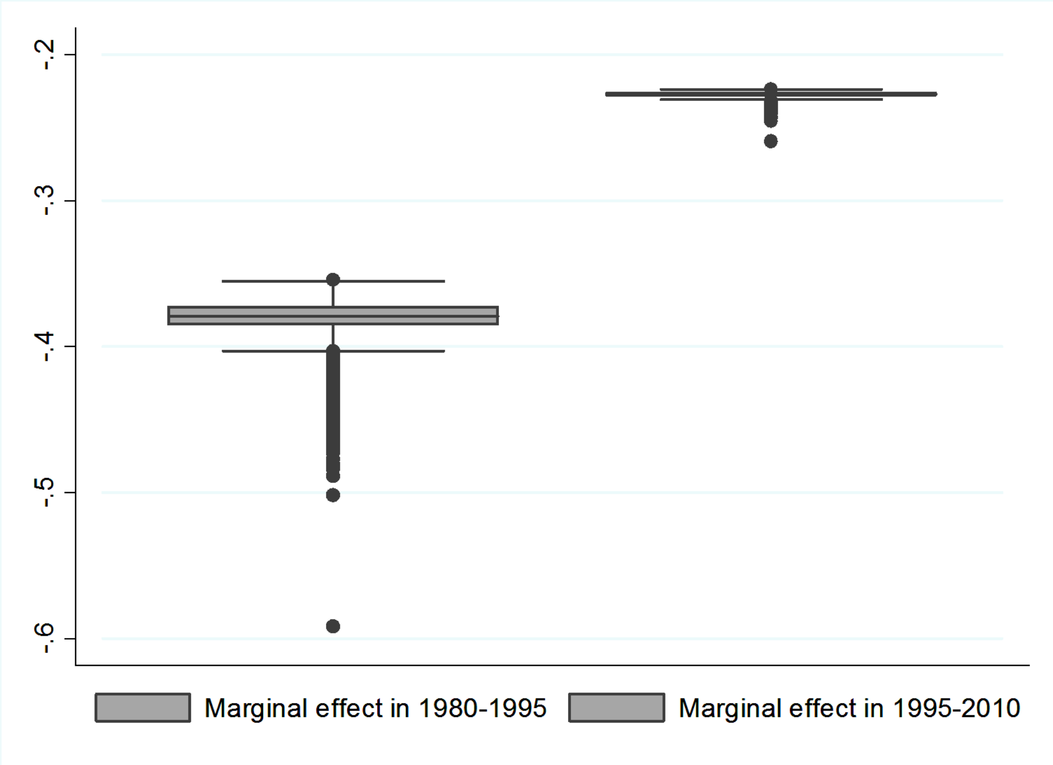


Figure 3.2 proposes a visual representation of all the marginal effects using a box-plot graph. In the left-side of the graph the period from 1980 to 1995 is presented and in the right-side from 1995 to 2010. The axes of the graph show the variability in the marginal effect of the variable $\ln y_{it-5}$ interacting with the distance to the U.S. border. This figure illustrates, how the process of convergence is significantly different across the distance to the border in the period 1980-1995. The β–convergence estimation goes from -0.592, in the north of the country, to -0.354 in the south part of the country. However, this difference becomes non significant in the period 1995-2010. It can be seen that the β–convergence takes a minimum value of -0.259 and a maximum of -0.224. As a result, the variability is much lower in this last period of analysis, as was predicted by Paelinck and Polèse (1999).

III.5. Some consequences

This estimation provides an example of the advantages of using the local level to include idiosyncratic processes of the local level in the estimation of convergence. This type of movement is especially important if we take into account that at least 70% of the variability of the GDP per capita is located with-in states in the sample.

The obtained results are consistent with those in the main studies recently done on Mexican convergence and show how convergence slows down in the post-NAFTA phase, indicating that this agreement has curbed the pace of reducing regional inequalities. However, when we look at how the distance to the Northern border affects convergence, we find that this factor was more significant before NAFTA, while losing importance after NAFTA. The parameter associated with distance is smaller in the later period, but still significant. This set of results, with lower convergence after NAFTA, but with distance to the U.S. border carrying less weight, is perfectly consistent with what Paelinck and Polèse (1999) predict. The existence of a national core in the center of Mexico (Mexico City) counteracts the economic pull to the border, the continental core, exerting a tension that benefits the territories between the border and Mexico City. However, for the South of the country is more difficult to follow the path to convergence.

We also notice that factors such as industrial specialization, level of education and growth of the labor market play an important role in the increase of regional productivity, which, in the long run, will intensify the convergence of the regions. These indicate that better policies will be needed to control the increase of regional disparities. Industrial, educational, and regional development policies must be quickly developed to set up the foundations for growth in all regions. Further research is necessary to determine what other factors influence regional convergence in Mexico. Factors that were previously considered fundamental in growth theory are quickly giving way to different and less known factors that are likely to shape the next phase of Mexico's regional development.

Our analysis opens the discussion about the suitability of econometric techniques that are not affected by MAUP problems. In this regard, multilevel estimation (see, among others, Goldstein, 1986; Hox, 2010; or Goldstein, 2011), which allow for using data at different scales is particularly interesting if we want to identify different spatial scales of convergence avoiding the potential bias derived from the data aggregation.

IV. MULTILEVEL ANALYSIS OF β-CONVERGENCE

IV.1.Introduction

We believe it is possible to contribute to the extensive previous literature on regional convergence by pointing out the relevance of the different trends at different spatial levels of disaggregation. Our idea is not completely new in the field of economic geography. Li and Wei (2010) apply it to analyze spatial inequalities in China. However, this is the first time it is used in a β-convergence model. Furthermore, it could constitute an important issue for convergence studies for a number of reasons. First, the theories explaining economic convergence do not specify the exact spatial level at which the economy operates. In other words, researchers do not have sufficient information to decide whether convergence should be measured at a county, regional, province or state level (Hoover and Giarratani, 1971). Moreover, researchers do not necessarily have to choose one level as in chapter II. Rather, the econometric strategy should be able to capture all these scales simultaneously in a unified framework. Here, multilevel modeling helps by mixing the entire scale in just one empirical approach. Second, the explicit and prior choice of a specific level of aggregation could have inevitable effects on the conclusions reached. While this problem has been the subject of a significant number of papers in fields like spatial econometrics, we think that its consequences on convergence modeling have not yet been extensively studied. Finally, the correct specification of multilevel structure in the convergence model helps to provide a better estimation of the parameters given the recognition of the clustered errors. So, explicit inclusion of the hierarchy avoid the possible estimation problems seen in part II.4.

Our objective is to propose an estimation of the standard β-convergence regression, but using a hierarchical geographic structure: namely, how regional convergence is shaped when we accept that counties and states interact simultaneously to define a catch-up condition. Our aim is to contribute to both these perspectives by: (i) presenting a proposal for the estimation technique which overcomes the limitations of standard procedures, simultaneously using different regional scales; and (ii) finding evidence regarding the possibility of there being overall patterns of convergence which are coherent with heterogeneous intra-regional trends. We use the widely known economy of the Europe as an example to compare our estimations with previous research, thus allowing the generation of debate using the generally accepted existing literature.

IV.2. Possibilities of the multilevel approach in convergence analysis

The multilevel technique (Goldstein, 1986, for more information see Hox, 2010; Goldstein, 2011) has been widely used in different disciplines. Most of the applications in Economics have been made in labor market studies (Andersson et al., 2013; Cohen, 1998). Nowadays, however, it is being successfully introduced in other types of economic studies (Li and Wei, 2010; Srholec, 2010).

This methodology has its advantages and drawbacks. Multilevel analysis achieves considerable simplicity and efficiency due to the numbers of parameters estimated. Compared with using dummy variables (see De la Fuente, 2002), we only need one parameter to introduce variability on the intercepts, three on both slopes and the intercepts. Furthermore, a general component and a parameter for each group can be observed in this estimation without the problem of multicollinearity. As a result, the effect on the dependent variable is completely separated into two components. Finally, as Goldstein (2011) states, 'there is a controversy about the proper unit of analysis. This problem is solved using solved by explicit hierarchical modelling.'

However, the multilevel methodology has a number of drawbacks for our analysis. It does not include any kind of spatial structure within the groups. Hence, a bias in the effects could appear, similar to an omitted variable (see Anselin, 1988). In order to apply the methodology to a more complex model, like conditional convergence, it would be necessary to introduce this dependency of the data. Furthermore, the problem of ecological fallacy is not completely solved. This method only takes into account the hierarchy that is explicit in the model; in our case, the relationship between counties and states. However, any other possible aggregation of the areas in a different scale or shape is not fixed by a multilevel model (see Openshaw, 1984).

Despite the problems of multilevel analysis, this approach can separate the variability in two components if we are able to identify the different levels. In our case, we can apply it using the different degrees of spatial desegregation: local and regional level.

We can depart from the β conditional convergence model. As stated in chapter II, this consists in estimating a simple regression of the income per capita or similar variable growth rate of a spatial unit at the initial level of that variable. The approach is thus based in equation (2.2):

$$\Delta \ln y_i = \alpha + \beta \ln(y_0)_i + \theta x_{i0} + \varepsilon_i \tag{4.1}$$

where x_i is a vector of different variables correlated with the steady state. Using this expression, a new hedonic equation with two levels can be introduced. Level i is the local level and level j represents the regions or states in which the local units are aggregated. Thus equation (4.2) would be the multilevel estimation of equation (2.4):

$$\Delta \ln y_{ij} = \propto_{0j} + \beta_0 \ln(y_0)_{ij} + \theta x_{ij0} + \varepsilon_{ij} \tag{4.2}$$

where:

$$\begin{gathered} \propto_{0j} = \propto_0 + u_j \\ u_j \approx N(0, \sigma_u^2) \qquad \varepsilon_{ijt} \approx N(0, \sigma_\varepsilon^2) \end{gathered} \tag{4.3}$$

This kind of model has several advantages over those discussed previously. In this chapter, however, we are particularly interested in only two of them. First, note that equation (4.3) presents two variance components: the variance at level 1 and at level 2. If we could estimate both parameters, then we could evaluate the composition of the total variance. This ratio of variance, namely between level 2 and level 1 variance, is known as intraclass correlation coefficient (ICC) and its magnitude is key to evaluating the necessity of a multilevel model. If the variance at level 2 is significant, then we can use this fact to estimate a random intercept and random slopes. This will enable us to estimate different coefficients for all the regions. As a result, this methodology allows us to study not only the process of convergence in the country as a whole as does the previous literature, but also the steady state of a particular region. Moreover, using the information regarding the variance, the random intercepts are then calculated in a second step. That is why this model also provides an advantage of efficiency, because fewer parameters need to be estimated. From a statistical perspective, the model is estimated using the Restricted Maximum Likelihood (REML) procedure. First, there are both a common coefficient of convergence (β_0) and an intercept for the country as a whole ($\propto_0$). Second, the coefficients of the random part, u_r and ε_{ir}, are σ_u^2 and σ_ε^2, i.e., the random parameters. The coefficients u_r are the regional intercepts. Thus, a positive (negative) intercept will indicate that the counties of that state have a higher (lower) steady state than the rest of the country.

Control variables in this estimation are introduced following Mankiw et al. (1992). They define a Solow model augmented with human capital. As a result, the steady state level ($\tilde{y}^*$) as a function of the rate of savings (s), growth in population (n), growth in technology (g), depreciation of capital (d) and the level of human capital.

The parameter σ_u^2 indicates the variability between groups. However, the Intraclass Correlation Coefficient (ICC) can be used for the purposes of interpretation. This coefficient indicates the percentage of total variation ($\sigma_u^2 + \sigma_\varepsilon^2$) which is due to the differences in regions. It can also be interpreted as the correlation in growth of random territories within a region.

$$\mathrm{ICC} = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2} \tag{4.4}$$

We can improve this initial simple approach, in which the variance could occur on the intercept, for a more complete formulation, in which the variation is possible on the intercept and on the slope (β parameter). The new equation is thus:

$$\Delta \mathrm{lny}_{ij} = \propto_{0j} + \beta_{0j}\ln(y_0)_{ij} + \theta x_{ij0} + \varepsilon_{ij} \tag{4.5}$$

but where:

$$\begin{gathered} \propto_{0j} = \propto_0 + u_j \qquad \beta_{0j} = \beta_0 + v_j \\ u_j \approx N(0, \sigma_u^2) \quad v_j \approx N(0, \sigma_v^2) \quad \mathrm{cov}(u_j, v_j) = \sigma_{uv} \quad \varepsilon_{ijt} \approx N(0, \sigma_\varepsilon^2) \end{gathered} \tag{4.6}$$

This new specification of the convergence equation is known as the Random Slope Model. It has two additional parameters, the variance of the random component on the slopes (σ_v^2) and the covariance between the random components of intercepts and slopes (σ_{uv}). The estimation of this kind of model could suppose an important step in improving unconditional convergence analysis in Regional Science, as it allows us to consider the hierarchy in the process of convergence. An overall coefficient of convergence can thus be estimated with this methodology, as well as an additional process of convergence within each region, aside from the general process of the country as a whole.

The independent variable has been centered in this second model. In the previous model, centering only affects the general intercept, but this second model offers important advantages. The parameters σ_u^2 and σ_{uv} are explained for the value 0 of the independent variable. In this case, that value is ln1, which is meaningless for our analysis. This transformation thus allows us to interpret at the mean. Moreover, the Variance Partition Coefficient (VPC) has to be calculated as in (4.7) and evaluated at a point. After centering, however, we can use (4.4) to evaluate in the mean. This transformation can also be made when there is variability in all the coefficients.

$$\mathrm{Var}[u_j + \ln(y_0)_{ij} v_j] = \sigma_u^2 + 2\ln(y_0)_{ij}\sigma_{uv} + [\ln(y_0)_{ij}]^2\sigma_v^2$$

$$\mathrm{VPC} = \frac{\sigma_u^2 + 2\ln(y_0)_{ij}\sigma_{uv} + [\ln(y_0)_{ij}]^2\sigma_v^2}{\sigma_u^2 + 2\ln(y_0)_{ij}\sigma_{uv} + [\ln(y_0)_{ij}]^2\sigma_v^2 + \sigma_\varepsilon^2} \tag{4.7}$$

Note that the application of the multilevel methodology in convergence research provides more in-depth knowledge of this process. Using this approach, we do not have only the country divided into counties or states; we have a twofold process of convergence that considers this hierarchy. Therefore, we can see how the territories of a state behave, excluding the overall process of convergence. However, this specification does not include any type of spatial effect.

The multilevel model proposed for the β-convergence equation with spatial effects can be viewed in equation (4.8).

$$\Delta \ln y_{ij} = \propto_{0j} + \beta_{0j}\ln(y_0)_{ij} + \theta x_{ijt} + \varepsilon_{ij}$$

$$u_j \approx N(0,\sigma_u^2) \quad v_j \approx N(0,\sigma_v^2) \quad \mathrm{cov}(u_j, v_j) = \sigma_{uv} \qquad \varepsilon_{ijt} \approx N(0,\sigma_\varepsilon^2\Lambda) \tag{4.8}$$

The spatial interactions of equation (2.6) are included in our model through the matrix of variances-covariances of the error term. Using the model described in Pinheiro and Bates (2000). In this model correlation decreases with the distance, indicating the level of correlation for each distance.

The matrix of variances-covariances can be divided in two, as in equation (4.9). The first component is the usual variability of the error term. The second component is the spatial correlation of the observation as a function of the distance between them, depending on the parameter ρ – known as the range. It indicates the minimum distance where there are spatial correlations.

$$\Lambda_i = A_i C_i A_i;\ \mathrm{Var}(\varepsilon_{ijt}) = \sigma^2 A_i^2\ ;\ \mathrm{Corr}(\varepsilon_{ij}, \varepsilon_{kj}) = C_{ikr} = h(d(p_{ij}, p_{kj}), \rho) \quad (4.9)$$

This model is estimated using a Generalized Linear Mixed Model (GLMM), where the correlation between errors depends on a function of the distance between observations. Thus, the Random Slope Model without spatial effects is a concrete case of this model where Λ=I.

In order to estimate the model, it is needed a transformation in the model. Since Λ is a positive definite matrix, $\Lambda^{1/2}$ is invertible. So, the transformated model is:

$$y_i^* = (\Lambda_i^{-1/2})^T y_i \quad , \varepsilon_i^* = (\Lambda_i^{-1/2})^T \varepsilon_i \quad (4.10)$$

So, the distribution of the transformed error of the model becomes:

$$\varepsilon_i^* \sim \mathrm{N}\left[(\Lambda_i^{-1/2})^T 0, \sigma^2 (\Lambda_i^{-1/2})^T \Lambda_i \Lambda_i^{-1/2}\right] = \mathrm{N}(0, \sigma^2 \mathrm{I}) \quad (4.11)$$

It can be seen that the differential of y_i^* is $\mathrm{d}y_i^* = \left|\Lambda_i^{-1/2}\right| dy_i$. So, the likelihood function for the estimation is obtained as:

$$\begin{aligned} L(\beta, \theta, \sigma^2, \lambda|\mathrm{y}) &= \prod p(y_i|\beta, \theta, \sigma^2, \lambda) = \prod p(y_i^*|\beta, \theta, \sigma^2, \lambda) \left|\Lambda_i^{-\frac{1}{2}}\right| \\ &= L(\beta, \theta, \sigma^2, \lambda|y^*) \left|\Lambda_i^{-1/2}\right| \end{aligned} \quad (4.12)$$

Where θ represents a vector of parameters of the variance-covariance matrix of the intercepts and slopes and λ is a vector of parameters which determine Λ_i. Last representation of the log likelihood function in equation (4.12) points to a basic linear multilevel model where standard algorithms for maximum likelihood estimation can be applied - see Hox (2010), Pinheiro and Bates (1996) or Pinheiro and Bates (2000) for further details.

We propose the widely known European economy as an example for the estimation of the multilevel convergence. Equation (4.5) and (4.6) are estimated for this economy. The results would allow us to see whether is suitable a mechanism with forces at both levels. In addition, it is applied a multilevel estimation of convergence which also includes spatial interactions – see equations (4.8) and (4.9).

IV.3.Spatial multilevel convergence in Europe

In this second case, we focus our efforts in the estimation of the multilevel process of convergence for the European economy. This scenario will be suitable for an estimation which also includes spatial interactions.

Since its inception, the EU project has been conceived as a project of not only economic integration but also social and economic cohesion among European nations and regions. One of the main objectives of the EU can be seen in the Annual Growth Survey for 2016 of the European Commission. According to the European Commission, "A renewed process of upward economic and social convergence is needed to tackle the economic and social disparities between Member States and within societies". Within the competences of the institutions of the European Union, there is even an explicit regional policy that attempts to reduce the inequality between rich and poor territories.

Our empirical analysis is carried out with data at the regional level for the European Union. Following Escriba and Murgui (2014), Cambridge Econometrics database provides homogeneous information of the income on all the regions in the European Union in constant terms of 2005. In addition, Eurostat database provides information of the control variables. We need data that are disaggregated at different levels, from the local to the national. Thus, we need information at the NUTS-III regions, the most disaggregated level with information on GDPpc. A total of 1,284 regions from 27 countries are included in the analysis[9]. The period of analysis is 2000-2011 –the longest period available using data at this level of spatial disaggregation–.

To view the spatial distribution of wealth, Figure 4.1 represents the GDPpc in 2000 for the NUTS-III in the EU corrected by Purchasing Power Parity (PPP) of the countries. It is clear that the central areas of the Union concentrate the largest levels of GDPpc. The spatial concentration of the growth and levels of GDPpc is extremely important, especially when it is mixed with the weak process of convergence described above. These two processes together describe a territory with a rich core that other regions find difficult to reach.

[9] Croatia is excluded due to a lack of data.

To apply the model proposed in equation (4.8) to the case of the European Union, we need data on the GDPpc in the PPP of each region and information on the main variables that describe the steady state. In addition, multilevel analysis requires information on the country to which each region belongs, whereas the spatial specification requires the geographic position of the centroids. Table 4.1 summarize the data included in the analysis in terms of equation (4.8) and provides the standard descriptive statistics.

Figure 4.1. Map of the GDPpc in PPS(*) of the NUTS-III in 2000.

Note: (*) Purchased Parity Standard provided at Country level.

The variables needed to describe the steady state are the rate of savings, the growth in labour force, the growth in technology, depreciation and human capital. The rate of investment is provided as the percentage of Gross Capital Formation with respect to GDP. The growth in labour force is estimated as the growth rate of the working population. The growth in technology is introduced through a proxy variable of total R&D expenditure as the percentage of the total GDP. Simultaneously, the depreciation is considered constant and equal to 3%, as in the previous literature. The information used for the $\frac{s}{n+g+d}$ coefficient were provided at a national scale. So, their coefficient will be common in all the sample. Finally, human capital is included through two proxy variables: percentage of population from 25 to 64 years in the

NUTS-II region with level 3-4 of education (Upper secondary and post-secondary non-tertiary education) and the percentage with level 5-8 (Tertiary education)[10].

Table 4.1. Descriptive statistics of the variables.

	Variable	Mean	Std. Dev.	Min	Max
Growth of GDPpc in PPS	$ln\left(\frac{Y_t}{L_t}\right)_{ij} - ln\left(\frac{Y_0}{L_0}\right)_{ij}$	0.131	0.181	-0.46	0.90
GDPpc (€) in PPS in the first period	y_{ij}^0	20491	9918	2631	137283
Percentage of population with level 3-4 of education	h_{1j}	46.936	15.126	7.3	77.8
Percentage of population with level 5-8 of education	h_{2j}	19.843	7.790	3.7	48.9
Gross capital formation as % of GDP	s_j	21.977	2.707	17.8	29.9
Mean growth of the working population	n_j	0.625	0.589	-1.154	2.837
R&D expenditure as % of GDP	g_j	1.732	0.825	0.25	4.13

As a first step before introducing the results of our approach, we explore the OLS estimation of unconditional and conditional β-convergence in our period of analysis. The conditional model includes the variables of the Solow model as shown in equation (4.8). Then, the model becomes more complex to introduce the hierarchy until the final GLMM estimation. These results are summarized in Table 4.2.

The β-convergence coefficient obtained is similar to the results found in the previous literature. The negative and significant coefficient indicates that the convergence hypothesis cannot be rejected. In addition, this result leads to a speed of unconditional convergence of 1.62%. This conclusion is actually similar to the significant and low convergence among the regions obtained the previous literature (e.g., Sala-i-Martin, 1994; Shioji, 1992; Coulombe and Lee, 1993; Cuadrado-Roura and García-Greciano, 1999; De la Fuente, 2002).

[10] educational attainments are defined according to ISCED 2011.

Table 4.2. Random Slope Model of unconditional β convergence of the NUTS-III (2000-2010).

	(1) OLS	(2) OLS	(3) Null	(4) Random slope	(5) GLMM	(6) GLMM
Constant	1.863 ***	-1.194 ***	0.189 ***	-0.604 ***	-0.711 ***	-0.311 ***
lny_{ij}^{2000}	-0.177 ***	-0.111 ***		-0.048 *	-0.044 *	-0.052 **
h_{1j}		0.133 ***		0.027	0.066 **	0.101 **
h_{2j}		0.151 ***		0.058 ***	0.054 ***	0.077 ***
$ln\left(\frac{s}{n+g+d}\right)_{ij}$		0.269 ***		0.336 ***	0.316 ***	0.316 ***
σ_0, Std. Dev. (constant)			0.198	0.138	0.121	0.109
σ_1, Std. Dev. (lny2000)				0.095	0.079	0.075
σ_2, Std. Dev. (h_{1j})						0.124
σ_3, Std. Dev. (h_{2j})						0.059
σ_ε, var (Residual)			0.106	0.102	0.105	0.104
VPC in the mean			77.7%	64.7%	57.0%	52.3%
Range					2.388	2.394
Nugget					0.781	0.794
N	1273	1273	1273	1273	1273	1273
M			27	27	27	27
Speed of general convergence	1.62%	0.98%		0.41%	0.37%	0.45%
Adjusted R^2	25.48%	46.8%				
Log Likelihood	561.6	774.4044	994.6567	1034.027	1061.757	1065.416
AIC	-1117.2	-1536.809	-1983.313	-2052.055	-2103.514	-2106.833
BIC	-1101.74	-1505.914	-1967.857	-2010.862	-2052.101	-2045.138

*Note: (*) Significant at 10%; ** significant at 5%; *** significant at 1%.*

Bearing this scenario in mind, the multilevel methodology is introduced into the analysis to take into account that the NUTS-III has a hierarchy. This estimation adds the importance of the hierarchy through the variances and the correlation of the intercepts and slopes. To explain the results, the independent variables with variability in their coefficient have been centred on their mean.

Following the previous literature (See Mankiw et al., 1992), the human capital and Solow coefficients were also introduced into our model. The coefficient for the different levels of education is positive and significant in the OLS as is the effect of the Solow coefficient. In addition, the coefficient of these variables remain significant in the rest of the models.

The simplest model to test the hierarchy is the null model. This model only includes an intercept coefficient and a random intercept. This model indicates that 77.7% of the variability is generated at the group level before including any information. However, this estimation can be improved with the information pointed by the economic model.

Compared to a linear regression, the multilevel model of random slopes is actually significant. Using a likelihood ratio test, the hypothesis of no significant differences must be rejected (χ^2_2=519.25). In this model, the VPC evaluated in the mean indicates that 64.7% of the total variance is caused by differences between states. It appears as both levels are important in the European Union

Moreover, the general convergence is reduced from a coefficient of -0.177 and -0.111 in the unconditional and conditional models to -0.048 in the conditional multilevel convergence. This change seems to indicate that most of the convergence process is not homogeneous, which is coherent with the high VPC. The VPC indicates that the different models of convergence within the states cause an important proportion of the variability. As a result, it is not surprising that the general level of convergence become less significant when these differences are taken into account.

However, as in the traditional unconditional β convergence estimation, spatial interactions could alter the results. The complete estimation of equation (4.8) is performed through a GLMM model. This procedure includes both the control variables of the steady state and the spatial interactions between neighbours. In addition, GLMM estimation with variability in all the coefficients is also included in Table 4.2. The hypothesis of non significant differences with the simple Random Slope model is rejected through a LR test (χ^2_2=55.46 and 62.78). Therefore, control varibles and spatial interactions seems to improve the model significantly.

The VPC is 57% and 52.30% in these models**.** This finding indicates that the fundamental variables explain a significant part of the differences between countries. The main consequence of this result is that the hierarchy could have a central importance in the process of convergence.

Table 4.3. Country-specific coefficient estimates for regional-level variables.

Country	Constant	$\ln y_{ij}^{2000}$	h_{1j}	h_{2j}
Portugal	-0.142 ***	-0.119 ***	-0.002	-0.035
Spain	0.008	-0.094 **	-0.087	0.003
Greece	-0.313 ***	-0.083 ***	0.076	-0.006
Latvia	0.149 ***	-0.076 ***	0.054	0
Finland	0.091 ***	-0.062	0	0.009
Austria	0.071 **	-0.057	0.009	-0.009
Germany	0.047 *	-0.032 *	0.1 *	-0.071 ***
Belgium	-0.006	-0.027	-0.002	0.005
Italy	-0.154 ***	-0.02	0.006	-0.014
Estonia	0.02	-0.006	0	0
Denmark	-0.136 ***	-0.005	-0.023	-0.01
Lithuania	0.173 ***	-0.005	-0.007	0.03
Sweden	0.149 ***	0.001	0.051	0.017
Cyprus	0.026	0.001	-0.004	0.002
Slovakia	0.193 ***	0.002	0.067	0.005
Romania	0.052	0.013	0.037	0.019
Luxembourg	0.03	0.013	-0.001	0
Malta	-0.016	0.015	-0.015	-0.003
Netherlands	-0.033	0.031	0.012	-0.058 *
Poland	0.065	0.036	0.084	0.017
Czech Republic	-0.075 *	0.043	-0.026	0.038
Ireland	-0.054	0.05	0.002	-0.015
United Kingdom	-0.015	0.053 **	-0.076	-0.014
Slovenia	-0.062 *	0.057	-0.028	0.003
France	-0.027	0.074 **	-0.206 ***	0.023
Hungary	-0.144 ***	0.086 **	-0.037	0.064 *
Bulgaria	0.106 **	0.112 ***	0.016	0

With the conditional multilevel model, we can also explore the concrete intercept and slope for each country. These parameters are obtained in a second estimation called 'shrinkage' (see Goldstein, 2011) with the last model. They represent whether the country has a significantly different estimate than the common one. To explore the variations between countries, the slopes v_{1j} could be considered as the phenomenon of convergence in that country apart from the general process. On the other hand, the intercepts v_{0j} represent the steady state of that particular country when the remaining variables are 0. A representation of the slopes and intercepts in the extended GLMM model can be found in Figure 4.2 and Figure 4.3. Detailed information of the intercepts and slopes in all the variables can be found in Table 4.3, in the appendix.

Figure 4.2. Slopes by country in the Random Slope Model with spatial effects.

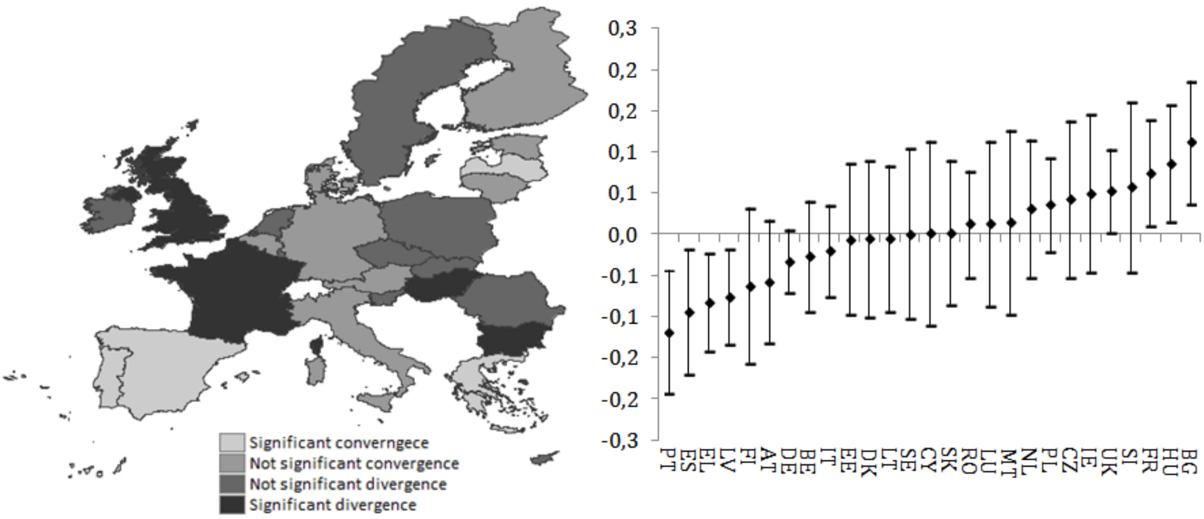
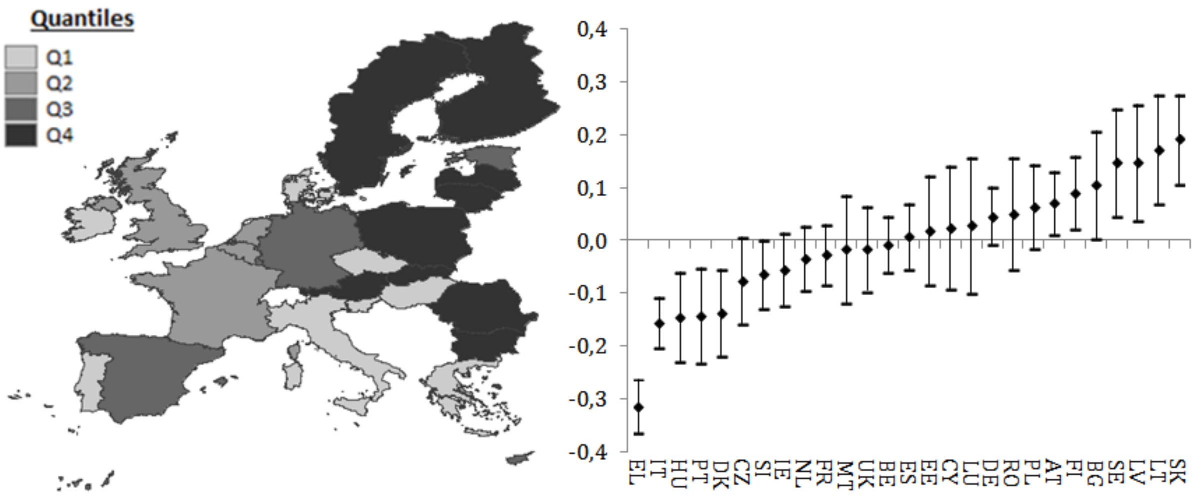


Figure 4.3. Caterpillar plot of the intercepts by country with spatial effects.



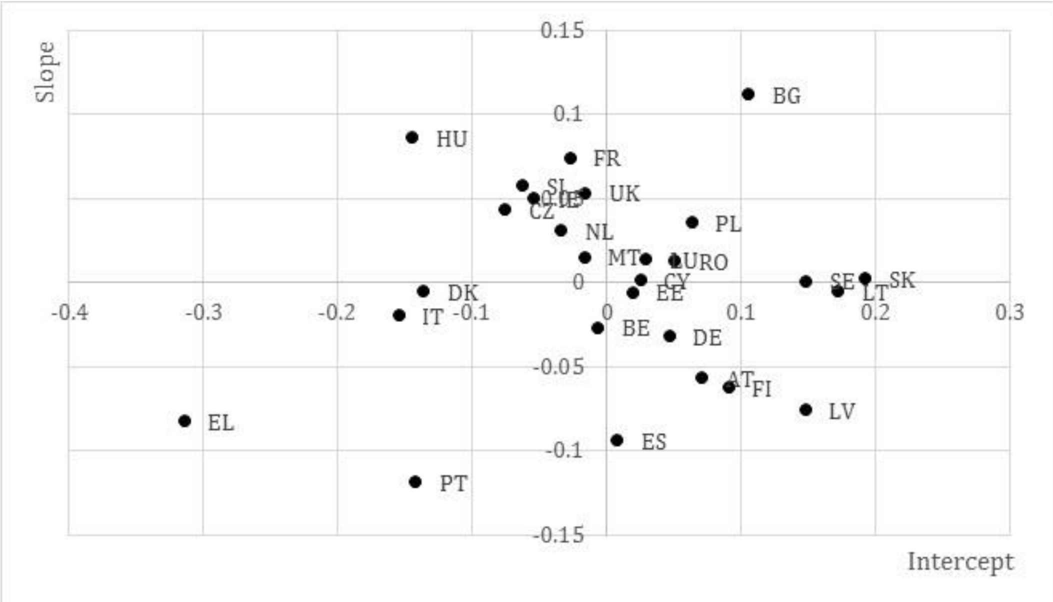
Using the country specific estimates, it seems that there are significant differences between the convergence models of the different countries of the European Union. For example, there are countries such as Bulgaria with a process of divergence, whereas others such as Spain, Greece and Portugal have a process of significant convergence. In addition, the caterpillar plot of the intercepts indicates the common characteristics of the territories of a group. Despite the differences in terms of the speed of convergence, it seems that the differences in the intercepts are much more important. This evidence could indicate that the lack of integration between the models of the different countries could lead the regions to actually different paths of growth.

The results reveal that the new members of the EU are defined by a higher rate of growth after you have taken into account the rest of the factors. The convergence rate of these countries are lower than in the rest of the EU. Artelaris et al. (2010) indicate that the communist period of these countries created an enormous homogeneity of wealth. So, they are catching up the rest of the EU in both, wealth but also spatial inequalities.

Country specific estimates of the human capital on the growth of the region seem to be homogeneous in the EU. It seems that Germany or France could have a lower impact of the human capital in one level of attainment, but it is compensated with a higher influence of the other category of attainment.

The variety of cases can be summarized in a scatter plot of the intercepts and slopes estimated for the different countries. Figure 4.4 illustrates the joined distribution of slopes and intercepts for the European scenario. This graph divides the scenarios in cases with a higher or lower process of convergence and the intercept – the potential of the economy. This figure highlights the special case of the countries in the east of the European Union such as Poland and Slovakia. These countries could be far from their potential steady state and may easily develop dynamics of polarization.

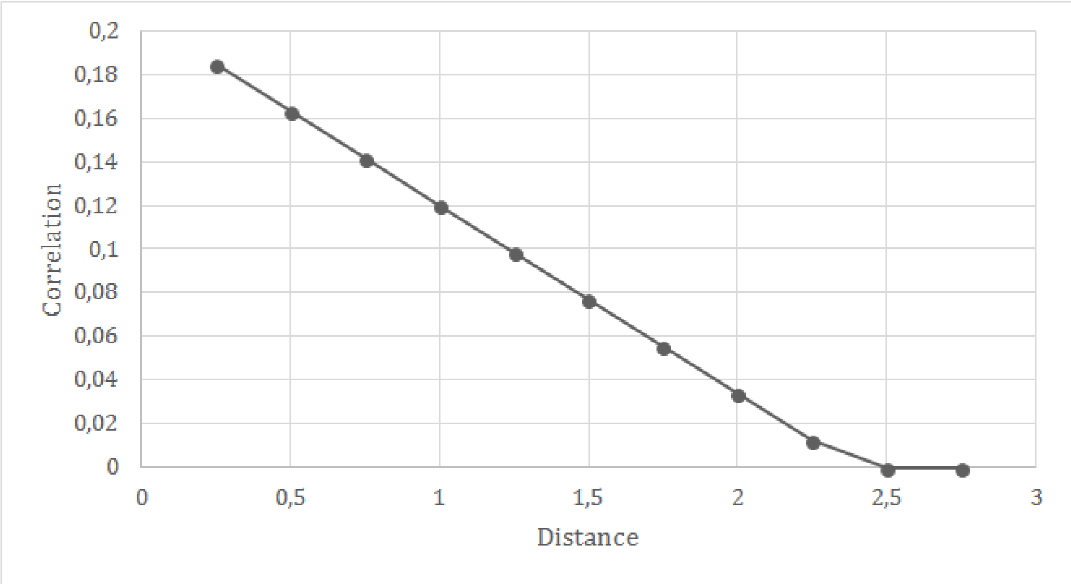
Figure 4.4. Scatter plot of the intercepts and the slopes.



Finally, the spatial interaction of the effects introduced into our model is significant, as expected. The coefficient called range establishes the maximum distance with spatial correlation. Thus, for a distance greater than 2.39 degrees, there no spatial interactions. The equivalent of this distance in kilometres depends on the location of the coordinates. As a reference, there are 162.5 kilometres measuring from Berlin to its west.

Our main interest in this chapter is the measurement of the importance of hierarchy. However, the inclusion of this spatial interaction had to be taken into account to obtain an accurate estimation of this weight. The estimation of these interactions can be used to calculate the function of the correlation, depending on the distance. This estimation indicates the types of spatial interactions that are found. Depending on the area of influence of the spatial interactions, the interaction can be inside each country or extend beyond the borders. Figure 4.5 illustrates the mathematical function of the correlation (see Littell et al., 2006).

Figure 4.5. Function of the correlation and the distance.



The function has a slope of -0.045. This estimation indicates that half of the correlation is lost after 81.3 kilometres and that there are no correlations after 162.5 kilometres. This profile indicates a phenomenon of spatial interactions located in a small area of influence. The area of influence in this analysis is not limited by the groups in the sample. So, the spatial interactions can cross through borders.

This result is expected for two reasons. Firstly, it is the effect of the spatial interactions after the hierarchy has been taken into account. As a result, any distant interaction is expected to be modelled by the heterogeneity of the hierarchy. Secondly, these are similar to the results found in the previous literature. For example, using a different model, Ramajo et al. (2008) and López-Bazo et al. (2004) also indicate that the spatial spillovers follow a national profile

IV.4. Conclusions

Under the Neoclassical theoretical framework, the regional division of data is not relevant because all spatial units should present the same pattern of decreasing returns. β-convergence analysis is consistently designed for the empirical testing of this theoretical framework. A great deal of attention has been paid to how the β parameter is technically estimated, but not to how the definition of spatial units is made. Subsequently, Economic Growth theories have substantially changed some assumptions. First, human capital, technological development and other factors were included in understanding the differences in development. Conditional convergence and panel data estimations aim to consider how these factors could affect the results. Endogenous Economic Growth models and the New Economic Geography focus on local growth processes. Under these new perspectives, spatial definition is very important. However, as seen in part II.4, spatial aggregation can generate significant problems of both, bias and inefficiency. These problems seem to appear, even in a simple Monte Carlo simulation.

Nevertheless, in the empirical framework, very different local patterns could be aggregated below the country, state or regional level. The only way to estimate these patterns is by working at the suitable scale with the process.

However, it seems too restrictive to think that all the generation process occurs at the disaggregated level. By means of multilevel analysis, it is possible to use different levels of spatial desegregation and observe how convergence changes at each of these levels. This is of major interest because it allows simultaneous and coherent analyses by region and local area. We apply different multilevel models to data from the Europe.

The multilevel methodology allows us to explore a larger picture of the complexity of the processes of convergence among countries and large regions. Using this type of estimation, several processes of convergence for each region have been calculated, taking into account the general process for the entire sample. As a result, the importance of the variance between and within region in the ß convergence phenomenon can be estimated.

The results of our empirical approach confirm that different behaviours can coexist. For the European case, an overall local convergence is confirmed, as found in previous studies. Using the multilevel approach, however, we observe that the internal rate of convergence in some states could be lower than the aggregate rate and could even show internal divergence. This is coherent with the New Economic Geography. Some areas, central areas, grow rapidly with a pattern of convergence and produce aggregate convergence of the state. However, others, i.e., the periphery, could present divergent behaviour.

To observe whether the differences in the models of convergence are caused by the fundamental factors explained in the neoclassical literature, the model was expanded with new information. However, even with this information, there remain important differences between the states, which could be historical, institutional, cultural or structural.

In this estimation, spatial interactions are also taken into account with a correlation between error terms that depends on the function between the territories. The estimation indicates that this type of interaction operates in close distances. This type of result could be related to different economic models, where the space is the key element in obtaining equilibrium since the NEG.

The contribution of this analysis is measuring convergence at global and local levels simultaneously to observe if different intra-regional patterns exist though hidden in the general trend.

Our study shows the importance of the spatial definition of the convergence study. Moreover, it could explain why general convergence could be found with a central-peripheral pattern occurring at a lower degree of spatial desegregation. From the political point of view, it is also worth stressing that regional policies should be designed at the local level, as peripheral local areas of rich regions could present processes of divergence as intense as those in poorer regions.

One of the key elements which could generate these processes of convergence or divergence could be the advantage in terms of productivity in the core of the region. Higher populated areas could generate advantages in terms of productivity caused by agglomeration economies which could create a core-periphery pattern.

Section 2

V. MEASURING THE AGGLOMERATION ECONOMIES

V.1. Urban agglomeration and economic growth

The process of convergence or divergence is directly linked with the core-periphery processes. In fact, the urbanization - defined as the concentration of the population in certain places: large metropolises and urban areas - is intrinsically linked to development and implies processes of strong concentration. At least since the mid-20th century, urban systems across nations and over time follow very stable general rules of concentration in large metropolises. These are situated at the top of the relative size distribution, above medium-sized and small cities. Despite the massive urbanization and technological changes that have taken place over the last 50 years, the relative size distribution of cities has remained 'rock stable': the relatively big stay big (Henderson, 2010). Even though the absolute number of cities in the top 5 percentile by size has grown over the last century in the US, Henderson (2003) shows that cities which were in that percentile 100 years ago are still there today. See also Eaton and Eckstein (1997) with respect to this point for the cases of Japan and France. Henderson and Wang (2007) demonstrate that the size distribution for US cities has remained almost identical over the last five decades.

What are the consequences of this strong concentration? How spatially concentrated should urbanization be? How much development should be located in megacities? What is better, promoting huge urban concentrations or spatial dispersion? Among other questions, these constitute highly relevant issues for present and future urban policy making. Answering these kinds of questions is decisive in the design of urban policies in all cases, but is especially important in developing countries. Many megacities are emerging in Asian, Latin American and sub-Saharan African states, giving rise to the most asymmetric urban systems in the world. Urban economic analysis should provide various types of evidence which aid policy makers when taking decisions and determining urban policies.

The main issue is to understand what is best in order to generate sustainable economic and social growth. According to Williamson (1965), agglomeration has positive effects in terms of GDP per capita at early stages of development, when transport and communication infrastructures are scarce and the reach of capital markets is limited. However, this author postulates that when infrastructures improve and markets expand, congestion externalities could become relevant, thereby making large concentration less efficient. Contrary to Williamson's hypothesis, the more widely accepted idea in the new theories of economic growth and geography is that spatial concentration and proximity are always good for economic growth. For example, among many other researchers, Martin and Ottaviano (1999) suggest that strong agglomerations and growth follow a self-reinforcing process, while Baldwin and Martin (2004) stress that spatial agglomeration is conducive to growth thanks to the spillovers and other positive effects of economic concentration. From the perspective of urban systems theory, however, a minimum degree of urban concentration is considered necessary and very positive for growth. However, policies should focus on avoiding too much concentration and reducing congestion externalities (Henderson, 2010).

The estimation strategy followed by Ciccone (2002) may be considered as one of the most popular ways of measuring agglomerations economies effect on local/regional productivity by estimating the effect of employment density on the generation of spatial externalities. More specifically, in Ciccone (2002) a model on which average labor productivity in one area depends on labor density –defined as labor units by unit of land– is derived. The empirical estimations of this model for the cases of Germany, UK, France, Italy and Spain find that the elasticity of labor productivity with respect to employment density is within the range of 4.5 and 5 percent, under several specifications at the scale of NUTS-3 regions.[11]

One potential issue in the measurement of spatial externalities on productivity is the geographical scale at which the empirical estimation of agglomeration economies is conducted. Ciccone (2002) argues that NUTS-3 regions is an appropriate spatial scale, since their median size in the set of countries studied is 1,511 km^2, which is slightly smaller than the median size of U.S. counties. However, administrative NUTS-3 division can be considered a highly aggregated spatial scale for the case of some countries. This is the case of Spain, which is divided into 50 NUTS-3 regions (Spanish Provincias) with sizes ranging from less than 2,000 km^2 to more than

[11] NUTS is the acronym of Nomenclature of Territorial Units for Statistics in French.

20,000 km^2 –the median size is 9,998 km^2–. This relatively large size of the regional units can be hiding a potential heterogeneity within the regions that can be conditioning the results of empirical model: by assuming the same average productivity and density figures the potential intra-regional heterogeneity is neglected.

The literature on the empirical quantification of agglomeration economies in Spain is not vast. Alonso-Villar et al. (2004) measured agglomeration economies for the manufacturing industries between 1993 and 1999 at the level of NUTS-2 and NUTS-3 regions, finding significant inter-industry differences in the scope of agglomerations and a positive correlation between their size with the technological intensity of the industries. Martínez-Galarraga et al. (2008) studied the productivity of industrial labor in Spain during the period 1860-1999, basing on the same estimation strategy as in Ciccone (2002) and taking NUTS-3 regions as units of analysis. They found that the elasticity with respect to employment density ranged between slightly less than 2% to more than 8%, depending on the time period and the estimator applied. More recently Jofre-Monseny (2009) conducted a similar exercise but for the specific case of Catalonia –a NUTS-2 Spanish region– for the period 1995-2002, finding agglomeration elasticities ranging between insignificant to more than 7% depending on the specific branch of the manufacturing industry. Oppositely to the previously mentioned papers, they base on information highly disaggregated at the spatial level –microdata at the scale of establishments– from data not publically available on registers in the Spanish National Social Security Registry. Alañón-Pardo and Arauzo-Carod (2013) also study the agglomeration effect, but they focus on the effect over the locations decisions. Their analysis highlights the agglomeration effects, accessibility and the spatial interactions between municipalities in the locations decisions.

In this section we estimate importance of the agglomeration with two different methodologies. The first one, in the spirit of Ciccone (2002), explaining average labor productivity in one spatial unit on its employment density. The novelty of the research is that, instead of estimating our empirical model at the level of NUTS-2 or NUTS-3 regions, we base our analysis on more disaggregated spatial data. Specifically, we take Income-tax microdata compiled by the Spanish Fiscal Studies Institute (Instituto de Estudios Fiscales) for calculating average compensations by worker at the scale of Local Labor Market (LLM), as defined in Boix and Galletto

(2008). Our claim is that taking highly disaggregated geographical units allows for considering an appropriate spatial scale to measure agglomeration economies, since spatially aggregated data implies assuming a high level of intra-regional homogeneity.

The process of agglomeration economies could be considered one of the most important explanations of the heterogeneity in the concentration across the space. However, a measurement still misses to answer how can it affect to the evolution of the country. What we propose is to ponder the degree to which large metropolitan areas may influence aggregate fluctuations. Could most of these fluctuations be explained just by the behavior of major urban agglomerations on the map? Might we find some kind of granular behavior? In other words, could the idiosyncratic behavior of some places on the map, the largest cities, explain a significant fraction of the aggregate volatility of the entire economy? This idea – in contrast to the previous approach – opens the process of agglomeration economies to an influence wider than the city.

The seed of the idea lies in a paper by Gabaix (2011), in which this author studies how relevant the idiosyncratic behavior of the largest firms might be in the aggregate fluctuations of the economy as a whole. Gabaix postulated the so-called 'granular' hypothesis: under several conditions related to the size distribution of firms, the main 'grains' in the economy, the largest firms, might play a significant role in many economic fluctuations. We propose to extend this idea of 'granular' effects to geographical units, namely to cities. Thus, the question that this study adds to the analysis of urban concentration processes is that of determining to what degree spatial 'granular' behavior can be confirmed. It also analyses the effects of this behavior in terms of economic volatility. Our specific aim is thus to apply this alternative methodology to regional economic analysis. We aim to achieve this goal using Gabaix´s methodology using the US territories instead of firms as the example for this methodology. The database that we use in our analysis is the Personal Income information of US Counties provided by the Bureau of Economic Analysis (BEA). This database has information from 1969 to 2011, sufficient for a sound econometric analysis.

V.2. Empirical estimation of agglomeration economies: The Spanish case.

Ciccone (2002), basing on Ciccone and Hall (1996), proposes a model with spatial externalities on productivity caused by the economic density of the territory. The point of departure is the following equation explaining the production by unit of land in the geographical unit or sub-region s that belongs to a larger region c:

$$q = \Omega_s f(nH, k, Q_{sc}, A_{sc}) = \Omega_s((nH)^{\beta} k^{1-\beta})^{\alpha} \left(\frac{Q_{sc}}{A_{sc}}\right)^{\frac{\lambda-1}{\lambda}} \qquad (5.1)$$

where q stands for the output per unit of land, Ω_{sc} is an index of total factor productivity in the area, A_{sc} is the total surface, n denotes economic density, H is the average level of human capital of workers per unit of land, and k stands for the density of physical capital. Parameter α captures the returns of capital and labor, whereas β is a distribution parameter. The empirical specification of equation (5.1) assumes that spatial externalities are driven by the density of production in the area Q_{sc}/A_{sc}, where $\frac{\lambda-1}{\lambda}$ represents the elasticity of output per unit of land with respect to economic density. In this specification, there are spatial externalities when $\lambda > 1$.

Some transformations are required in order to have an estimable version of equation (5.1) Assuming that the distribution of labor and capital is uniform within each spatial unit s across c, aggregate production may be written as $Q_{sc} = A_{sc}q$. Defining N_{sc} and K_{sc} as the levels of employment and physical capital in s, respectively, and assuming that the demand function of capital follows the expression:

$$K_{sc} = \frac{\alpha(1-\beta)}{r_c} Q_{sc} \qquad (5.2)$$

where r_c stands for the price of capital that is assumed constant in every sub-region s within the large region c. Under this assumption, labor productivity (Q_{sc}/N_{sc}) is given by:

$$\frac{Q_{sc}}{N_{sc}} = \Lambda_c \Omega_{sc}^{\omega} H_{sc} \left(\frac{N_{sc} H_{sc}}{A_{sc}}\right)^{\theta} \qquad (5.3)$$

In equation (5.3) ω is a constant and Λ_c depends on the rental price of capital and is assumed common for all the geographical units within c. Moreover, θ is defined as:

$$\theta = \frac{\alpha\lambda - 1}{1 - \alpha\lambda(1 - \beta)} \tag{5.4}$$

Parameter θ measures the effect on labor productivity of the density of employment in sub-region s. The value of θ in equation (5.4) can be estimated from data on production, employment density and human capital by assuming that differences in Λ_c across large regions are captured by dummy variables at the level of these larger regions.

By taking logarithms in (5.3) and including dummy variables for large regions, the final equation to be estimated is:

$$\log\left(\frac{Q_{sc}}{N_{sc}}\right) = \text{Large region dummies} + \theta \log\left(\frac{N_{sc}}{A_{sc}}\right) + \gamma \log H_{sc} + u_{sc} \tag{5.5}$$

Equation (5.5) relates labor productivity (Q_{sc}/N_{sc}) to employment density (N_{sc}/A_{sc}) in the spatial unit s, controlling by the effect of the stock of human capital (H_{sc}) by means of parameter γ, dummies that account for differences in total factor productivity and rental prices of capital between large regions and a disturbance term u_{sc}.

V.3. Database: fiscal data for Local Labor Markets in Spain (2011)

The empirical work draws on data on employment density and indicators of human capital and labor productivity at a spatial scale more disaggregated than NUTS-3 regions. Estimating equations like (5.5) from data collected at the scale of NUTS-3 administrative regions can imply working at a too highly aggregated scale, since average indicators of labor productivity or employment density can be hiding large intra-regional heterogeneity. This could be an issue, especially for those NUTS-3 regions on which the largest Spanish cities are located. As an example, the NUTS-3 province of Madrid is divided into 179 municipalities. Data from the 2001 census on population density (employment figures at municipal scale were not published in the 2011 census) showed huge disparities on this variable: the average population density in the province was approximately 800 inhabitant per km^2, but at municipal level population densities ranged from less than 2 to more than 7,000.

With the purpose of avoiding this problem, the data to estimate (5.5) are taken from a database that allows a more detailed spatial disaggregation. In particular, we base on microdata at the individual level in a cross-sectional sample of income-taxpayers published on a yearly basis by the Spanish Fiscal Studies Institute (Instituto de Estudios Fiscales), an institution dependent on the Spanish Ministry of Economy. This database provides information on wages reported on their income-tax declarations by the sample of individuals. The micro-data released from the sample in 2011 of approximately 549,000 individuals have been analyzed and taken as the main source of information for this study. One disadvantage of this specific database is that it does not provide data on variables as education level or years of tenure, for example, which could be useful when controlling for individual characteristics. On the other hand, it allows for deriving average indicators of labor productivity and employment density at a highly disaggregated spatial scale and covering all the population range. Since the model takes as unit of analysis spatial sub-regions s, this database is specially convenient for estimating models like (5.5) for Spain. Wage reported in the sample is taken as indicator of labor productivity, and average wage figures can be derived at the scale of Zip codes.[12] Similarly, it is possible to derive employment figures at the same level and then they can be aggregated at the desired spatial scale. Indicators of population or human capital, however, are not available at this same scale, which prevents the use of ZIP codes areas as the spatial unit of analysis. The most detailed spatial classification to estimate (5.5) is at the scale of municipalities, since the Housing and Population Census publishes information on the academic level of workers at municipal level.[13] Information on the surface of municipalities is available in the Housing and Population Census as well.

Even when our databases will allow us to take municipalities as the spatial sub-regions s in our model, we opted for aggregating these areas into larger spatial units for several reasons. One is the huge number of municipalities present in the Spanish spatial configuration –more than 8,000–. Consequently, for many of them the number of individuals sampled is too small to have reliable estimates of the variables

[12] Ciccone (2002) bases his study on data on value added, which can be considered a better indicator of labor productivity. However, this variable is not observable at the desired spatial scale and wages are taken instead. See Combes et al. (2011) or Melo et al. (2009) for examples of previous research when this approach is followed.

[13] The 2011 census has not released information on educational levels of workers for all the municipalities, which prevents using 2011 data. The census conducted in 2001, however, released this information and it will be the basis for recovering indicators of human capital in our estimations.

of interest. Secondly, the sample of income tax-payers provides information on their place of residence, not the place where these individuals work, and the labor density should be referred to the place where economic activity is located. For these reasons, ZIP codes are aggregated at the level of Local Labor Markets (LLMs).

LLMs are analytical areas resulting from aggregating municipalities among which the commuting flows are especially intense. A LLM is a group of municipalities designed to maximize flows of commuting intra-LLMs and, conversely, commuting flows between LLMS are minimized. The specific procedure for defining LLMs applied in this chapter corresponds to the definition given by the Italian Statistical Agency (ISTAT) and applied later by Boix and Galletto (2008) for Spain. This technique groups contiguous municipalities with the condition that at least 75% of people living within a LLM work there as well.[14] Consequently, the individuals in the sample are assigned to some of the 763 LLMs on which the Spanish territory is classified. The 763 areas do not cover all the Spanish territory because individuals paying their taxes in Basque Country and Navarra are not sampled, since these NUTS-2 regions have their own fiscal system, the so-called Haciendas Forales. The full set of variables, their definitions and sources and a summary of descriptive statistics is set out in Table 5.1.

Table 5.1. Variable definition, sources of information and descriptive statistics.

Variable	Definition	Source	Mean	Median	St. Dev
Q_{sc}/N_{sc}	Average wage (€/year)	Sample of income-taxpayers	14,529.08	14,153.31	3,499.68
N_{sc}/A_{sc}	Number of jobs by km^2	Sample of income-taxpayers; Housing and Population Census	36.3	13.32	85.06
H_{sc}	Percentage (%) of workers with college degree	Housing and Population Census	7.7	7.12	3.03

[14] Details on the specific algorithm used by these authors can be found in Boix and Galletto (2008).

V.4. Estimation Strategy

Equation (5.5) is estimated from information contained in the previously described databases. Average wages by LLM (Q_{sc}/N_{sc}) are derived from the sample of income-taxpayers; human capital (H_{sc}), defined as the fraction of workers with a college degree, is extracted from the 2001 Housing and Population Census; and the indicator of labor density (N_{sc}/A_{sc}) is derived by combining both statistical sources. Each LLM is assigned to one larger region c defined at the scale of the 15 NUTS-2 regions sampled in the survey.[15] This could be problematic if LLMs were formed by grouping municipalities belonging to different NUTS-2 regions. In practical terms, however, this is not an issue since there are only a limited number of cases –41 out of 763 LLMS– where this problem happens. It has been solved by assigning that specific LLM to the NUTS-2 region on which most of its population is located.

With these considerations in mind, the final equation to be estimated is:

$$log\left(\frac{Q_{sc}}{N_{sc}}\right) = \text{NUTS-2 } dummies + \theta\, log\left(\frac{N_{sc}}{A_{sc}}\right) + \gamma H_{sc} + u_{sc} \qquad (5.6)$$

In this chapter we describe several approaches to estimate parameter θ in (5.6). The simplest procedure consists of an ordinary least-squares (OLS) estimation from the 763 LLMs. However, an endogeneity problem caused by reverse causality can emerge if more productive LLMs attract more workers by unit of land (see Graham, 2006). Consequently, the OLS estimator of (5.6) would become inconsistent. Ciccone (2002) address endogeneity by adopting a Two Stage Least Squares (2SLS) estimator, where employment densities of the European regions analyzed are instrumented by their total land area. The argument for this choice is that land area is a variable historically predetermined and not conditioned by current productivities. This instrument would not be valid in our estimation, since LLMs are constructed grouping municipalities strongly interconnected by commuting flows, so

[15] The dummies variables are introduced at NUTS-2 level because this is the level in which the autonomous government (Spanish Autonomous Communities) works. At this level the Autonomous Communities have independency to carry out social, educational and health care policies among others. Provinces (NUTS-3 level) are in Spain an administrative division without political independence.

the total land area of one LLM is not exogenous but determined by the economic characteristics of the municipalities.

Alternatively, we follow the approach of Ciccone and Hall (1996), Rice et al. (2006), Graham and Kim (2008) or Artis et al. (2012), where current levels of density are instrumented by long lags of density. The justification is that modern densities are conditioned by past densities, being these not correlated with current productivities. Applying this approach to our problem requires data on historical densities at the spatial scale of LLMs. From the 1950 Spanish Housing and Population Census, which is the oldest one providing information on population densities at a municipal level, we recover the data necessary to define our instrument and (5.6) is estimated by 2SLS.[16] Additionally, a second set of instrumental variables that exploits weather differences throughout Spain has been considered as well, following the ideas presented in Combes and Gobillon (2015), which allows for performing formal exogeneity tests. In particular, we have taken as potential regressors on the first stage equations the Euclidean distance from the centroid of each LLM to the nearest point in the cost and the difference between the maximum and minimum temperatures –on average from 1987 to 2007-.[17] Both OLS and 2SLS approaches are estimated in their respective versions robust to heterokedasticity.

Besides OLS and 2SLS estimations, Quantile Regression (QR) estimations of (5.6) have been obtained as well. This estimation strategy has been previously applied in the context of quantifying agglomeration economies, as in Combes et al. (2009) or Briant (2010). Developed by Koenker and Bassett Jr (1978), the QR approach allows for estimating a coefficient at each conditional quantile of the dependent variable, not only at its conditional mean like in OLS and 2SLS estimators (see Koenker, 2005, for a more recent overview). The coefficient estimates by QR show the reaction at different points in the conditional distribution of the dependent variable to changes in the regressors. Estimating a modified version of (5.6) by QR will assess the conditional effect of labor density at several quantiles of the distribution of labor

[16] Data on labor density is not available at this geographical scale until the 1981 census. This forces us to take as instrument population density instead.
[17] Information on distances are taken from the website of the Spanish National Geographic Institute (www.ign.es), while data on temperatures come from the Spanish National Weather Agency (www.aemet.es).

productivity. Denoting as τ these quantiles, the QR equations to be estimated will be:

$$log\left(\frac{Q_{sc}}{N_{sc}}\right)_{\tau} = \text{NUTS-2 } dummies_{\tau} + \ \theta_{\tau}\, log\left(\frac{N_{sc}}{A_{sc}}\right) + \gamma_{\tau} H_{sc} + u_{sc_{\tau}} \qquad (5.7)$$

where the coefficients θ_{τ} and γ_{τ} measure respectively the effect to labor density and human capital at the τth quantile of labor productivity. QR estimates can be affected by the same endogeneity problems commented for the case of OLS. The solution to this issue lies in applying the Instrumental Variable Quantile Regression (IVQR) estimator developed by Chernozhukov and Hansen (2005 and 2006). The IVQR estimation of (5.7) is based on the same instruments –1950 population densities, distance to the coast and the range between maximum and minimum temperatures– for current labor densities as in the 2SLS estimation.

V.5. Main results

We have proceeded to estimate the models exactly as depicted in in equation (5.5), but also adding and removing some control variables.[18] First, we have applied the four estimation strategies described above to a version of (5.5) that does not include the effect of educational human capital. Table 5.2 and Table 5.3 summarize the results. Table 5.2 reports the OLS estimation of (5.6) in the first column, together with QR estimates of (5.7) at quantiles 10th, 25th, 50th, 75th and 90th in columns 2 to 6, while Table 5.3 shows the 2SLS in its first column and the IVQR estimates in columns 2 to 6.

[18] Ciccone (2002) extends its model –see page 219– by including externalities across regions derived from a spatial autoregressive process: higher productivities in neighboring regions could increase the own productivity in one area. The self-contained nature of the LLMs in our analysis excludes theoretically the presence of these spatial effects, since commuting flows between two different LLMs are close to zero. However, a Moran's test has been conducted to test for spatial autocorrelation in labor productivity, basing on a distance based and a binary contiguity matrix among LLMs. The respective Moran's-I statistics were -0.001 and 0.01, not rejecting in neither case the null hypothesis of absence of spatial autocorrelation.

Table 5.2. OLS and QR estimations.

		QR				
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	**0.10**	**0.25**	**0.50**	**0.75**	**0.90**
Labor density (θ)	0.093***	0.075***	0.080***	0.089***	0.097***	0.102***
Constant	5.264***	5.157***	5.157***	5.296***	5.330***	5.313***
N	763	763	763	763	763	763
R^2	.371					
Pseudo R^2		.217	.216	.237	.251	.223

Note: *, ** and *** represent estimates significantly different of zero at 10%, 5% and 1%, respectively.

Table 5.3. 2SLS and IVQR estimations.

		IVQR				
	(1)	(2)	(3)	(4)	(5)	(6)
	2SLS	**0.10**	**0.25**	**0.50**	**0.75**	**0.90**
Labor density (θ)	0.071***	0.062***	0.070***	0.065***	0.069***	0.073***
Constant	5.289***	5.076***	5.228***	5.336***	5.348***	5.450***
N	763	763	763	763	763	763
R^2	.367					
F(1,747)	613.5***					
Score χ^2 for overidentifying restrictions[(b)]	0.744 (p = 0.689)					

Notes: *, ** and *** represent estimates significantly different of zero at 10%, 5% and 1%, respectively. (a) F(1,747) represents the F-statistic for the first stage equation. The result is significant at 1%. (b) Wooldridge's score χ^2 for overidentifying restriction tests whether the instruments are uncorrelated with the error term. The value in parenthesis reports the p-value. The result for our specification is not significant at a 10% level, indicating that we should not reject the null hypothesis that our instruments are valid.

The estimate of θ in the mean is significantly different from zero, being the point estimate equal to 9.5 percent in the OLS estimation and 7.1 percent in the 2SLS version that uses the log of labor density in 1950, the distance to the coast and the range of temperatures as exogenous regressor for the first stage equations.[19]

[19] Tests for relevance of the instruments and for the no-correlation with the structural error term have been conducted in the 2SLS estimation. The auxiliary regression of employment density on the rest regressors –including those added to the first stage equations- is useful to test the quality of this set of instruments. The F-statistic in the first stage equation that explains the labor density on the rest of regressors and the mentioned additional regressors is significant at 1%. Moreover, the results of a Wooldridge's score test of overidentifying restriction were not significant at a 10% level, indicating that we should not reject the null

Estimates from QR and IVQR estimation in columns 2 to 6 show the impact of labor density along the conditional distribution of labor productivity. Estimates of θ_τ quantify the change in the conditional labor productivity quantile caused by a shift in LLM employment density. A noticeable result under both strategies of estimation is that the effect of employment density is estimated to be significantly positive at any of the quantiles reported. In this reduced formulation of equation (5.5), the effect of labor density on labor productivity ranges from 7.5 to 10.2 percent between the 1st and the 9th decile in the QR formulation and between 6.2 to 7.3 percent if the IVQR estimator is applied. The effect of changes in labor density on productivity is generally increasing along quantiles, indicating a pattern similar to the results found in other recent literature that measure agglomeration economies along the distribution of wages (see for example Briant, 2010; Matano and Naticchioni, 2015).

The estimation described above has been extended to control for educational human capital as originally formulated in (5.5). The estimates are reported in Table 5.4 and Table 5.5, where the OLS and QR estimates are shown in Table 5.4, while the 2SLS and the IVQR results are reported in Table 5.5.

Table 5.4. OLS and QR estimations.

		QR				
	(1) **OLS**	(2) **0.10**	(3) **0.25**	(4) **0.50**	(5) **0.75**	(6) **0.90**
Labor density (θ)	0.078***	0.094***	0.080***	0.078***	0.073***	0.068***
Human capital (γ)	0.438***	0.373***	0.452***	0.477***	0.540***	0.545***
Constant	5.264***	5.204***	5.214***	5.271***	5.301***	5.316***
N	763	763	763	763	763	763
R^2	.485					
Pseudo R^2		.266	.285	.321	.348	.319

Note: *, ** and *** represent estimates significantly different of zero at 10%, 5% and 1%, respectively.

hypothesis that our instruments are valid. All the estimations that are based on instrumental variables pass both test.

Table 5.5. 2SLS and IVQR estimations.

		IVQR				
	(1)	(2)	(3)	(4)	(5)	(6)
	2SLS	**0.10**	**0.25**	**0.50**	**0.75**	**0.90**
Labor density (θ)	0.054***	0.066***	0.047***	0.059***	0.047***	0.044***
Human capital (γ)	0.463***	0.367***	0.456***	0.516***	0.542***	0.617***
Constant	5.289***	5.142***	5.206***	5.287***	5.339***	5.355***
N	763	763	763	763	763	763
R^2	.476					
F(1,746)	639.9***					
Score χ^2 for overidentifying restrictions(b)	1.144 (p = 0.564)					

Notes: *, ** and *** represent estimates significantly different of zero at 10%, 5% and 1%, respectively. (a) F(1,746) represents the F-statistic for the first stage equation. The result is significant at 1%. (b) Wooldridge's score χ^2 for overidentifying restriction tests whether the instruments are uncorrelated with the error term. The value in parenthesis reports the p-value. The result for our specification is not significant at a 10% level, indicating that we should not reject the null hypothesis that our instruments are valid.

The effect of labor density on the mean of productivity is again estimated to be significantly different from zero, being the point estimate 7.8 percent when we apply an OLS estimator and 5.4 percent in the 2SLS version that uses the same instrument as before, which are similar to the estimates found by Artis et al. (2012) for the British counties in the period 2001-2005. However, the effect of labor density is now estimated to be decreasing along the conditional distribution of productivity, ranging from 9.4 percent in the 1st decile to less than 7 percent at the 9th decile in the QR estimation, and from 6.6 percent in the 1st decile to 4.4 percent at the 9th decile in the IVQR formulation. The contribution of human capital, in the other hand, is estimated as significant and increasing along the distribution of labor productivity for both QR and IVR estimators. This result can be interpreted as a signal that LLMs at the upper-end of the conditional distribution of labor productivity benefit less than those at the lower-end of the distribution from a shift in employment density, but they get comparatively higher growths on labor productivity as consequences of shifts in their educational human capital.

As a robustness test, we have repeated the estimation including as additional regressors indicators of industry specialization in order to account for industrial composition effects. In particular, we have taken from the 2001 Census data on employment for each LLM classified by industry, and we have calculated location quotients for the agricultural (LQ_{agr}), mining (LQ_{min}) and construction (LQ_{con}) sectors that are included as explanatory variables in (5.5).[20] The results of applying all the estimators previously described to this new specification are presented in Table 5.6 and Table 5.7.

Table 5.6. OLS and QR estimations with controls to industry specialization.

	(1)	QR (2)	(3)	(4)	(5)	(6)
	OLS	**0.10**	**0.25**	**0.50**	**0.75**	**0.90**
Labor density (θ)	0.058***	0.068***	0.056***	0.053***	0.062***	0.057***
Human capital (γ)	0.268***	0.233***	0.204***	0.271***	0.322***	0.381***
LQ_{agr}	-0.055***	-0.057***	-0.060***	-0.058***	-0.050***	-0.045***
LQ_{min}	-0.034**	-0.012	-0.046**	-0.048***	-0.055***	-0.013
LQ_{con}	-0.135***	-0.098***	-0.166***	-0.117***	-0.123***	-0.158***
Constant	5.337***	5.248***	5.283***	5.338***	5.377***	5.399***
N	763	763	763	763	763	763
R^2	.552					
Pseudo R^2		.324	.335	.366	.383	.357

Note: *, ** and *** represent estimates significantly different of zero at 10%, 5% and 1%, respectively.

Table 5.7. 2SLS and IVQR estimations with controls to industry specialization.

	(1)	IVQR (2)	(3)	(4)	(5)	(6)
	2SLS	**0.10**	**0.25**	**0.50**	**0.75**	**0.90**
Labor density (θ)	0.036***	0.037***	0.032***	0.036***	0.045***	0.031**
Human capital (γ)	0.273***	0.158**	0.214***	0.302***	0.378***	0.443***
LQ_{agr}	-0.059***	-0.063***	-0.064***	-0.063***	-0.050***	-0.043***
LQ_{min}	-0.029*	-0.027	-0.042*	-0.048**	-0.038*	0.015
LQ_{con}	-0.128***	-0.098	-0.157***	-0.131***	-0.107***	-0.132***
Constant	5.358***	5.132***	5.298***	5.313***	5.388***	5.486***
N	763	763	763	763	763	763
R^2	.546					
F(1,743)	634.4***					
Score χ^2 for overidentifying restrictions[(b)]	1.342 (p = 0.511)					

Notes: *, ** and *** represent estimates significantly different of zero at 10%, 5% and 1%, respectively. (a) F(1,743) represents the F-statistic for the first stage equation. The result is significant at 1%. (b) Wooldridge's score χ^2 for overidentifying restriction tests whether the instruments are uncorrelated with the error term. The value in parenthesis reports the p-value. The result for our specification is not

[20] We do not use data from the more recent 2011 census because it does not release information on employment by industry at the desired spatial scale, while the census conducted in 2001 does.

significant at a 10% level, indicating that we should not reject the null hypothesis that our instruments are valid.

The consequences of accounting for these industrial effects are not important in terms of general picture depicted, even when the size of the estimates of labor density is lower than in the other specifications. The contribution of the educational human capital is again increasing along the quantiles of productivity. Once we control for the potential industrial composition effects -measured as a concentration of employment in traditionally lower productivity sectors as agriculture, mining or construction- the productivity in the LLMs still benefit from higher labor densities. The estimates associated to these coefficients are, generally speaking, significantly negative at the mean and along the distribution of labor productivity for both QR and IVQR estimators. Under this specification, there is no evidence of an increasing effect of labor density along the quantiles, which was present in the most basic formulation of the model. This effect is estimated in the neighborhood of 6 percent at the mean and the different quantiles in the case of the OLS-QR estimators. The 2SLS formulation gets an estimate of 3.6 percent at the mean, while the IVQR estimator finds little variability around this number along the quantiles. The results under this extended specification of (5.5) are consistent with those in Table 5.4 and Table 5.5, not finding conclusive empirical evidence of higher effects of shifts in employment densities for the most productive LLMs.

V.6. Conclusions

In this analysis we have estimated two models to quantify the effect of the agglomerations in the economy. In particular, we follow the model developed in Ciccone (2002), but oppositely to previous empirical research that takes NUTS-2 or NUTS-3 regions as spatial units of analysis, we base our analysis on Local Labor Markets (LLMs). Additionally, we study to what extent the idiosyncratic behavior of some cities affect national aggregate fluctuations.

For the first model, we use microdata on wages reported in a sample of income taxpayers in 2009 that is disaggregated at the level of ZIP codes, together with information from the 2001 census, to calculate indicators of labor productivity, labor density and human capital at the desired spatial scale.

The estimable equation on which the empirical analysis bases is estimated by four different approaches. First, Ordinary Least Squares (OLS) and Quantile Regressions (QR) estimators are obtained for quantifying, respectively, the effect of employment density in the conditional mean and along the conditional quantiles of labor productivity. Additionally, in order to avoid endogeneity problems, Instrumental Variable (IV) versions of these estimators are applied as well. More specifically, a Two-Stages Least Squares (2SLS) and Instrumental Variable Quantile Regressions (IVQR) estimators are obtained by using as instrument of current densities population densities taken from the 1950 and geographic attributes.

Our empirical analysis finds a significantly positive effect of agglomeration in any of the approaches described. The effect of employment density in the mean is around 3 percent, with small differences found between the OLS and 2SLS estimators. The two quantile regression approaches, the ordinary QR and the IVQR estimator, show a similar pattern of the effect of employment density on the conditional distribution of labor productivity. Both estimators reveal a decreasing effect –but always significantly positive– of density along the conditional quantiles of labor productivity: QR and IVQR estimates of this effect at the 10th quantile are respectively 4.56 and 3.49 percent, whereas they are estimated in 2.44 and 2.83 percent respectively at the 90th quantile.

VI. THE GRANULAR HYPOTHESIS, A SPATIAL PERSPECTIVE

VI.1.The granular hypothesis

Gabaix (2011) proposes a simple origin for the volatility of aggregate fluctuations: as most economies are dominated by the largest firms, the idiosyncratic shocks of these firms can explain an important fraction of aggregate volatility. In his view, the main 'grains' in the economy, the largest firms, play a significant role in many economic fluctuations, and the so-called 'granular' hypothesis offers a micro-foundation for aggregate shocks.

The granular hypothesis is rooted in the size distribution of the units of analysis (firms in the original argument). In an economy with N identical firms with independent shocks, idiosyncratic movements vanish in the aggregate if N is a large number, as it is in modern economies. However, if the firm size distribution is sufficiently heavy-tailed, diversification effect may not be applied and the idiosyncratic shocks will not be cancelled out in the aggregate.

We can assume that the growth of the unit i is determined by:

$$g_{it} = \beta' X_{it} + \varepsilon_{it} \tag{6.1}$$

where g_{it} is the growth rate of unit i between t-1 and t, X_{it} is a vector of factors that may depend on unit characteristics at time t-1 and on factors at time t, and ε_{it} is the idiosyncratic shock. The granular residual is defined as the sum of the idiosyncratic shocks of the K largest units, weighted by size:

$$GR_t = \sum_{i=1}^{K} \frac{U_{i,t-1}}{Y_{t-1}} \varepsilon_{it} \tag{6.2}$$

where $U_{i,t-1}$ is the output of unit i in t-1 and Y_{t-1} is the total output in the same period. The idiosyncratic component, ε_{it}, would be extracted as $\hat{\varepsilon} = g_{it} - \hat{\beta} X_{it}$ after the estimation of (6.2) for the largest $Q \geq K$ units. However, the simplest specification is to control for the mean growth rate, $\overline{g}_t = Q^{-1}\sum_{i=1}^{Q} g_{i,t}$. Hence, the granular residual used in the empirical calculation will be:

$$GR_t = \sum_{i=1}^{K} \frac{U_{i,t-1}}{Y_{t-1}} (g_{it} - \overline{g}_t) \qquad (6.3)$$

In this configuration, the parameters K and Q have to be fixed, with Q ≥ K. Gabaix (2011) chose K=Q=100 firms, representing about one-third of the US GDP, for the baseline scenario.

After computing the granular residual for the K largest units, we are interested in knowing to what extent it can explain aggregate fluctuations, i.e., the growth in national output. So we regress the latter variable on the granular residual and an intercept. The main interest lies on the measure of fit of the proposed regression.

VI.2.Empirical application to US case

Our aim is to propose an extension of the general idea of the granular hypothesis from a geographical perspective using data from the US economy (US counties). In this case, counties with a large population, major cities, constitute the local 'grains' and the remaining medium-sized and small size cities and rural counties, the whole sample.

The Bureau of Economic Analysis (BEA) provides information on Personal Income (hereinafter, PI) in levels for every county throughout the US (3,138 counties). According to the BEA, PI is defined as 'the income received by, or on behalf of, all the residents of an area from all sources'. It should be borne in mind that each year is just one observation in our variables. As a result, the period of time studied in this research is the maximum available: from 1969 to 2011. Thus, one advantage of this database for the US economy is that it provides a number of years large enough to undertake this kind of research with confidence.

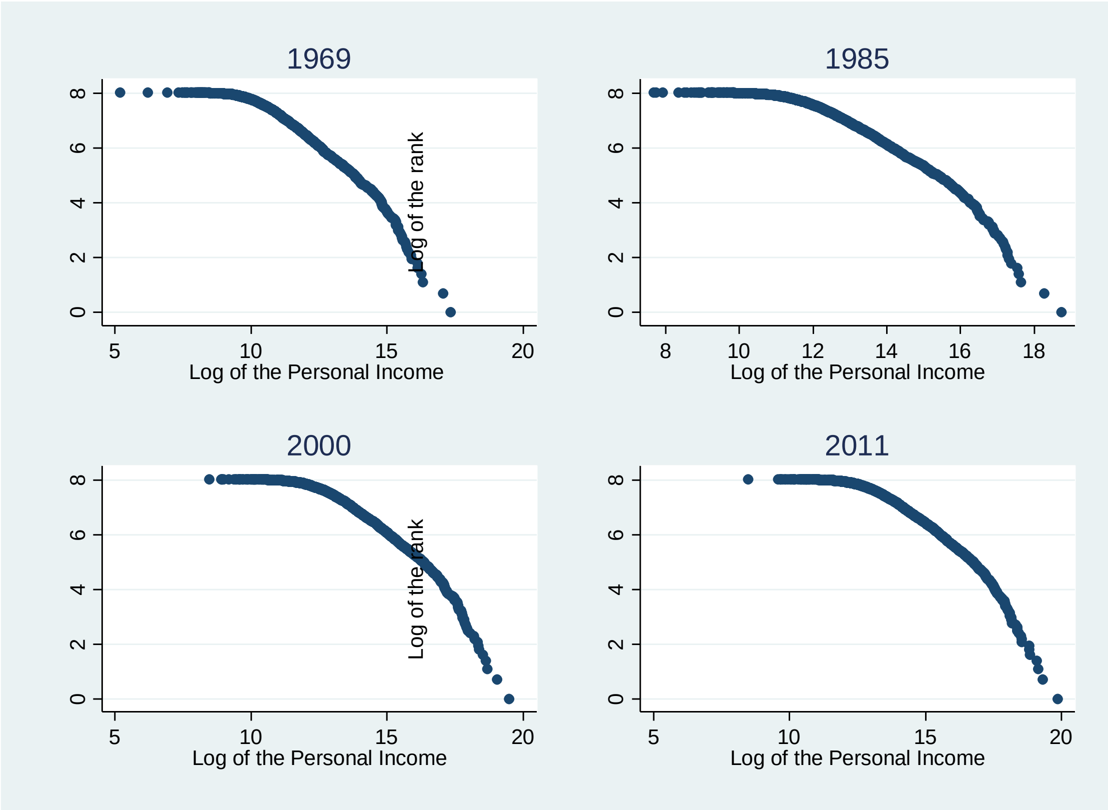
Information on Metropolitan Areas (MAs) for this period can also be found in the BEA database. At this point, a reasonable doubt might arise as to the use of MAs instead of counties seeing as MAs are the units generally used in regional analysis. However, counties are the natural 'grains' to test the granular hypothesis because their aggregation in MAs could dilute the behavior we wish to test. Moreover, the drawback of using MAs is that the population of units is small (366 in the BEA database), so only a few of them will represent a significant share of the total (the top 5 MAs alone represent between 25% and 30% of US Personal Income). Hence, the ideal small unit for this research would be the core of MAs, which could be captured by means of counties. Although this division raises problems even for this research study (in a few cases, for example, the core of a city could be divided in two counties), it could be the closest division to the idea of the `granular hypothesis´ for economic and methodological reasons.

As already stated, the starting point of Gabaix's granular hypothesis is a collection of units whose rank-size distribution obeys a power law. In his argumentation, the author focuses on the benchmark case of Zipf's Law when the exponent is equal to 1[21]. As is well known, there is abundant literature, especially for the US, showing that city size distribution obeys a power law, and, in many cases, that it follows Zipf's distribution. Numerous authors have verified this empirical regularity using population to measure city size. In the recent literature, Krugman (1996) and Gabaix (1999) himself estimate a Pareto exponent approximately equal to 1. Gabaix and Loannides (2004) extend these estimations to several countries. The literature is also reviewed in Eeckhout (2004) and González-Val (2010), who confirm that Zipf's Law holds depending on the truncation point (the number of the largest units of the subsample), a general result in empirical estimations of the power law exponent of size distributions in economics and finance (Gabaix, 2009).

In this analysis we shall use US PI by counties instead of population, so it is necessary to know whether PI by counties has a fat-tailed distribution, and in particular whether the size distribution obeys a power law. Figure 6.1 shows the scatter plot of ln(rank) versus ln(PI) for four periods: 1969, 1985, 2000 and 2011. The other years present a similar pattern.

[21]As Gabaix (2011, p. 744) remarks, it should be noted that the arguments do not depend on this assumption. He uses the case of Zipf's Law because it is empirically relevant and theoretically appealing.

Figure 6.1. Rank-size plots for select years.



After running a log-rank, log-size regression[22] for every year in the sample period, we find that the PI of US counties can be power-law distributed if the truncation point is fixed around 5% of the counties. In this case, the exponent is close to 1.5. However, the estimations depend on the truncation point. If we take the top 5% of the distribution, the estimated exponent is 1.57 (averaged over 1969-2011); if we take the top 10%, it is 1.23. These results are in line with those analyzed in the aforementioned papers by Eeckhout (2004) and González-Val (2010). Although the granular hypothesis does not depend on a Zipf distribution, we want to know whether this can be a sufficient outcome to expect granular behavior in the data.

Figure 6.2 as well as Figure 6.3 depict other features of the distribution of US PI by counties. Figure 6.2 reports the sum of the PI of the top 25 and 50 counties as a fraction of US PI. The share of the top 25 counties (less than 1% of the total number

[22] We run the following OLS regression: $\ln(R_i - ½) = \text{constant} - \beta \ln PI_i + \varepsilon_i$, where R_i is the rank of county i, PI_i is its Personal Income, and ε_i is white noise. We use the optimal shift ½ proposed by Gabaix and Ibragimov (2011).

of counties) is about 25% of US PI, whereas the top 50 counties comprise 35%. However, as both shares are decreasing over time, income is less concentrated in the top counties. The same pattern can be observed with the Herfindahl Index shown in Figure 6.3.

Figure 6.2. Shares of the top 25 and 50 counties in US Personal Income (1969-2011).

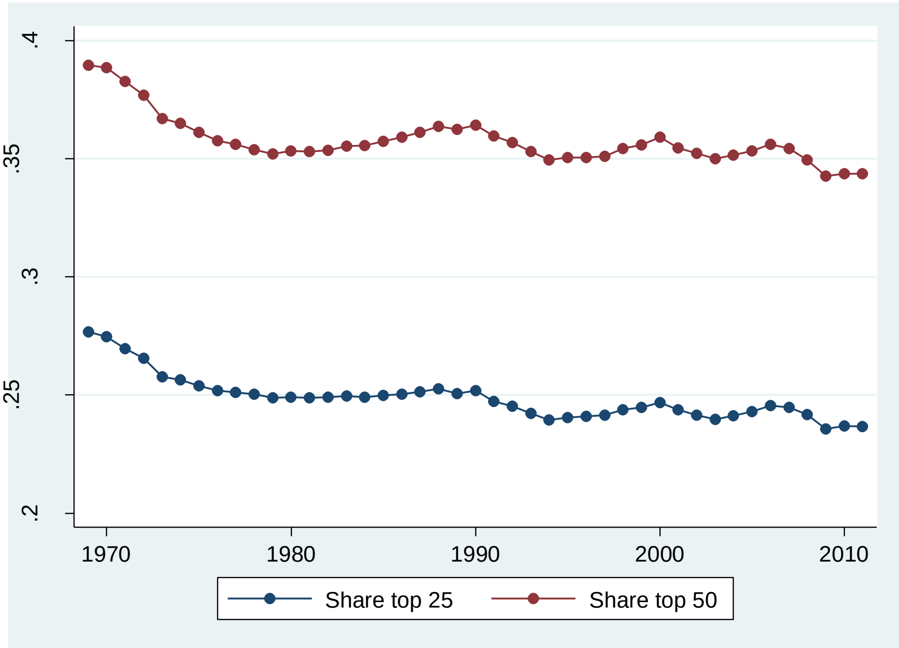
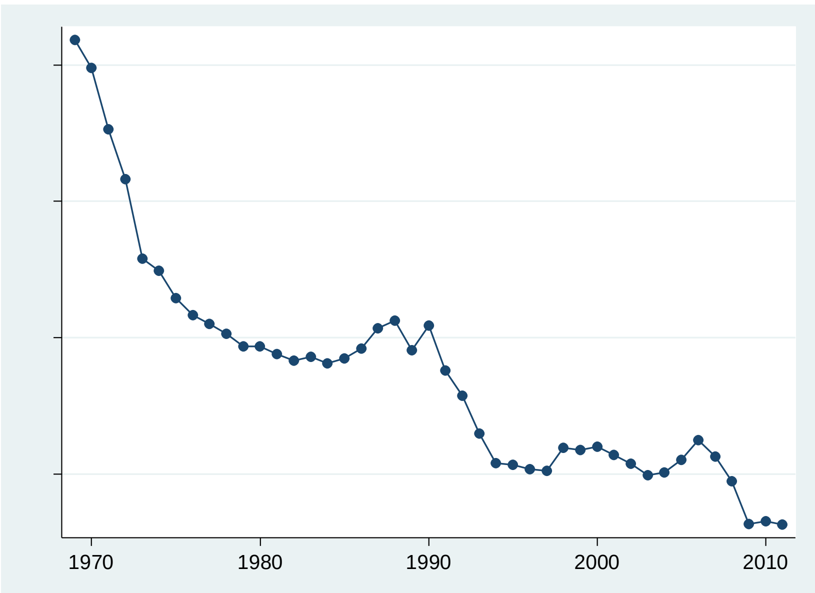


Figure 6.3. Herfindahl index of US Personal Income by counties (1969-2011).



Taking into account the granular hypothesis and the observed evolution of the US urban system, we expect to find some kind of granular behavior. The aim of the following part is to estimate this possible granular effect.

VI.3.Main results

The methodology outlined previously is applied to the data obtained from the BEA, the distribution of PI by counties for the 1969-2011 sample period.

In the first step, we compute the granular residual of PI by counties. We have to choose the parameters K and Q in order to compute the granular residual of the top K counties with the demeaning based on averaging over the top Q counties. The final selection is K=35 (a number of counties that concentrates 29.8% of the US PI, as an average for the 1969-2011 sample period) and K=50 (35.7% of the US PI in the 1969-2011 average). For the parameter Q, we choose Q=K or Q=2K, and for the case of K=35, we also test the largest value, Q=10K=350.

In the second step, we have to regress the growth of US PI on the granular residual. Due to the nature of the variables involved (a growth rate as the regressand and an average of demeaned growth rates as the regressor), a complex dynamic specification was not expected to be necessary, although a lag of the granular residual was be included. After running the OLS regressions, however we found that the residuals are serially correlated. The Prais-Winsten estimator (Prais and Winsten, 1954) is used instead, and the results are shown in Table 6.1

Table 6.1. Explanatory Power of the Granular Residual (Personal Income by Counties). Prais-Winsten Estimator Results.

	K=Q=35		K=35; Q=70		K=Q=50		K=50; Q=100		K=35; Q=350	
GRt	-0.0973	5.844	11.24***	10.79***	7.586	10.75**	8.659***	7.542***	5.168***	4.543***
GRt-1	-11.13*		1.093		-5.564		2.796		0.85	
_cons	0.0628***	0.0687***	0.0750***	0.0741***	0.0673***	0.0716***	0.0765***	0.0726***	0.0782***	0.0739***
R-sq	0.147	0.071	0.339	0.337	0.197	0.176	0.284	0.249	0.336	0.306
adj. R-sq	0.102	0.048	0.304	0.32	0.155	0.155	0.246	0.23	0.301	0.289
ρ	0.651	0.64	0.756	0.759	0.676	0.696	0.743	0.718	0.795	0.771
Durbin-Watson	0.703	0.78	0.615	0.627	0.684	0.736	0.656	0.656	0.683	0.699
Durbin-Watson	2.044	2.148	2.269	2.272	2.085	2.171	2.291	2.257	2.263	2.215
N	41	42	41	42	41	42	41	42	41	42

Note. Standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

As can be seen, the granular residual when Q≥K is statistically significant and the adjusted R^2s are relatively high, reaching a peak at 30.4% when K=35 and Q=70. The results are similar with K=35 and Q=10K (with a low β), and somewhat lower with K=50 and Q=100. The explanatory power decreases when K=Q. This outcome is not surprising, given that Q has to be equal to or greater than K, and indicates that a large Q has to be used for the demeaning in order to extract the component of the growth of the top counties. However, the overall impression is that the granular residual has explanatory power over the growth of US PI, so the results may support the granular hypothesis, in the sense that the idiosyncratic shocks of the top counties can explain a significant fraction (up to 30%) of the volatility of the US PI.

VI.4. Conclusions

Taking into account the important concentration generated by the agglomeration economies, we tried to understand how it affect to the whole economy. In the US economy Gabaix (2011) postulated and tested the 'granular hypothesis', measuring the relation between concentration in firms and evolution of the national economy. We extend the idea of this 'granular' effect to cities using the Personal Income distribution among US counties provided by BEA databases. After verifying that Personal Income by counties is power-law distributed, we calculate the granular residual and estimate which part of aggregate fluctuation is explained by this residual. The overall results show that our study may provide support for the granular behavior of US counties in the sense that the idiosyncratic shocks to the top counties can explain a significant fraction (up to 30%) of the volatility of US Personal Income.

Urban concentration is both a consequence and cause of economic development. There is a strong relationship between urbanization and growth, or vice versa, that has been widely discussed in the literature. These conclusions simply contribute a very specific point. Our evidence provides support for the hypothesis that this concentration also influences aggregate fluctuations in the sense that the more concentrated the economy, the greater the influence the major metropolises will have on aggregate volatility.

This idea of a ‘spatial granular hypothesis’ confirms the relevance of studying the differences between the economic structure and cycles of major metropolises compared to the remaining medium-sized and small cities. Henderson (1988) and Ellison and Glaeser (1997) show that small to medium-sized cities in the US are highly specialized, fundamentally in industrial sectors, while the large metropolitan areas are diversified and show a higher presence of culture and creative industries, R&D and global services. These differences tested in previous studies add further relevance to our conclusion. If we confirm some type of asymmetric structure between large cities and the rest, in addition to recognizing the effect of major metropolises on aggregate fluctuations, we are introducing new insights into the causes of territorial unbalance that should be addressed.

VII. SUMMARY AND FINAL REMARKS

Empirical regional economic analysis should be aware about the relevance of the spatial scale and spatial unit of investigation. The level of aggregation should not be considered as a minor issue. If the level of analysis is not consistent with the theoretical framework and assumptions the conclusions could be meaningless. This is the central hypothesis of this thesis. The main aim of this work was to evaluate how relevant was the scale in regional economic analysis as well as to propose new approaches and methodologies that could take advantage of the increasing information at highly disaggregated level and recent improvements in statistical tools.

The spatial unit of investigation and the scale of analysis is far from being a minor decision limited by the availability of data. The level of disaggregation should be motivated by the research question itself and be consistent with the theoretical framework and its assumptions. Regarding to this, it is important to observe that the differences of the relevance of the spatial scale are in the core of the main theoretical frameworks. Neoclassical Economics models are based on the key assumption of decreasing returns. This assumption operates no matter the scale, so, conclusions at an aggregated scale can be directly applied to lower scales. However, most of the Urban Economics models or the New Economic Geography framework operate in a local level, therefore an analysis made at regional or national level could hide the expected different behaviors among central and peripheral places. Under these approaches the scale became a crucial component. The assumption of homogeneity within groups generates problems capturing the true value of the coefficient and the miss-specification of the scale could lead to a measurement of a different concept.

To explore and evaluate the relevance of the spatial scale in the conclusions of empirical analysis we focus our attention in two main fields of the Regional Economics: (i) the analysis of the evolution of the economic differences among territories by means of the widely extended Beta-Convergence studies and (ii) the studies of productivity and growth and the particular effect of agglomeration economies or accumulative processes.

Theoretical analysis–shows to which extent MAUP has an effect on ß-convergence equations. Firstly, we show how aggregation of spatial data can generate a problem of bias in the OLS estimator of β-convergence equations from cross-sectional data, as well as inflating its variance. Second, by means of a numerical simulation, we quantify the effect of geographical aggregation on the estimates of β-convergence. Our experiment is based on real spatial structures of aggregated and disaggregated data for different countries and it numerically illustrates how a modification in the spatial scale has a significant effect on this type of studies.

Advantages of local estimation can be seen in a more complex model presented for the Mexican case. We propose a conditional spatial β-convergence model that uses as regressor the distance to the U.S. border. This model is applied to the period from 1980 to 2010 using data at the local level (by municipalities). Unlike previous papers, working with municipal-level data allows us to more clearly observe convergence patterns across space and identify the effects of location. The extension of the time period considered makes possible to distinguish between before and after the NAFTA agreement. Results show that regions near the U.S. border grew faster than those further away. In addition, the rate of convergence near the U.S. border is significantly higher than in the rest of the country.

Our main results highlight that the conclusions obtained with aggregated data are not always valid at the local level. A high spatial level of aggregation could hide relevant different behaviors that could be happening inside of heterogeneous regions. We propose the multilevel approach to measure the relevance of these differences. This technique allows us to measure the importance of the different scales simultaneously. In addition, different models are shown in order to incorporate other variables and spatial interactions. European Union is used as an empirical example of the possibilities of this methodology in regional convergence. The results show clearly that the convergence process is driven by forces operating in different levels. They also indicate that there are also processes of convergence or divergence within the states after taking into account the general process for the whole sample. This analysis shows how the concept of multilevel convergence can be applied to a scenario in order to observe whether there are processes operating at different scales at the same time. In the specific case of the EU, it seems reasonable to think that different scales with operative government figures would have a significant role in the process of convergence. In general, the application of multilevel approach to beta-convergence studies shows the importance of the different spatial scales in the process of convergence. It appears as though the differences between countries avoid a unique movement. Therefore, the regions are also influenced by the movements within the country. These important differences between countries could be a relevant source of asymmetric shocks. This result could indicate that the process of integration is far from over. In addition, many countries have no evidence of significant convergence. This outcome directly contradicts the main objective of the regional policy of the European Union. It highlights the idea of the necessity of cooperation between administrations of each level in order to generate a homogeneous process of convergence. Otherwise, processes of divergence could emerge within the countries despite the regional policy decisions made at a supra-national level. Other similar analysis, not finally included in this thesis, were done to the cases of U.S. and Brazil obtaining similar general conclusions and relevant policy implications for each one of this cases.

The second part of the research tries to understand the reasons that could be behind the significant importance of the local level. One of the most important processes which may be identified as idiosyncratic of the local level is the agglomeration economies. This type of analysis focuses on the advantages in terms of productivity created in the cities that generates a network or cities and urban areas. The last chapter explores this process trying to use the most suitable geographical scale. The combination of the two data sources described in the thesis allows to estimate the model explained in Ciccone (2002) at a highly disaggregated geographical scale, which implies theoretical and empirical advantages. Empirically, working at a more disaggregated spatial scale -when the data required is available- increases the number of units of analysis and allows studying differential responses to shifts in density along the distribution of labor productivity, being this analysis nearly impossible at an aggregated level for a country like Spain. In addition, this is the most suitable scale with the theoretical specification of agglomeration economies.

The study has been made for the specific case of Spain. This type of analysis has not been conducted before due to the lack of available data at the city level in this country. However, productivity at this level was obtained through microdata of the income-taxes. The estimation indicates a positive and significant effect of the agglomeration as well as an important heterogeneity along the distribution of the productivity.

As a consequence, it could not be denied the existence of agglomeration economies at the local level. This result highlights the differences between rural and urban territories. In fact, they may be so dramatic that the entire economy could be linked to the evolution of a few influential cities. The methodology in Gabaix (1999) and Gabaix (2011) has been proposed in order to answer this type of queries. In the example of the US economy, it could be seen that a significant percentage of the GDP growth is explained by the characteristic evolution of the biggest cities in the country. 30% of the variation of the GDP could be explained by the idiosyncratic shocks of the top counties. This result is extremely interesting in order to understand the influence of the local level. Local processes are not only important in order to explain the evolution of the cities. The cities interact with the rest of the country and they can condition the entire economy.

As a summary, this dissertation indicates that movements at a disaggregated scale might be hidden when the data are aggregated at a large scale. In fact, it could be identified a significant importance of the different scales of the hierarchy. In addition, this research identifies significant processes of agglomeration at a local scale of economic areas. These conclusions indicate several considerations in terms of policy implications.

The first conclusion that we can extract from our work is the relevance of having more information at the local level in order to choose the proper unit of analysis in each particular empirical study. A few decades ago, this problem was not possible to solve due to the lack of data, but the new technologies are reducing the accessibility cost to information. As a result, this variable would not be so important in the future in order to process data at a disaggregated scale. Nowadays, the main problem with this requisite is the confidentiality of the information. Statistical institutes will have to face this type of problem in order improve future empirical analysis in the field of regional economics. Availability of enormous sources of information seems to be, in fact, growing over the time, so it is becoming easier for the researches to build their own indicators from individual databases.

Secondly, regional policy tries to promote poorer regions in order to reduce the income gap between territories. In fact, this policy is extremely important in well-known economies like the European Union –44.9% of the total budget–. But it only measures problems of inequality between aggregated regions. So, regional policies of the European Union could be boosting the core of the poor regions in expense of the rural and poor areas. In fact, this research indicates that this type of processes would be extremely difficult to observe and control.

Thirdly, disaggregation of the analysis could improve decision making in different steps of the process. Evaluation of the inequalities would be able to introduce the differences with-in the regions as an important variable of the regional policy. As a result, policy decision would not only be based on the GDP per capita of the region but also on its internal inequality. In addition, local evaluation of territories would allow to detect crisis of local communities that may need assistance of the regional policy but are located outside objective regions.

Fourthly, distribution of the resources at the local level may have its advantages. Promotion of local projects may be much more efficient because local governments may have a better knowledge about the necessities of that territory. However, they should not be treated as independent entities. Spatial interactions in the regions should be taken into account. According to NEG, territories tend to interact in a core-periphery mechanism, so, a group linked territories should be treated as a whole in efficient and modern regional policy. For example, promotion of environmental projects becomes useless when the contiguous municipality is doing opposite policies. Nowadays, the territories are linked and their policies can affect the surrounding neighbors. So, they should try to have a global and structured project through the regional policy.

And finally, evaluation and control of the regional policy could become more precise through an analysis with local data. This research proposes the multilevel technique as a suitable methodology to isolate the influence of the different levels of the hierarchy. This advantage of the methodology becomes clear in scenarios with multiple government categories – e.g. the European Union. The estimated results could be interpreted as the behavior of the territories apart from the general movements in the European Union.

Several future research lines can be proposed after our analysis. It focuses on the results created by a single aggregation, but there is no analysis about the consequences that different aggregations would create on the results. This type of study would focus on the other component of the MAUP, the zoning effect. The consequences of spatial aggregation in the context of estimators applied to dynamic panels are an important issue that should be included in the research agenda on the estimation of β-convergence and agglomeration equations.

In addition, convergence is not the only field of regional economics where the hierarchy might have a significant effect. Other phenomenon can be measure using this methodology with an appropriate specification to the special problem. So, it is expected that future regional research would try to investigate the importance of the structure through this technique. There are also relevant issues not studied here that would require further research. This thesis points to the necessity of using the correct definition of region for each problem in regional economics. A new framework of research would emerge when the national institutes allow a higher disaggregation of the available data. Improved and adapted classifications of the territory would be more common in future empirical analysis due to new information at local levels. For example, the European Union has no homogeneous databases at a local level, which avoids any comparative analysis across the national borders of regional mechanics.

Conclusiones en español

Los análisis de economía regional empírica deberían tener en cuenta la importancia de la escala y la unidad espacial de investigación. Si el nivel del análisis no es consistente con el marco teórico y los supuestos, las conclusiones pueden carecer de significado. El objetivo principal de este trabajo ha sido evaluar lo relevante que es la escala en el análisis económico regional, así como proponer nuevos enfoques y metodologías que puedan aprovechar la creciente información a un nivel desagregado y las recientes mejoras en las herramientas estadísticas.

La unidad espacial en la investigación y la escala del análisis está lejos de ser una decisión menor, limitada por la disponibilidad de los datos. El nivel de desagregación debe estar motivado por la pregunta a investigar en sí misma y debe ser consistente con el marco teórico y los supuestos. En relación con esto, es importante observar que las diferencias de la importancia de la escala espacial están en el centro de los principales marcos teóricos. Los modelos de Economía Neoclásica están basados en el supuesto clave de rendimientos decrecientes. Este supuesto opera sin importar la escala, así que las conclusiones a un nivel agregado pueden aplicarse directamente a niveles inferiores. Sin embargo, la mayoría de los modelos de Economía Urbana o en el marco de la Nueva Geografía Económica operan a nivel local, por lo que el análisis hecho a escala regional o nacional puede esconder comportamientos distintos en zonas centrales y periféricas. Bajo estos enfoques la escala se ha convertido en un componente crucial. Los supuestos de homogeneidad entre los grupos generan problemas capturando el valor real de los coeficientes y una elección incorrecta de la escala podría llevar a la medida de un concepto diferente.

Para explorar y evaluar la importancia de la escala espacial en las conclusiones de los análisis empíricos centramos la atención en dos campos principales de la Economía Regional: (i) *El análisis de la evolución* de las diferencias económicas entre territorios por medio de extensos estudios de Beta-Convergencia y (ii) los estudios de productividad y crecimiento y en particular, el efecto de las economías de aglomeración o los procesos acumulativos.

El análisis teórico muestra hasta qué punto el MAUP ha tenido un efecto en las ecuaciones de ß-convergencia. Primero, mostramos cómo la agregación de datos espaciales puede generar un problema de sesgo en el estimador MCO de las ecuaciones de ß-convergencia a partir de datos de sección cruzada, así como aumentar su varianza. En segundo lugar, por medios de una simulación numérica, podemos cuantificar el efecto de la agregación geográfica en la estimación de ß-convergencia. Nuestro experimento se basa en estructuras espaciales reales de datos agregados y desagregados para diferentes países e ilustra numéricamente cómo una modificación en la escala espacial tiene un efecto significativo en este tipo de estudios.

Las ventajas de la estimación local pueden verse en un modelo más complejo presentado para el caso mexicano. Proponemos un modelo espacial de ß-convergencia condicional que usa como regresor la distancia a la frontera de Estados Unidos. Este modelo se aplica para el período de 1980 a 2010 con datos a nivel local (por municipios). A diferencia de estudios anteriores, trabajar con datos a nivel municipal nos permite observar más claramente los patrones de convergencia a través del espacio e identificar los efectos de la localización. La extensión del período de tiempo considerado hace posible distinguir entre antes y después del NAFTA. Los resultados muestran que las regiones cerca de la frontera con Estados Unidos crecieron más rápido que las más alejadas. Además, la ratio de convergencia cerca de la frontera es significativamente más alta que en el resto del país.

Nuestros principales resultados destacan que las conclusiones obtenidas con datos agregados no siempre son válidas a nivel local. Un alto nivel espacial de agregación puede esconder diferentes comportamientos relevantes que pueden estar sucediendo en regiones heterogéneas. Proponemos el enfoque multinivel para medir la importancia de esas diferencias. Esta técnica permite medir la importancia de las diferentes escalas simultáneamente. Además, se muestran diferentes modelos para incorporar otras variables e interacciones espaciales. La Unión Europea se usa como ejemplo empírico de las posibilidades de esta metodología en convergencia regional. Los resultados muestran claramente que el proceso de convergencia es conducido por fuerzas operando a diferentes niveles. También indican que hay procesos de convergencia o divergencia en los estados tras tener en cuenta el proceso general de la muestra completa. Este análisis muestra cómo el concepto de convergencia multinivel puede ser aplicado a un escenario para observar si hay un proceso operando a diferentes escalas al mismo tiempo. En el caso concreto de la UE, parece razonable pensar que las diferentes escalas con figuras gubernamentales operativas tienen un papel importante en el proceso de convergencia. En general, la aplicación del enfoque multinivel a los estudios de beta-convergencia muestran la importancia de las diferentes escalas espaciales en el proceso de convergencia. Parece como si las diferencias entre países evitan un movimiento único. En consecuencia, las regiones también se ven influenciadas por los movimientos dentro del país. Esas importantes diferencias entre países pueden ser una fuente relevante de shocks asimétricos. Este resultado puede indicar que el proceso de integración está lejos de acabar. Además, en muchos países no hay evidencia de convergencia significativa. Esto contradice directamente el objetivo principal de la política regional de la Unión Europea. Resalta la idea de que es necesaria la cooperación entre administraciones de cada nivel para generar un proceso homogéneo de convergencia. De otra manera, podrían aparecer procesos de divergencia entre países a pesar de las decisiones de política regional hechas a nivel supra-nacional. Otro análisis similar, pero no incluido finalmente en la tesis, ha sido hecho para los casos de Estados Unidos y Brasil, obteniendo conclusiones generales parecidas e importantes implicaciones políticas para cada uno de los casos.

La segunda parte de la investigación intenta comprender las razones que podrían estar detrás de la importancia significativa del nivel local. Uno de los procesos más importantes que puede identificarse como idiosincrático del nivel local son las economías de aglomeración. Este tipo de análisis se centra en las ventajas en términos de productividad creada en las ciudades que genera una red o ciudades y áreas urbanas. El último capítulo explora este proceso tratando de usar la escala geográfica más adecuada. La combinación de dos fuentes de datos descrita en la tesis permite estimar el modelo explicado en Ciccone (2002) a la escala geográfica más desagregada, lo que implica ventajas teóricas y empíricas. Empíricamente, trabajar con una escala espacial más desagregada -cuando los datos requeridos están disponibles- aumenta el número de unidades de análisis y permite estudiar diferentes respuestas a cambios en la densidad a lo largo de la distribución de la productividad del trabajo, siendo el análisis casi imposible de hacer a un nivel agregado para un país como España. Además, esta es la escala más adecuada con la especificación teórica de economías de aglomeración.

El estudio ha sido hecho para el caso concreto de España. Este tipo de análisis no se ha realizado antes debido a la falta de información disponible a nivel local para este país. Sin embargo, la productividad a este nivel se ha obtenido a través de microdatos de los ingresos fiscales. La estimación indica un efecto positivo y significativo de la aglomeración, así como una importante heterogeneidad a lo largo de la distribución de productividad.

Como consecuencia, no se puede negar la existencia de economías de aglomeración a nivel local. Este resultado resalta las diferencias entre los territorios rurales y urbanos. De hecho, pueden ser tan dramáticas que la economía entera podría estar vinculada a la evolución de unas pocas ciudades influyentes. Se ha propuesto la metodología de Gabaix (1999) y Gabaix (2011) para responder a este tipo de preguntas. Utilizando el caso de la economía estadounidense como ejemplo, se observó que un porcentaje significativo del crecimiento del PIB se explica por la evolución característica de las mayores ciudades del país. El 30% de la variación del PIB podría explicarse por los shocks idiosincrásicos de los principales condados. Este resultado es muy interesante para explicar la evolución de las ciudades. Las ciudades interactúan con el resto del país y pueden condicionar la economía.

En resumen, esta tesis indica que movimientos en una escala desagregada pueden estar ocultos cuando los datos están agregados a un mayor nivel. De hecho, pudo identificarse una importancia significativa de las diferentes escalas en la jerarquía. Además, esta investigación identifica procesos significativos de aglomeración a escala local de áreas económicas. Estas conclusiones tienen varias implicaciones en términos de políticas.

La primera conclusión que podemos extraer de nuestro trabajo es la importancia de tener más información a nivel local para escoger adecuadamente la unidad de análisis en cada estudio empírico concreto. Pocas décadas atrás, este problema no se podía solventar por la falta de información, pero las nuevas tecnologías están reduciendo el coste de accesibilidad a los datos. Como consecuencia, esta variable podría no ser tan importante en el futuro para procesar datos a una escala desagregada. Hoy en día el principal problema con este requisito es la confidencialidad de la información. Los institutos de estadística tendrán que afrontar este tipo de problemas para mejorar los análisis empíricos futuros en el campo de la economía regional. La disponibilidad de grandes fuentes de información parece ser, de hecho, creciente con el tiempo, así que se está volviendo más fácil para los investigadores construir sus propios indicadores desde bases de datos individuales.

En segundo lugar, la política regional intenta promover las regiones más pobres para reducir la brecha en ingresos entre los territorios. De hecho, esta política es extremadamente importante en economías como la Unión Europea -44,9% del presupuesto total-. Pero sólo mide problemas de disparidad entre regiones agregadas. Así que las políticas regionales de la Unión Europea podrían estar impulsando el centro de regiones pobres a expensas de las zonas rurales. De hecho, esta investigación muestra que este tipo de procesos podría ser muy difícil de observar y controlar.

En tercer lugar, la desagregación del análisis puede mejorar la toma de decisiones en diferentes etapas del proceso. La evaluación de las desigualdades podría ser capaz de introducir las diferencias dentro de las regiones como una variable relevante de la política regional. Como consecuencia, la decisión política podría estar basada no sólo en el PIB per cápita de la región sino también en sus desigualdades internas. Además, la evaluación local de territorios podría permitir la detección de crisis en comunidades locales que pueden necesitar la ayuda de la política regional pero que están ubicadas fuera de regiones objetivo.

En cuarto lugar, la distribución de los recursos a escala local puede tener ventajas. Promover proyectos locales podría ser mucho más eficiente porque los gobiernos locales podrían tener un mayor conocimiento de las necesidades de cada territorio. Sin embargo, no deben ser tratados como entidades independientes. Hay que tener en cuenta las interacciones espaciales en las regiones. Según la NGE, los territorios tienen de interactuar en un mecanismo de centro-periferia, así que un grupo conectado de territorios debería tratarse como un todo en políticas regionales eficientes y modernas. Por ejemplo, la promoción de proyectos medioambientales se vuelve inútil cuando los municipios contiguos están llevando a cabo políticas opuestas. Actualmente los territorios están conectados y sus políticas pueden afectar a los barrios del entorno. Así que deben tratar de tener un proyecto global y estructurado a través de la política regional.

Y finalmente, la evaluación y el control de la política regional puede volverse más precisa mediante el análisis con datos locales. Este estudio propone la técnica multinivel como una metodología adecuada para aislar la influencia de los diferentes niveles de jerarquía. Esta ventaja metodológica se vuelve más clara en escenarios con múltiples categorías gubernamentales -por ejemplo, la Unión Europea. Los resultados estimados podían interpretarse como el comportamiento de los territorios al margen de los movimientos generales en la Unión Europea.

Tras este análisis se pueden proponer múltiples líneas de investigación para el futuro. Se centra en los resultados creados por una única agregación, pero no hay un análisis sobre las consecuencias que diferentes agregaciones podrían crear sobre los resultados. Este tipo de estudio se centraría en otro componente del MAUP, el efecto zona. Las consecuencias de la agregación espacial en el contexto de estimadores aplicados a paneles dinámicos son un tema importante que debe incluirse en la agenda de investigación sobre la estimación de ecuaciones de β-convergencia y aglomeración.

Además, la convergencia no es el único capo de la economía regional donde la jerarquía puede tener un efecto significativo. Otros fenómenos pueden medirse utilizando esta metodología con una especificación adecuada al problema espacial. Así que es de esperar que el análisis regional en el futuro intente investigar la importancia de la estructura a través de esta técnica. Hay también temas importantes que no se estudian aquí, que requerirían mayor investigación. Esta tesis apunta a la necesidad de usar la definición correcta de región para cada problema en economía regional. Un nuevo marco de investigación surgirá cuando los institutos nacionales permitan una mayor desagregación de los datos disponibles. Las clasificaciones mejoradas y adaptadas del territorio serán más comunes en futuros análisis empíricos debido a la nueva información a nivel local. Por ejemplo, la Unión Europea no tiene bases de datos homogéneas a escala local, lo que impide cualquier análisis comparativo a través de las fronteras nacionales de mecánica regional.

VIII. REFERENCES

Alañón-Pardo, Á., Arauzo-Carod, J.M., (2013) Agglomeration, accessibility and industrial location: evidence from Spain. *Entrepreneurship & Regional*, 25(3–4), 135–173.

Alonso-Villar, O., Chamorro-Rivas, J.M., González-Cerdeira, X., (2004) Agglomeration economies in manufacturing industries: the case of Spain. *Applied Economics*, 36(18), 2103–2116.

Andersson, L., Hammarstedt, M., Hussain, S., (2013) Ethnic origin, local labour markets and self-employment in Sweden: a multilevel approach. *The Annals of Regional*, 50(3), 885–910.

Anselin, L., (1988) *Spatial econometrics: methods and models*, Springer.

Arbia, G., Petrarca, F., (2011) Effects of MAUP on spatial econometric models. *Letters in Spatial and Resource Sciences*, 4(3), 173–185.

Aroca, P., Bosch, M., Maloney, W.F., (2005) Spatial dimensions of trade liberalization and economic convergence: Mexico 1985-2002. *World Bank Economic Review*, 19(3), 345–378.

Artelaris, P., Kallioras, D., Petrakos, G., (2010) Regional inequalities and convergence clubs in the European Union new member-states. *Eastern Journal of European*.

Artis, M.J., Miguelez, E., Moreno, R., (2012) Agglomeration economies and regional intangible assets: an empirical investigation. *Journal of Economic Geography*, 12(6), 1167–1189.

Azzoni, C.R., (2001) Economic growth and regional income inequality in Brazil. *The Annals of Regional Science*, 35, 133–152.

Baldwin, R., Forslid, R., (2003) *Economic geography and public policy*, Princeton University Press, Princeton.

Baldwin, R., Martin, P., (2004) Agglomeration and regional growth. In J. V Henderson & J. F. Thisse (Eds.), *Handbook of regional and urban economics 4*, 2671–2711, Elsevier.

Ballas, D., Tranmer, M., (2012) Happy People or Happy Places? A Multilevel Modeling Approach to the Analysis of Happiness and Well-Being. *International Regional Science Review*, 35(1), 70–102.

Barro, R.J., Sala-I-Martin, X., (1992) Convergence. *Journal of Political Economy*, 100(2), 223–251.

Barro, R.J., Sala-I-Martin, X., Blanchard, O.J., Hall, R.E., (1991) Convergence across states and regions. *Brookings Papers on Economic Activity*, 1991(1), 107–182.

Baumol, W.J., (1986) Productivity growth, convergence, and welfare: what the long-run data show. *The American Economic Review*, 76, 1072–1085.

Beardsell, M., Henderson, V., (1999) Spatial evolution of the computer industry in the USA. *European Economic Review*, 43(2), 431–456.

Behrens, K., Thisse, J.F., (2007) Regional economics: A new economic geography perspective. *Regional Science and Urban Economics*, 37(4), 457–465.

Boix, R., Galletto, V., (2008) Marshallian industrial districts in Spain. *Scienze Regionali*, 7(3), 29–52.

Briant, A., (2010) *Marshall's scale economies: a quantile regression approach*, Paris School of Economics.

Canova, F., Marcet, A., (1995) The poor stay poor: Non-convergence across countries and regions. *London CEPR Discussion Papers*, 1265.

Carrion-i-Silvestre, J.L., German-Soto, V., (2009) Panel data stochastic

convergence analysis of the Mexican regions. *Empirical Economics*, 37(2), 303–327.
Castells, M., (1996) *The rise of the network society*, Blackwell, Malden.
Chernozhukov, V., Hansen, C., (2005) An IV model of quantile treatment effects. *Econometrica*, 73(1), 245–261.
Chernozhukov, V., Hansen, C., (2006) Instrumental quantile regression inference for structural and treatment effect models. *Journal of Econometrics*, 132(2), 491–525.
Christaller, W., (1933) *Die zentralen Orte in Süddeutschland: eine ökonomisch-geographische Untersuchung über die Gesetzmässigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen*, University Microfilms.
Ciccone, A., (2002) Agglomeration effects in Europe. *European Economic Review*, 46(2), 213–227.
Ciccone, A., Hall, R., (1996) Activity, Productivity and the Density of Economic. *American Economic Review*, 86(1), 54–70.
Cohen, P.N., (1998) Black concentration effects on black-white and gender inequality: Multilevel analysis for US metropolitan areas. *Social Forces*, 77(1), 207–229.
Combes, P.P., (2000) Economic structure and local growth: France, 1984–1993. *Journal of Urban Economics*, 47(3), 329–355.
Combes, P.P., Duranton, G., Gobillon, L., (2008) Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2), 723–742.
Combes, P.P., Duranton, G., Gobillon, L., (2011) The identification of agglomeration economies. *Journal of Economic Geography*, 11(2), 253–266.
Combes, P.P., Duranton, G., Gobillon, L., Puga, D., Roux, S., (2009) The productivity advantages of large markets: Distinguishing agglomeration from firm selection. *C.E.P.R. Discussion Papers*, 7191.
Combes, P.P., Gobillon, L., (2015) The Empirics of Agglomeration Economies. In G. Duranton, J. V Henderson, & W. C. Strange (Eds.), *Handbook of Urban and Regional Economics 5*, Elsevier.
Coulombe, S., Lee, F., (1993) *Regional economic disparities in Canada*, , University of Ottawa.
Cuadrado-Roura, J.R., (2001) Regional convergence in the European Union: From hypothesis to the actual trends. *The Annals of Regional Science*, 35(3), 333–356.
Cuadrado-Roura, J.R., García-Greciano, B., Raymond, J.L., (1999) Regional convergence in productivity and productive structure: The Spanish case. *International Regional Science Review*, 22(1), 35–53.
De la Fuente, A., (2002) On the sources of convergence: A close look at the Spanish regions. *European Economic Review*, 46(3), 569–599.
Desmet, K., Fafchamps, M., (2005) Changes in the spatial concentration of employment across US counties: a sectoral analysis 1972–2000. *Journal of Economic Geography*, 5, 261–284.
Díaz-Dapena, A., Fernández-Vázquez, E., Rubiera-Morollón, F., (2016) The role of spatial scale in regional convergence: the effect of MAUP in the estimation of beta-convergence equations. *The Annals of Regional Science*, 56(2), 473–489.
Dixon, R., Thirlwall, A.P., (1975) A model of regional growth-rate differences on Kaldorian lines. *Oxford Economic Papers*, 27(2), 201–214.
Duranton, G., Puga, D., (2000) Diversity and specialisation in cities: why, where and when does it matter? *Urban Studies*, 37(3), 533–555.
Eaton, J., Eckstein, Z., (1997) Cities and growth: Theory and evidence from France

and Japan. *Regional Science and Urban Economics*, 27(4), 443–474.
Eeckhout, J., (2004) Gibrat's law for (all) cities. *The American Economic Review*, 94(5), 1429–1451.
Ellison, G., Glaeser, E., (1997) Geographic concentration in US manufacturing industries: a databoard approach. *Journal of Political Economics*, 105, 889–927.
Escriba, J., Murgui, M.J., (2014) La base de datos BD.EURS (NACE Rev.1). *Investigaciones Regionales*, 28, 173–194.
Fujita, M., Krugman, P., Venables, A., (2001) *The spatial economy: Cities, regions, and international trade*, MIT press.
Gabaix, X., (1999) Zipf's law for cities: an explanation. *Quarterly Journal of Economics*, 739–767.
Gabaix, X., (2009) Power laws in economics and finance. *Annual Review of Economics*, 1, 255–293.
Gabaix, X., (2011) The granular origins of aggregate fluctuations. *Econometrica*, 79(3), 733–772.
Gabaix, X., Ibragimov, R., (2011) Rank− 1/2: a simple way to improve the OLS estimation of tail exponents. *Journal of Business & Economic Statistics*, 29(1), 24–39.
Gabaix, X., Loannides, Y., (2004) The evoluction of city size distribution. In *Handbook of regional and urban economics 4*, 2341–2378, Elsevier.
Glaeser, E.L., (1994) Why does schooling generate economic growth? *Economics Letters*, 44, 333–337.
Glaeser, E.L., (1998) Are cities dying? *The Journal of Economic Perspectives*, 12, 139–160.
Goldstein, H., (1986) Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika*, 73(1), 43–56.
Goldstein, H., (2011) *Multilevel statistical models*, John Wiley & Sons.
Gómez-Zaldívar, M., Ventosa-Santaulària, D., (2009) Liberación comercial y convergencia regional del ingreso en México. *El Trimestre Económico*, LXXVI(2)(301), 215–235.
González-Val, R., (2010) The Evolution Of US City Size Distribution From A Long-Term Perspective (1900–2000). *Journal of Regional Science*, 50(2), 952–972.
González Rivas, M., (2007) The effects of trade openness on regional inequality in Mexico. *The Annals of Regional Science*, 41(3), 545–561.
Graham, D.J., (2006) *Wider economic benefits of transport improvements: link between agglomeration and productivity*, Imperial College London.
Graham, D.J., Kim, H.Y., (2008) An empirical analytical framework for agglomeration economies. *The Annals of Regional Science*, 42(2), 267–289.
Greene, W.H., (2012) *Econometric Analysis*, New York University.
Hall, P., (2000) Creative cities and economic development. *Urban Studies*.
Henderson, J. V, (1988) *Urban development: theory, fact and illusion*, Oxford University Press, Oxford.
Henderson, J. V, (2003) *The urbanization process and economic growth: The so-what questionJournal of Economic growth*, Springer.
Henderson, J. V, (2010) Cities and development. *Journal of Regional Science*, 50(1), 515–540.
Henderson, J. V, Wang, H.G., (2007) Urbanization and city growth: The role of institutions. *Regional Science and Urban Economics*, 37(3), 283–313.
Hirschman, A.O., (1958) *The Strategy of Economic Development*, Yale University Press.
Hoover, E., Giarratani, F., (1971) *An introduction to regional economics*, Knopf.

Hox, J., (2010) *Multilevel analysis: Techniques and applications*, Routledge, New York.
Isard, W., (1956) *Location and Space-economy; a General Theory Relating to Industrial Location, Market Areas, Land Use, Trade, and Urban Structure*, MIT press.
Islam, N., (1995) Growth empirics: a panel data approach. *The Quarterly Journal of Economics*, 110(4), 1127–1170.
Islam, N., (2003) What have We Learnt from the Convergence Debate? *Journal of Economic Surveys*, 17(3), 309–362.
Jacobs, J., (1969) *The economy of cities*, Random House, New York.
Jofre-Monseny, J., (2009) The scope of agglomeration economies: Evidence from Catalonia. *Papers in Regional Science*, 88(3), 575–590.
Kaldor, N., (1957) A model of economic growth. *The Economic Journal*, 67(268), 591–624.
Koenker, R., (2005) *Quantile Regression*, Econometric Society Monographs.
Koenker, R., Bassett Jr, G., (1978) Regression quantiles. *Econometrica: Journal of the Econometric Society*, 46(1), 33–50.
Koo, J., Kim, Y.Y., Kim, S., (1998) Regional income convergence: evidence from a rapidly growing economy. *Journal of Economic Development*, 23(2), 191–203.
Krugman, P., (1991) Increasing Returns and Economic Geography. *The Journal of Political Economy*, 99(3), 483–499.
Krugman, P., (1996) *The Self-Organizing Economy*, Blackwell, Cambridge.
Krugman, P., (1998) What's new about the new economic geography? *Oxford Review of Economic Policy*, 14(2), 7–17.
Krugman, P., Venables, A.J., (1995) Globalization and the Inequality of Nations. *The Quarterly Journal of Economics*, 110(4), 857–880.
Li, Y., Wei, Y.H.D., (2010) The spatial-temporal hierarchy of regional inequality of China. *Applied Geography*, 30, 303–316.
Littell, R.C., Stroup, W.W., Milliken, G.A., Wolfinger, R.D., (2006) *SAS for mixed models*, SAS Institute. Inc.
López-Bazo, E., Vayá, E., Artís, M., (2004) Regional externalities and growth: Evidence from European regions. *Journal of Regional Science*, 44(1), 43–73.
Lucas, R.E., (1988) On the mechanics of economic development. *Journal of Monetary Economics*, 22(1), 3–42.
Lucas, R.E., (2001) Externalities and Cities. *Review of Economic Dynamics*, 4, 245–274.
Mankiw, N.G., Romer, D., Weil, D.N., (1992) A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, 107(2), 407–437.
Marshall, A., (1890) *Principles of political economy*, Maxmillan, New York.
Martin, P., Ottaviano, G.I.P., (1999) Growing locations: Industry location in a model of endogenous growth. *European Economic Review*, 43(2), 281–302.
Martínez-Galarraga, J., Paluzie, E., Pons, J., Tirado-Fabregat, D.A., (2008) Agglomeration and labour productivity in Spain over the long term. *Cliometrica*, 2(3), 195–212.
Matano, A., Naticchioni, P., (2015) What drives the urban wage premium? evidence along the wage distribution. *Journal of Regional Science*, 56(2).
Melo, P.C., Graham, D.J., Noland, R.B., (2009) A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics*, 39(3), 332–342.
Miller, J.R., Genc, I., (2005) Alternative regional specification and convergence of U.S. regional growth rates. *Annals of Regional Science*, 39(2), 241–252.
Myrdal, G., (1957) *Economic Theory and Under-developed Regions*, Duckworth,

London.
Openshaw, S., (1984) *The modifiable areal unit problem, Concepts and techniques in modern geography*, GeoBooks.
Ottaviano, G., Thisse, J., (2004) Agglomeration and economic geography. In J. V Henderson & J. F. Thisse (Eds.), *Handbook of regional and urban economics 4*, Elsevier, New York.
Paelinck, J.H.P., Polèse, M., (1999) Modelling the regional impact of continental economic integration: lessons from the European Union for NAFTA. *Regional Studies*, 33(8), 727–738.
Parr, J., (2002) Agglomeration economies: ambiguities and confusions. *Environment and Planning A*, 34, 717–731.
Pinheiro, J.C., Bates, D.M., (1996) Unconstrained parametrizations for variance-covariance matrices. *Statistics and Computing*, 6(3), 289–296.
Pinheiro, J.C., Bates, D.M., (2000) *Mixed-Effects Models in S and S-PLUS*, Springer.
Porter, M.E., (1990) The competitive advantage of notions. *Harvard Business Review*.
Prais, S.J., Winsten, C.B., (1954) Trend Estimators and Serial Correlation. *Cowles Comission Discussion Paper*, 383, 1–26.
Raiser, M., (1998) Subsidising inequality: economic reforms, fiscal transfers and convergence across Chinese provinces. *The Journal of Development Studies*, 34(3), 1–26.
Ramajo, J., Márquez, M.A., Hewings, G.J.D., Salinas, M.M., (2008) Spatial heterogeneity and interregional spillovers in the European Union: Do cohesion policies encourage convergence across regions? *European Economic Review*, 52(3), 551–567.
Resende, G.M., (2011) Multiple dimensions of regional economic growth: The Brazilian case, 1991-2000. *Papers in Regional Science*, 90(3), 629–662.
Rey, S.J., Montouri, B.D., (1999) US regional income convergence: a spatial econometric perspective. *Regional Studies*, 33(2), 143–156.
Rice, P., Venables, A.J., Patacchini, E., (2006) Spatial determinants of productivity: analysis for the regions of Great Britain. *Regional Science and Urban Economics*, 36(6), 727–752.
Rodríguez-Pose, A., Sánchez-Reaza, J., (2005) Economic polarization through trade: Trade liberalization and regional growth in Mexico. *Spatial Inequality and Development*, 237–259.
Romer, P.M., (1990) Capital, labor, and productivity. *Brookings Papers on Economic Activity. Microeconomics*, 337–367.
Romer, P.M., (1994) The origins of endogenous growth. *The Journal of Economic Perspectives*, 8(1), 3–22.
Rosenthal, S.S., Strange, W.C., (2001) The determinants of agglomeration. *Journal of Urban Economics*, 50(2), 191–229.
Sala-I-Martin, X., (1994) Cross-sectional regressions and the empirics of economic growth. *European Economic Review*, 38(3), 739–747.
Sala-I-Martin, X., (1996) The classical approach to convergence analysis. *The Economic Journal*, 1019–1036.
Sánchez-Reaza, J., Rodríguez-Pose, A., (2002) The impact of trade liberalization on regional disparities in Mexico. *Growth and Change*, 33(1), 72–90.
Secretaría de Comunicaciones y Transportes, (2008) Traza tu ruta. .
Shioji, E., (1992) *Regional growth in Japan*, Yale University.
Solow, R.M., (1956) A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 65–94.

Srholec, M., (2010) A Multilevel Approach to Geography of Innovation. *Regional Studies*, 44(9), 1207–1220.
Theil, H., (1954) *Linear aggregation of economic relations*, North Holland Publishing, Amsterdam.
Tsionas, E.G., (2000) Regional growth and convergence: evidence from the United States. *Regional Studies*, 34(3), 231–238.
Villarreal, C.C., Tykhonenko, A., (2007) Convergencia regional e inversión extranjera directa en México en el contexto del TLCAN, 1994-2002. *Investigación Económica*, 66(259), 15–41.
Williamson, J.G., (1965) Regional inequality and the process of national development: a description of the patterns. *Economic Development and Cultural Change*, 13(4), 1–84.
Wooldridge, J.M., (2011) *Introductory Econometrics*.
Zipf, G.K., (1949) *zipf human behaviour and the principle of least effort*, Addison Wesley Press, Cambridge.