The Annals of Regional Science (2020) 64:421–454
https://doi.org/10.1007/s00168-019-00960-9

SPECIAL ISSUE PAPER

Spatiotemporal methods for analysis of urban system dynamics: an application to Chile

Andrés Vallone[1,2] · Coro Chasco[3,4]

Received: 19 October 2018 / Accepted: 30 November 2019 / Published online: 10 December 2019

Abstract

This paper presents a methodological procedure to evaluate the influence of spatial proximity on evolution of cities to detect regional differences in their spatiotemporal dynamics. The six-step method based on a set of statistical methods can be computed with a new R package: *estdaR*. The first step consists of the usual characterization of the cross-sectional distribution of the urban areas by means of nonparametric estimations of density functions for a set of significant years. In the second and third steps, the growth process is modeled as a first-order stationary Markov chain to evaluate the effect of global and local spatial autocorrelation on the transition probabilities with a set of indices based on the spatial version of the standard Markov chain. The fourth, fifth, and sixth steps perform in-depth analysis to detect the existence and interaction of spatial regimes in the movement direction and ranking mobility of urban distribution. We apply this novel strategy for the period 1930–2002 to analyze the entire Chilean urban system—not only the Central Zone, in which most of the population and economic activities are concentrated, but also other urban zones in the country.

JEL Classification C14 · C21 · O18

✉ Coro Chasco
coro.chasco@uam.es

Andrés Vallone
avallone@ucn.cl

1 Dpto. Economía Aplicada, Universidad Autónoma de Madrid, Madrid, Spain

2 Escuela de Ciencias Empresariales, Universidad Católica del Norte, Antofagasta, Chile

3 Dpto. Economía Aplicada, Universidad Autónoma de Madrid, Av. Francisco Tomás y Valiente 5, 28049 Madrid, Spain

4 Universidad Nebrija, Madrid, Spain

1 Introduction

From an ecological point of view, individual cities are conceived as locationally separate from neighboring cities because they do not compete for territory. Hence, in terms of the consumption of space, cities would be largely independent of one another, with no constraint based on the development of a neighboring city (Parr 2012). Nevertheless, in urban subsystems with cities experiencing increasing population and expansion, competition for space increases, producing city concentration and economies of scale (Ye and Xie 2012). In fact, the formation of cities is closely related to Christaller's central places theory, which posits a functional relationship between the population of a central place and its complementary area (Mutlu 1986).

Hence, the urbanization process cannot be considered as spatially homogenous. It is closely related to economic development since urbanization occurs as countries shift from rural-agricultural activity to urban-industrial activity (Davis and Henderson 2003). Factors such as geography (Henderson et al. 2001), factor endowments (Venables 2005), spatial proximity among human settlements (Ioannides and Overman 2004), and public policies (Desmet and Henderson 2015) affect the evolution of urban systems. They may generate either regionally concentrated urban settlements (Antrop 2004) or polycentric urban poles (Paci and Usai 2008) as a consequence of spatial interaction processes in population growth.

In Latin American countries, including Chile, population and economic activities are especially concentrated in the capital city and its corresponding metropolitan area (Rodríguez 2007), leading in many cases to overconcentration of the population and economic activity in this area. People in Chile's provinces are fond of saying, 'God is everywhere, but his office is in Santiago.' In fact, the literature tends to focus on this megalopolis, forgetting the rest of the country.[1] With the exception of Escolano Utrilla et al. (2007), who analyze the entire group of Chilean cities, studies on this topic focus on urban dynamics of either the Metropolitan Region of Santiago (Rodríguez et al. 2009) or other individual cities (e.g., Bustos Validiva 2013; Escolano Utrilla and Ortiz Véliz 2004; Santiago et al. 2016). To the extent of our knowledge, no research has quantified the effect of space on Chilean urban dynamics over the last century.

This paper seeks to propose a methodology to identify spatiotemporal patterns in the evolution of cities and establish the influence of spatial proximity by assembling a set of six spatial statistical approaches. With the exception of the kernel density functions and the standard Markov chain, these methods have been applied primarily to study the evolution of spatial systems for crime and income distributions. In the urban literature, in particular, the spatial Markov chain (SMC) method has been used to analyze the historical development of Spanish cities (Le Gallo and Chasco 2008) and of Phoenix, Arizona (Kane et al. 2014), but no applications have been found for the other methods proposed.

This strategy is a six-step procedure based on a set of statistical methods that can be computed with a new R package, *estdaR*, which has been developed by the

[1] This phenomenon is also common in other Latin American countries with highly centralized governments exerting greater control over resources (Willis et al. 1999).

authors and is available under a GPL-2 license from the site https://github.com/amvallone/estdaR.[2] We employ the procedure to detect different trends and spatial clusters in the development of Chilean cities over the period 1930–2002, focusing specifically on how spatial proximity affects relative sizes and rankings. We seek to know to what extent certain urban processes, such as urban sprawl and population convergence, are homogeneous across the entire Chilean city system or whether a city's population grows faster or slower depending on its neighbors' growth speed.

To this end, we first analyze the cross-sectional distribution of urban population by means of standard statistical analysis and nonparametric estimations of density functions for a series of years, a method proposed by Quah (1996) and followed by many other authors (e.g., Xu and Zhu 2009; Xiufang et al. 2015). Second, the growth process is modeled as a first-order stationary Markov chain. Third, the role of geographical space in the transition probabilities of population growth is evaluated with a set of methods based on a spatial version of the standard Markov chain (Rey 2001). Fourth, we perform an in-depth analysis to detect local patterns in the joint transition of a city and its neighbors in the Chilean urban system. On the one hand, we use the LISA transition matrix (Rey and Janikas 2006), based on the Local Moran statistic proposed by Anselin (1995); on the other, we capture the co-movement[3] directions of cities and neighbors across the Moran scatterplot quadrants with the directional LISA (Rey et al. 2011) approach. Fifth, we study the existence of spatial regime differences in the ranking mobility of the Chilean urban distribution using the Global Indicators of Mobility Association (GIMA) (Rey 2016). Finally, we determine the ranking decomposition (Rey 2004) as a cohesion measure that detects synchronic rank movements among spatial regimes. To study Chilean urban dynamics, we selected the sample of contemporary cities and functional market areas from Chile's decennial censuses of 1930, 1940, 1952, 1960, 1970, 1982, 1992, and 2002.

The paper is organized as follows. In Sect. 2, we develop the methodology. In Sect. 3, we present the database and the main results obtained with this method for the Chilean urban system. The paper ends with a concluding section and the references.

2 The method

We propose a modeling strategy that improves knowledge of urban systems by evaluating the influence of spatial proximity among human settlements on cities' evolution to detect regional differences and interactions in their spatiotemporal dynamics. The strategy, the six-step method represented in Fig. 1, will be explained in the following subsections.

[2] *estdaR* must be installed in the R console: devtools::install_github("amvallone/estdaR").

[3] Throughout this paper, the terms "movement" and "mobility" refer to movements across the population distribution, as is common in the social inequality literature (Kang and Rey 2019). In this context, population mobility could thus be viewed as a re-ranking phenomenon in which cities switch population positions. Mobility could also be viewed as occurring, however, whenever cities move away from their previous city size levels. The former is termed absolute mobility and the latter relative mobility.

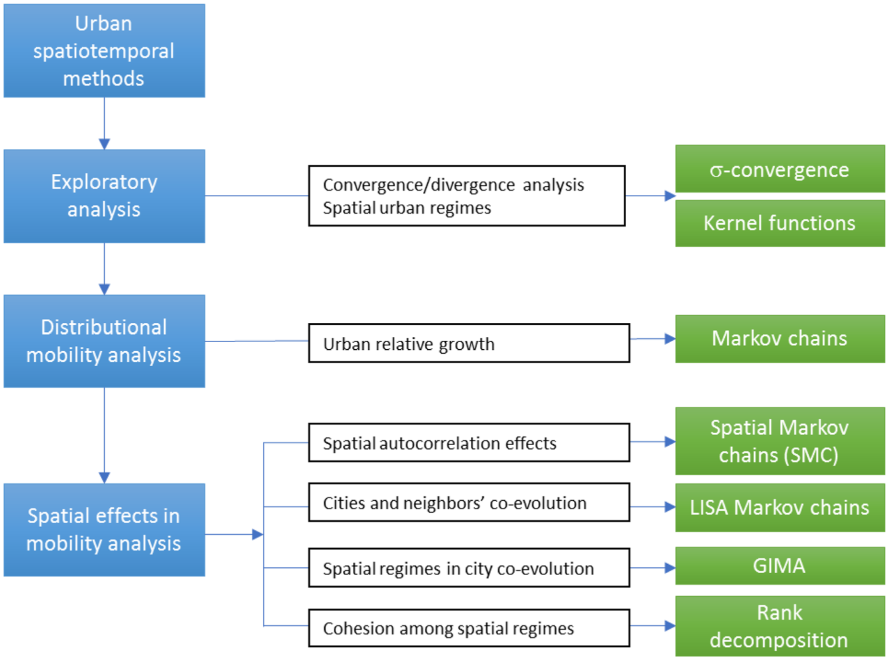

Fig. 1 Methodological procedure for evaluating urban spatiotemporal dynamics

2.1 Exploratory analysis of city-size distribution

Typical spatiotemporal exploratory data analysis does not permit inferences about patterns in the intertemporal evolution of the full cross-sectional distribution of cities in terms of size relative to the rest of the urban system. Such analyses also fail to take into account spatial interdependence between regions (Quah 1996). One traditional exploratory measure of urban convergence is the σ-convergence, which indicates a reduction of dispersion within city-size cross-sectional distribution over time. It is an interesting simple method concerned only with spread dispersion (second moment) of the population distribution. σ-convergence determines the dynamics of approximation between urban populations in a specific period by computing the standard deviation or coefficient of variation of relative log city population (Wu and He 2017). As stated above, since dispersion indicators provide no information about behavior of the overall population distribution, alternative concepts of convergence must be used.

We thus follow a strand of the literature that estimates nonparametric kernel density of urban population distributions for different periods. These density plots may be interpreted as the continuous equivalent of a histogram in which the number of intervals has been set to infinity and then to the continuum (Le Gallo 2004). We consider relative size distributions by normalizing the log of population size for each decade, divided by the log average size. Specifically, they are used to examine relative city-size distribution at present and the way this distribution has changed since a starting period to analyze its characteristics of mono- or multimodality. An increase or progressive concentration of the central probability mass of urban distributions in time can be interpreted as evidence of population convergence, while secondary distribution modes correspond to clubs of regional cities converging to a higher/lower population mean.

2.2 Analysis of city-size distributional mobility in urban dynamics using Markov chains

The density functions enable characterization of the evolution of the global population distribution, but they provide no information about the movements of the cities within this distribution over the course of the study period. That is, they do not say whether the right tail of the initial year distribution contains the same cities as the right tail in the final year distribution. One way to answer these questions is to track the evolution of each city's relative size over time by estimating transition probability matrices associated with Markov chains.

In this context, a Markov chain consists of a set of discrete states or discretization of the population distribution: $S=\{s_1, s_2,\ldots, sr\}$. As the population is a continuous variable, we must first discretize the continuous state-space of this variable, for example, by quantiles. This method describes a process that starts in one of these states (year) and moves successively from one state to another. Each move is called a step. If a city is currently in state s_i, then it moves to state s_j at the next step with a probability denoted by p_{ij}, which is called transition probability. If a city remains in the state it was in initially, this occurs with probability p_{ii} (Grinstead and Snell 1997).

The transition probabilities can be arranged in a square array called a transition matrix, P, such that the values in a specific row (e.g., the first) represent the probabilities that a city in a specific quantile i will move to this first quantile. The probability maximum likelihood estimator is defined as:

$$\hat{p}_{i,j} = \frac{\sum_t n_{i,j,t}}{\sum_t \sum_j n_{i,j,t}} \tag{1}$$

where $n_{i,j,t}$ is the number of times a sample chain started in state i in period t and transitioned to state j in the next period (Rey 2015).

Since improper discretization of a continuous variable could have the undesired effect of removing the Markov property and producing very misleading results (due to a certain degree of arbitrariness), discretization methods must satisfy two conditions. First, the initial classes should include a similar number of observations. Second, this discretization must perform best in the first-order test for Markovian property. As stated in Bickenbach and Bode (2003), the Markov property requires the transition probabilities, p_{ij}, to be of order 1, that is, to be dependent at the beginning of state t–1 (serially autocorrelated or order 1) and independent of states at the beginning of previous periods $t-2$, $t-3$,... If the chain is of a higher order, the transition matrix will be misspecified because it will contain only part of the information necessary to describe the true evolution of the population distribution. There are two first-order tests, the likelihood ratio (LR) and Pearson asymptotic Chi-square test statistics of $n-1$ degrees of freedom for n the city sample. They test the null hypothesis of the absence of serial autocorrelation against first-order autocorrelation. Hence, the best discretization method is the one that maximizes these tests from a group of alternative methods.

Finally, it is also important to compute the ergodic, steady-state, or limit distribution that can be interpreted as the city-size distribution in the long-run equilibrium. This function is used to assess the form of convergence in a distribution. A concentration of the frequencies of the ergodic distribution in a certain state or class implies convergence (if in the middle class, convergence to the mean). Concentration of the frequencies in some of the classes—that is, a multimodal limit distribution—can be interpreted as a tendency toward stratification into different convergence clubs. Finally, the dispersion of this distribution among all classes is interpreted as divergence.

2.3 The role of spatial dependence in city-size distributional mobility using spatial Markov chains

Spatial data have special properties and must be analyzed differently from non-spatial ones. Previous methods have not explicitly taken into account this spatial dimension of urban growth. In effect, population growth may spread around neighboring areas in a metropolization process or it may experience a chronic depopulation evolution, corresponding both situations to positive spatial autocorrelation,[4] that is, the presence of urban growth interactive processes of agglomeration or dispersion, respectively. On its part, negative spatial autocorrelation implies a systematic dispersed polarization of large cities in a country (Paci and Usai 2008).

Markov chain analysis enables integration of spatial dependence of the data by estimating the spatial Markov transition matrix or SMC (Rey 2001).[5] This method reports the probability of a particular transition conditioned by the populations of the city's neighbors in the preceding period. As noted in Rey (2015), for a quintile discretization of the population distribution, the maximum likelihood estimator of transition probabilities is:

$$\hat{p}(l)_{i,j} = \frac{\sum_t n(l)_{i,j,t}}{\sum_t \sum_j n(l)_{i,j,t}} \quad (2)$$

where $n(l)_{i,j,t}$ is the number of times a sample chain with a spatial lag in quintile l started in state i in period t and transitioned to state j in the next period.

[4] Spatial autocorrelation and spatial dependence are used as interchangeable terms, though in strong sense, the first is a specific type of the second.

[5] Spatial Markov chains have also been applied in other contexts, such as employment of disabled people (Agovino 2014), pro-environmental behavior (Agovino et al. 2016) and quality of life (Delmelle et al. 2016).

In the SMC, the transition dynamics are divided across cities with spatial lags in the different classes at the preceding year. Hence, for a quintile discretization of the transition matrix, we compute five transition probability matrices: $l(1)$ is the chain corresponding to the spatial lags of population values located at the lowest quintile of this distribution and $l(5)$ the chain of spatially lagged population values in the upper quintile. Computing the spatial lags requires a spatial weight matrix ('W') to capture the potential spatial interaction among cities in the urban system. It is an $n \times n$ matrix, for n the total number of cities, with the main diagonal elements (w_{ii}) set to zero by definition and the rest of the nonzero cells (w_{ij}) capturing the degree of spatial dependence among observations i, j. There is a rich variety of ways to specify the structure of these weights (Anselin and Rey 2014).

To contrast the influence of space on the transition and the homogeneity across lagged classes, we compute two statistics: Pearson's Q test and the likelihood ratio (LR) test (Rey et al. 2016), both distributed as an asymptotic Chi-square. These statistics test the null hypothesis that the initial transition probabilities of the cities in the population distribution are spatially independent, that is, that they are not influenced by the values of their corresponding surrounding cities (spatial lag).

2.4 Analysis of the co-evolution of cities and spatial neighbors using LISA methods

The SMC provides insight into the role of spatial neighbor cities at the beginning of the transition, but it cannot analyze the joint evolution of cities and neighbors in the urban system dynamics. To analyze these issues, we use two different methods: the LISA Markov Chain (Rey and Janikas 2006) and the directional LISA (Rey et al. 2011).

2.4.1 Local indicator of spatial association (LISA) Markov chain

The LISA transition matrix (Rey and Janikas 2006) is based on the local Moran statistic proposed by Anselin (1995) to identify local clusters and spatial outliers. The LISA Markov chain computes the joint transition of a city and its neighbors in the distribution by measuring their movements across the four quadrants of the Moran scatterplot. Conventionally, the upper-right quadrant and the lower-left quadrant correspond to positive spatial autocorrelation (similar values at neighboring locations) and are referred to, respectively, as high–high (HH) and low–low (LL) spatial autocorrelation. The lower-right and upper-left quadrants, in contrast, correspond to negative spatial autocorrelation (dissimilar values at neighboring locations), referred to, respectively, as high–low (HL) and low–high (LH) spatial autocorrelation.

The states of the LISA Markov chains are the four quadrants of the Moran scatterplot in a given period. In each period, a city can be classified into four mutually exclusive categories HH, LH, LL, and HL where, for example, HL indicates a city above the system average for that period while its neighbors' mean size is below the average. From period to period, a city's position in the Moran scatterplot may change among the quadrants, with 16 possible transitions. A formal test for co-movement dependence—based on (Rey et al. 2012)—can also be performed by decomposing the LISA Markov chain into a pair of chains, one for the city and the other for the neighbors. Each chain has two states: H and L. The statistics follow an asymptotic χ^2 distribution, where the null hypothesis is the independence of the two chains (Rey 2015).

2.4.2 Directional LISA

The LISA Markov method computes the probability that cities and spatial neighbors will move from the Moran scatterplot states, but it is not possible to observe this evolution on a diagram. One way of capturing the co-movements of cities and neighbors graphically across the Moran scatterplot is the directional LISA approach (Rey et al. 2011). This method visualizes these co-movements by means of the origin-standardized movement vector, obtained by comparing two Moran scatterplots corresponding to two different periods of time, for example, the first and last period of analysis.

This technique is very appropriate to test for different dynamics between spatial regimes or urban subsystems, as Gregory and Patuelli (2015) do for the German regions. To obtain a clearer view of the movement patterns' heterogeneity, rose diagrams are also very helpful (Rey et al. 2011; Gutiérrez and Rey 2013). The rose plot, based on circular statistics, is a circular histogram that shows the frequency of moves across different directions based on angular notation.

Rose diagrams are also used to represent the directional LISA inference results graphically.[6] The null hypothesis is that the distribution of the vectors across the segments reflects independence in the movements of the focal unit and its spatial lag. Inference is based on random spatial permutations under the null hypothesis, and it is necessary to identify the statistically significant values of directional LISA arrows and rose graph sectors.

2.5 Computation of spatial regime disparities in the co-evolution of cities and neighbors using GIMA

Rey (2016) proposes the Global Index of Mobility Association (GIMA), whose antecedent is the Kendall's spatial τ (Rey 2004). This indicator, which is based on the rank correlation coefficient (Kendall 1962), enables measurement of disparities or inequalities in city size using concordance or discordance of the ranks of this variable in two periods of time $\tau\left(y_t, y_{t+1}\right)$, as follows:

[6] A rose diagram is a circular chart to display data that contain direction and magnitude variables. They normally comprises of 8 or 16 radiating spokes, which represent degrees of a circle or compass points North, East, South, West and their intermediate directions. Each direction axis has values increasing outwards and similar to pie charts, the data are divided into proportional slices or sectors. The arc length of each slice is proportional to the quantity it represents.

$$\tau\left(y_t, y_{t+1}\right) = \frac{c - d}{\frac{n(n-1)}{2}} \tag{3}$$

where c is the number of concordant pairs, d the number of discordant pairs, and n the city sample size. The τ index ranges from 0 (perfect discordance) to 1 (perfect concordance). From this expression, we can build the GIMA mobility index as follows:

$$M = \frac{\tau\left(y_t, y_{t+1}\right) - 1}{-2} \tag{4}$$

M varies from 0 to 1, where 1 implies full ranking mobility and 0 complete stability of the ranking. As this statistic yields a single value for the amount of rank mobility in the entire population distribution over two points in time, it is considered a global indicator.

It is thus possible to decompose the τ index to capture the effect of space in the ranking changes as follows[7]:

$$\tau\left(y_t, y_{t+1}\right) = \varphi\tau_W\left(y_t, y_{t+1}\right) + (1 - \varphi)\tau_{\bar{W}}\left(y_t, y_{t+1}\right) \tag{5}$$

where $\varphi = \frac{i'Wi}{i'(W+\bar{W})i}$, with i a unit vector of order ($n \times 1$), W a spatial weights matrix containing the neighboring relationships, and $\bar{W} = ii' - W - I_{n\times n}$ a matrix capturing the non-neighboring relationships.

Equation (5) presents the spatial τ index of concordant and discordant rank pairs as a decomposition of two τ indexes: one for pairs of neighboring cities and the other for pairs of non-neighboring cities. This procedure enables us to identify the different correlation patterns between neighboring and non-neighboring cities. Different ranking patterns may be inferred based on random spatial permutations of the attributes to develop a distribution for τ_W under the null hypothesis of spatial homogeneity in the correlation patterns (Rey 2016).

As with Kendall's τ, it is possible to construct a spatial mobility index as follows:

$$M_W = \frac{\tau_W\left(y_t, y_{t+1}\right) - 1}{-2} \tag{6}$$

Equation (6) also allows for additive decomposition of overall mobility, which gives the option of comparing different levels of mobility between neighboring and non-neighboring cities, as follows:

$$M = \varphi M_W + (1 - \varphi)M_{\bar{W}} \tag{7}$$

It is possible to partition the cities into regimes, which can be used to operationalize neighbors using the so-called block weights (Anselin and Rey 2014, p. 37) such that $w_{ij} = 1$ if R(i) = R(j), otherwise $w_{ij} = 0$, where i, j are cities and R the regimes.

7 Rey (2016) presents the full mathematical decomposition of this index.

2.6 Rank decomposition of city size by spatial regimes

The rank decomposition index $\Theta_{t_1-t_0}$ (Rey 2004) is defined as the sum of rank changes, from period t_0 to t_1, within a regime over the sum of the overall rank changes. Formally, if we set $\theta_{i,t}$ as the position in the ranking of city i in period t and assume the existence of R spatial regimes, the rank decomposition index Θ is calculated as follows:

$$\Theta_{t_1-t_0} = \frac{\sum_R \left| \sum_{i\in R} \theta_{i,t_1} - \theta_{i,t_0} \right|}{\sum_i \left| \theta_{i,t_1} - \theta_{i,t_0} \right|} \tag{8}$$

The denominator of this measure is the sum of the absolute rank changes over the period.

For a sequence of time periods, θ measures the extent to which rank changes for a variable measured over n locations are in the same direction within mutually exclusive and exhaustive partitions (regimes) of the n locations. The cohesion index will take the value 0 in the case of complete absence of cohesion (i.e., when all changes in the ranking occur only inside the same regime). At the other end, Θ will take the value 1 when all movements in the ranking are 100% 'cohesive' within the regimes. In this case, all cities from one spatial regime will be increasing their ranks at the expense of cities belonging to another regime.

In this context, cohesion can be understood as a process of migration between regimes, such that the size of cities in one regime increases (ascending in the ranking) at the expense of the size of cities in the other regimes, which decline in the ranking. Full cohesion thus implies a perfect population transfer between regimes.

Although Θ is a nonparametric test, it is possible to construct an inferential process based on random spatial permutations under the null hypothesis of spatial homogeneity (Rey 2004). One drawback of this index, however, is that it cannot provide information about the direction of the migration flow, which must be derived from alternative information sources.

3 Application on city growth in Chile

How to define cities and their consistency over time is a question that arises often in the literature. Some authors use the official statistics, based on the authorities' definition of city boundaries (Soo 2014; Lanaspa et al. 2003). Since these statistics may not coincide with the economically meaningful definition of city, however, other authors estimate metro areas to cover the local labor market of a core city, accommodating changes in geographic definitions over time (Henderson 2005; Schmidheiny and Suedekum 2015, among others).

To explore the spatiotemporal dynamics of Chilean cities from 1930 to 2002, we need a data set with urban areas defined consistently over this period. Evolution of the population distribution is analyzed using the Census data over the eight decades under consideration: 1930, 1940, 1952, 1960, 1970, 1982, 1992, and 2002.[8] The data on population are extracted from the Chilean Office for Statistics (INE) databank.

In the first place, we use the official definition proposed by the Chilean Office for Statistics (Instituto Nacional de Estadísticas 2005), which defines a 'city' as any urban entity with more than 5000 inhabitants.[9] This definition is consistent with the definition of city as a continuously built-up area, different from the so-called large city, which represents the metropolitan area (Parr 2012). This sample of cities, which are located in all regions of the country, has the advantage of including a sufficient number of urban entities for the estimations and tests required by the methodology, some of which are based on asymptotic assumptions (Hamilton 1994). Since what was a group of different middle-size cities in the past may now be part of a large city, however, there is a risk of overestimating the spatial effects of autocorrelation and heterogeneity (spatial regimes). For this reason, we also use the Labor Market Areas (LMAs) recently defined by Casado-Díaz et al. (2017) as a robustness check for the influence of spatial proximity on relative city sizes and regimes.

Based on the official definition, we identified 184 Chilean cities in the 2002 Census that existed in the 1930 Census[10] to study their evolution during the eight periods considered (more than 70 years). These present-day cities are located throughout the Chilean territory. The LMAs are a group of 62 functional regions, built with an evolutionary computational method, as a sum of municipalities, to capture the extent of commuting fields of residents. Due to the frequent changes experienced by the municipalities in their boundaries during the period 1930–2002 (Rowe 2017), the LMAs of this panel were approximated as aggregations of the official cities. A summary of cities and LMAs is presented in Fig. 2, and the full list is given in 'Appendix 1.'

Next, we present the results obtained by applying the modeling strategy to the Chilean urban system.

---

[8] Despite the existence of previous censuses, we choose 1930 as the first period of analysis because the Southern city of Aysen was founded in 1928. We include Aysen in the sample because cities are scarce and sparsely disseminated in the far South of Chile. We excluded information from the 2012 Chilean census due to significant methodological problems. The new 2017 census data on entities are not yet available. For more information, see Instituto Nacional de Estadísticas (2014).

[9] The Republic of Chile is politically divided into regions, provinces, "comunas" (municipalities) and censal districts. Each municipality contains different "entities": cities, towns, villages, and hamlets, among others.

[10] There are 10 cities registered in the 2002 Census but not in the 1930 Census: Alto Hospicio (Region I), Estación Zaldivar (Region II), El Salvador (Region III), El Quisco (Region V), Quirihue (Region VIII), Padre de las Casas and Labranza (Region IX), Panguipulli and Los Muermos (Region X), and Padre Hurtado (Metropolitan Region). To homogenize the panel database, we added the population of these new cities to their corresponding originals. For example, since Alto Hospicio became independent of Iquique before the 2002 Census, the population of the former was added to that of the latter.

Spatial regimes	Region		Cities	LMAs
North	XV	Arica and Parinacota	1	1
	I	Tarapacá	2	1
	II	Antofagasta	7	3
	III	Atacama	7	3
	IV	Coquimbo	11	3
Center	V	Valparaíso	30	6
	RM	Metropolitan Region	22	2
	VI	O'Higgins	19	7
	VII	Maule	13	8
	VIII	Biobío	30	7
South	IX	Araucanía	18	7
	XIV	Los Ríos	5	3
	X	Los Lagos	15	8
	XI	Aysén	2	1
	XII	Magallanes	2	2

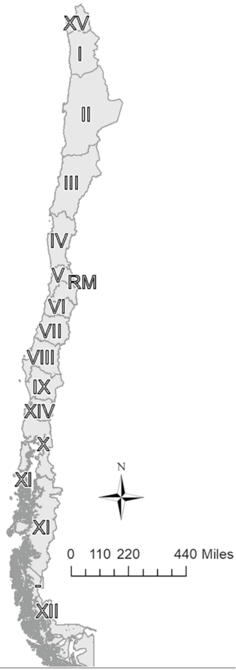

Fig. 2 Cities and Labor Market Areas (LMAs) by Chilean regions and spatial regimes

3.1 Exploratory analysis of the evolution of the Chilean urban system

Figure 3 shows the coefficient of variation, which is more robust than the standard deviation, to measure the σ-convergence or dispersion in the Chilean cities and LMAs, as well as the kernel distributions in 1930, 1952, 1970, and 2002. These periods correspond to the start and end-points of the period plus two other important events in Chilean history. The first is the decade of the 1950s, when the Import Substitution Industrialization (ISI) protectionist policy was implemented. The second is the 1970s, a decade of tremendous changes: Pinochet's accession to power, abandonment of the ISI model, and growth of the tertiary sector to the detriment of industry, which was progressively concentrating in Santiago (Henríquez et al. 2006).

The results for the σ-convergence charts show that the coefficient of variation follows an almost persistent decline trend in the period, whether for cities or LMAs. This provides initial evidence of urban convergence in the overall Chilean urban system. The kernel density functions for the four periods selected enable us to examine the relative population distributions of cities and LMAs in 1930 and the way these distributions changed over time until 2002. On the horizontal axis, the value 1 indicates average city/LMA size in Chile, 1.5 a value 50% higher than this average, and so on. Compared with 1930, more cities/LMAs reported values below the Chilean average in 2002, and there is evidence for bimodality at the end of the period, with a second small mode situated at around 50% of the average.

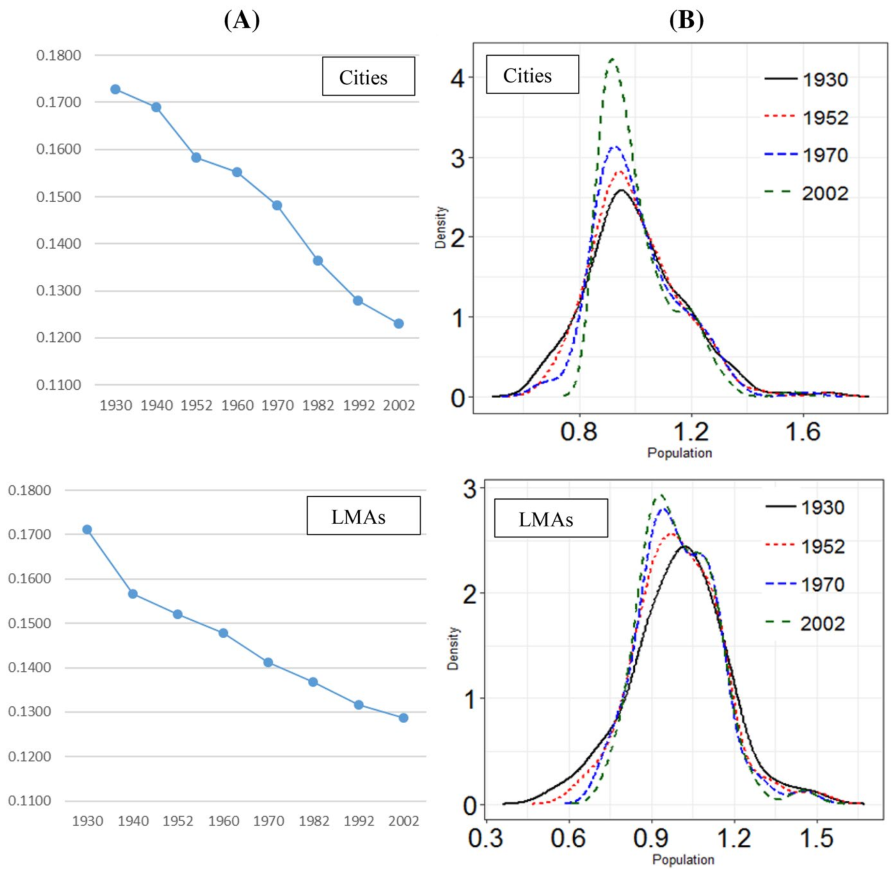

Fig. 3 Coefficient of variation (left) and kernel log-densities (right) of city and LMA populations in Chile, 1930–2002

These results may reflect a general convergence process in the Chilean urban population, as already shown by the σ-convergence measures. We also find, however, a divergent group of regions with an urban size well above average, converging toward a higher population level than the other regions. This is a group of the 16 largest cities in the country, 11 of which belong to the 'Central Zone' of Chile. This zone was defined by the Chilean Economic Development Agency (CORFO) in 1950 and contains five regions: Valparaíso, Metropolitan Region, O'Higgins, Maule, and Biobío ('Appendix 2'). In this natural area, we find an urban subsystem composed of a network of well-integrated and communicated cities of Spanish foundation, with linear integration and very high population growth (Olave 2005).

The Central Zone can be considered as a spatial regime that divides the country into two other regimes: the regions to the north (Arica and Parinacota, Tarapacá, Antofagasta, Atacama, and Coquimbo) and the regions to the south (Araucanía, Los Ríos, Los Lagos, Aysén, and Magallanes). The Northern regime is characterized by the existence of a sparser collection of cities, some of which were built by the Spanish conquerors as a linear terrestrial corridor connecting Santiago with Lima. These cities owe their economic development to their rich saltpeter, nitrate, and copper mines (Geisse 1977). The regions in the Southern regime are characterized by having being separated—for centuries—from the rest of the country because of the Mapuche insurrection. These zones have a rich variety of natural resources and great vegetal, animal, and fishing wealth. The cities and LMAs located in this regime are connected with each other and with the center primarily by the Austral Highway, with the exception of the Magallanes region, which has terrestrial communication through Argentina only.

Figure 4 represents the kernel distributions of these three regimes for cities and LMAs. Some interesting results to highlight concern the evolution of the central cities, which follow a pattern similar to that of the national group, in contrast to the rest of the country's urban system. First, in the three urban groups, the central mass of distributions increases more or less significantly in 1970, peaking in the 2002 distribution. This progressive concentration of probability mass can be interpreted as evidence of population convergence, consistent with the evolution of many other developed and developing countries (e.g., Anderson and Ge 2005 in China or Nitsch 2001 in some European countries).

The convergence trend seems to be more acute in the Southern city group, although not uniformly so, due to an observed second mode starting in 1952 that corresponds to a club of regional cities converging to a higher population mean. The convergence trend is due to an increase in the size of mid-size cities, such as Puerto Varas and Coihaique, which have grown at greater rates than large cities (Henríquez et al. 2006). The pattern of diversification of this urban subsystem registered by this outcome is due, however, to higher population growth in the main regional cities of this regime (Temuco, Puerto Montt, Osorno, and Punta Arenas), which experienced agglomeration economies and population concentration.

Cities in the Northern regime have a changing bimodal shape over the period. In 1930, a main mode of city clubs was converging to a higher population mean, with a second mode of cities converging to a lower population mean. This situation changes from 1952 to 2002, when the main mode represents most of the cities of this cluster converging to the mean population value of this regime, with a second mode of larger cities converging to a higher population value. We thus find a club of larger cities at the beginning of the period—mainly in region II (Tocopila, Chuquicamata, María Elena, Taltal), which decreases gradually in population until 2002—, to which the smaller-sized club catches up. These results could demonstrate the dominance of some consolidated largest regional cities (Antogasta and La Serena, in this case) over the urban subsystem distribution, a phenomenon also observable in some European countries (Nitsch 2001).

As a robustness check, we also represent the kernel density functions of the LMAs, which exhibit a shape similar to that of their city counterparts, with the exception of the Northern regimen, probably due to a very small sample size (only 11 functional areas). These results show that the convergence trend observed for the full set of Chilean cities is mainly determined by the central regime cities, which experienced periurban growth conditioned by the urban core of Santiago (Puertas

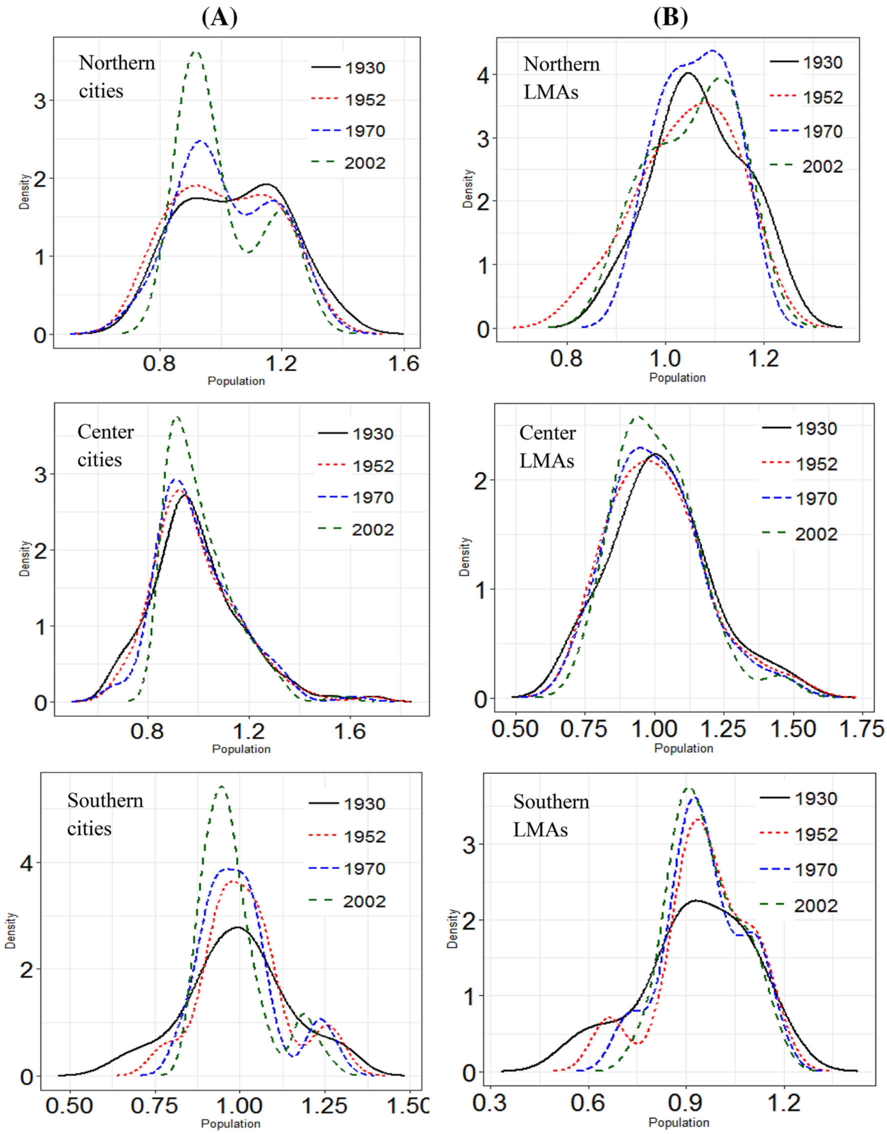

Fig. 4 Kernel log-densities of city (**a**) and LMA (**b**) populations in Chile by spatial regimes, 1930–2002

et al. 2014). The cities in the Northern and Southern regimes remain more or less polarized in two groups of diverging cities, however.

3.2 Analysis of city-size distributional mobility in the Chilean urban system

In order to estimate the transition matrix for the Chilean cities in the study period, we discretize the population variable using quintiles because they satisfy two

Table 1 Markov probability matrices

Cities						LMAs					
State	Q1	Q2	Q3	Q4	Q5	State	Q1	Q2	Q3	Q4	Q5
Q1	**0.846**	**0.127**	0.023	0.004	0.000	Q1	**0.912**	**0.088**	0.000	0.000	0.000
Q2	**0.139**	**0.687**	**0.170**	0.004	0.000	Q2	**0.095**	**0.821**	**0.071**	0.012	0.000
Q3	0.012	**0.191**	**0.698**	**0.099**	0.000	Q3	0.000	**0.083**	**0.881**	0.036	0.000
Q4	0.004	0.000	**0.100**	**0.826**	**0.070**	Q4	0.000	0.000	0.048	**0.881**	**0.071**
Q5	0.000	0.000	0.000	**0.070**	**0.931**	Q5	0.000	0.000	0.000	**0.066**	**0.934**
π	*0.201*	*0.201*	*0.196*	*0.201*	*0.201*	π	*0.210*	*0.194*	*0.194*	*0.194*	*0.210*

Note: In italics, the ergodic distribution values and in bold, probabilities equal or greater 0.05

conditions. First, the initial classes include a similar number of observations (a total of 296 cities per quintile, with the exception of 3 quintiles with 288 cities); second, the discretization performs best in the first-order test for Markovian property from a set of alternative discretization methods (quartiles, sextiles and natural breaks). For the group of Chilean cities, the test values corresponding to the quintile discretization are χ^2_{26} = 35.76, p value = 0.04 and = 34.16, p value = 0.10 and for the LMAs, they are χ^2_{18} = 39.54, p value = 0.00 and LR = 32.75, p value = 0.02.

Table 1 shows the results of the first-order probability matrix transition of the Chilean cities and LMAs. To aid in interpretation, the probabilities over 0.1 (before rounding) have been highlighted in the matrices. These outcomes yield three remarkable observations. First, as in many other social science phenomena, like income inequality analysis, path dependence plays an important role in dynamics with very low inter-class mobility, since the main diagonal of both matrices has the highest probability value for all states. Second, the mobility of the system occurs in small and medium-size cities. On the one hand, small and small-to-medium cities are more likely to move upwards in the distribution.

For example, during the decades from 1930 to 2002, the majority of small cities with population in the first quantile (84.6%) remained in that size class at the end of each decade, while 1.3% moved up one class by the end of the decade. On the other hand, the medium and medium-to-large cities were more likely to move downwards. This was the case of cities with population values in the third quintile, which had an estimated probability of 0.170 to move to the second quintile. Third, the last row of Table 1 shows the ergodic or steady-state distribution (π) tending to a uniform distribution, demonstrating the presence of urban divergence. Detailed examination shows that the largest values of this distribution are located in the extreme quintiles, possible evidence of a slight tendency toward stratification into two convergence clubs—small versus large cities or LMAs—more or less consistent with the results for the kernel functions shown in Fig. 3.

Table 2 Spatial Markov chain matrices for the Chilean cities and LMAs

Cities

	Q 1	Q 2	Q 3	Q 4	Q 5
l(1) Q 1	0.788	0.173	0.038	0.000	0.000
l(1) Q 2	0.178	0.600	0.222	0.000	0.000
l(1) Q 3	0.000	0.216	0.667	0.118	0.000
l(1) Q 4	0.000	0.000	0.086	0.914	0.000
l(1) Q 5	0.000	0.000	0.000	0.000	1.000
l(1) π	0.181	0.215	0.254	0.349	0.000
l(2) Q 1	0.826	0.130	0.043	0.000	0.000
l(2) Q 2	0.134	0.731	0.134	0.000	0.000
l(2) Q 3	0.000	0.190	0.746	0.063	0.000
l(2) Q 4	0.000	0.000	0.130	0.857	0.013
l(2) Q 5	0.000	0.000	0.000	0.138	0.862
l(2) π	0.247	0.320	0.282	0.138	0.013
l(3) Q 1	0.844	0.125	0.000	0.031	0.000
l(3) Q 2	0.196	0.696	0.109	0.000	0.000
l(3) Q 3	0.016	0.175	0.746	0.063	0.000
l(3) Q 4	0.000	0.000	0.136	0.818	0.045
l(3) Q 5	0.000	0.000	0.000	0.104	0.896
l(3) π	0.326	0.245	0.193	0.165	0.072
l(4) Q 1	0.862	0.123	0.015	0.000	0.000
l(4) Q 2	0.119	0.729	0.153	0.000	0.000
l(4) Q 3	0.023	0.159	0.705	0.114	0.000
l(4) Q 4	0.000	0.000	0.116	0.721	0.163
l(4) Q 5	0.000	0.000	0.000	0.125	0.875
l(4) π	0.201	0.199	0.184	0.180	0.235
l(5) Q 1	0.874	0.103	0.023	0.000	0.000
l(5) Q 2	0.071	0.643	0.262	0.024	0.000
l(5) Q 3	0.032	0.226	0.548	0.194	0.000
l(5) Q 4	0.017	0.000	0.033	0.817	0.133
l(5) Q 5	0.000	0.000	0.000	0.026	0.974
l(5) π	0.047	0.034	0.033	0.143	0.744

Probabilities: 1.00, 0.75, 0.50, 0.25, 0.00

LMAs

	Q 1	Q 2	Q 3	Q 4	Q 5
l(1) Q 1	0.941	0.059	0.000	0.000	0.000
l(1) Q 2	0.045	0.864	0.091	0.000	0.000
l(1) Q 3	0.000	0.250	0.700	0.050	0.000
l(1) Q 4	0.000	0.000	0.125	0.750	0.125
l(1) Q 5	0.000	0.000	0.000	0.062	0.938
l(1) π	0.300	0.389	0.141	0.057	0.113
l(2) Q 1	0.923	0.077	0.000	0.000	0.000
l(2) Q 2	0.111	0.833	0.056	0.000	0.000
l(2) Q 3	0.000	0.042	0.958	0.000	0.000
l(2) Q 4	0.000	0.000	0.000	1.000	0.000
l(2) Q 5	0.000	0.000	0.000	0.250	0.750
l(2) π	0.382	0.265	0.353	0.000	0.000
l(3) Q 1	0.818	0.182	0.000	0.000	0.000
l(3) Q 2	0.125	0.812	0.062	0.000	0.000
l(3) Q 3	0.000	0.000	0.875	0.125	0.000
l(3) Q 4	0.000	0.000	0.091	0.818	0.091
l(3) Q 5	0.000	0.000	0.000	0.026	0.974
l(3) π	0.000	0.000	0.140	0.193	0.667
l(4) Q 1	0.889	0.111	0.000	0.000	0.000
l(4) Q 2	0.133	0.800	0.067	0.000	0.000
l(4) Q 3	0.000	0.000	0.889	0.111	0.000
l(4) Q 4	0.000	0.000	0.000	0.920	0.080
l(4) Q 5	0.000	0.000	0.000	0.176	0.824
l(4) π	0.000	0.000	0.000	0.688	0.312
l(5) Q 1	0.895	0.105	0.000	0.000	0.000
l(5) Q 2	0.154	0.692	0.077	0.077	0.000
l(5) Q 3	0.000	0.043	0.957	0.000	0.000
l(5) Q 4	0.000	0.000	0.050	0.900	0.050
l(5) Q 5	0.000	0.000	0.000	0.000	1.000
l(5) π	0.000	0.000	0.000	0.000	1.000

Probabilities: 1.00, 0.75, 0.50, 0.25, 0.00

3.3 The role of spatial dependence in Chilean city-size distributional mobility

As shown previously, the urbanization process cannot be considered as spatially homogenous because it is closely related to economic development. Population tends to locate in the largest cities producing city concentration and economies of scale. To shed light on this issue, we analyze the influence of spatial proximity in terms of city-size intra-distributional mobility. The estimation of SMC transition matrices makes this evaluation possible for Chilean urban dynamics. We first built the chain's spatial lags, $l(1)$ to $l(5)$, for which we tested several definitions of the spatial weight matrix. We finally selected an inverse distance weights matrix, which is the impedance function of the gravity model, since distance is a key factor in city growth and spatial pattern of city sizes (Ioannides and Overman 2004).[11] Table 2 reports the five transition probability matrices corresponding to the spatial lag of city size and LMAs divided by quintiles.

Spatial independence in the distributional transitions has been contrasted with the Q statistic and the LR test, which take the following values: for the set of Chilean cities, $\chi^2_{52} = 70.131$, p value = 0.048 and LR = 70.938, p value = 0.024; for the LMAs, $\chi^2_{36} = 36.470$, p value = 0.447 and LR = 40.47, p value = 0.280. These results enable us to reject the null hypothesis of spatial homogeneity behavior for the Chilean cities. The nonsignificant values of these tests for the LMAs must be interpreted with caution, however, due to the asymptotic nature of the statistical inference.

We might infer a predominance of path dependence, independently of lag size, since the main diagonal probability matrix always assumes the major probability value for either cities or LMAs. Specifically, for the set of Chilean cities, the results in Table 2 provide clear evidence that the probability of an upward or downward move for a city in the population distribution will differ depending on its urban area

[11] As a robustness check, similar results were obtained with other spatial weight specifications, such as driving distance and other neighborhood measures. Complete computations are available from the authors upon request.

context. For example, the probability that a middle city (Q3) will move up to the next quintile (Q4) in the hierarchy is 11.8% when its spatial lag (in the preceding decade) contained on average a population of small cities (Q1), versus 19.4% probability when it contained on average a population of big cities (Q5). Conversely, the probability that a small-to-middle-sized city (Q2) will lose population, moving down to the first quintile in the hierarchy, is 17.8% when its spatial lag (in the preceding decade) contained on average a population of small cities (Q1), versus only 7.1% (more than two times lower) when it contained on average a population of big cities (Q5). In most cases, cities tend to move up or down in the population hierarchy when their corresponding city neighbors are on average larger or smaller, respectively, which would be evidence of the existence of positive spatial autocorrelation in urban population growth. In other words, the bigger Chilean cities are likely to agglomerate and promote growth in their neighborhood while the smaller peripheral ones suffer in general from endemic shrinking processes, inducing growing inequalities among cities in Chile.

The results are similar but less significant for the set of LMAs, particularly when their corresponding neighbors are in the intermediate quintiles of the population distribution (Q2–Q4). For example, the probability that a middle-sized LMA (Q3) will lose population, moving down to the second quintile in the hierarchy, is 25% when its spatial lag (in the preceding decade) contained on average a population of small LMAs, but only 4.3% when it contained on average a population of large LMAs. The effect of spatial dependence in LMAs population growth may reveal a certain degree of labor market interdependence in Chile. This phenomenon is promoted by the cooperation in economic, administrative and cultural functions, and the absence of barriers to factor mobility and trade, in a way that a perturbation in one LMA's economy is felt in its neighboring LMAs (Hoàng 2013; Rozenblat and Pumain 2018).

In addition, there is an absorbing state in the $l(5)$ of the largest LMAs' spatial neighbors, in the large LMA quintile. That is, distributional movements in the group of large LMAs are practically null independently when they were surrounded (in the preceding decade) by large neighboring LMAs, respectively. This absorbing state situation also takes place in large cities neighbored by small ones. That is, the smallest and largest cities and LMAs are not affected, in terms of population size, by neighbor cities' size.

Finally, the ergodic or steady-state distributions, π, for neighboring cities of small and small-to-medium size, $l(1)$–$l(3)$, are left-skewed distributions. This means that having small-to-medium-size cities nearby increases the probability that a city will shrink and become a smaller town in the long run. The $l(4)$ steady-state distribution corresponding to the medium-large neighboring cities tends to a uniform distribution, as in the non-spatial case, but the $l(5)$ chain of the largest neighboring cities has a right-skewed distribution, implying a long-term concentration process. The results for the LMAs are quite similar. Hence, the highest probability that a city/LMA will become a metropolis occurs when it has very large neighboring cities/LMAs. These outcomes confirm the existence of general urban agglomeration economies in Chile, as recently observed by Soto and Paredes (2016). Hence, there is a process of positive spatial autocorrelation in urban population, which tends to increase when they are surrounded by larger cities and LMAs, while they lose size when they are neighbored by smaller towns

Table 3 LISA transition matrices of Chilean city and LMA population, 1930–2002

Cities					LMAs				
State	HH	LH	LL	HL	State	HH	LH	LL	HL
HH	**0.838**	*0.103*	0.000	*0.059*	HH	**0.957**	0.000	0.044	0.000
LH	0.023	**0.959**	0.016	0.002	LH	0.000	**0.991**	0.000	0.009
LL	0.000	0.027	**0.970**	0.003	LL	0.006	0.025	**0.969**	0.000
HL	0.043	0.000	*0.057*	**0.901**	HL	0.000	*0.105*	0.000	**0.895**
π	*0.083*	*0.461*	*0.384*	*0.072*	π	*0.000*	*0.919*	*0.000*	*0.082*

Note: In italics, the ergodic distribution values and in bold, probabilities equal or greater 0.05

or LMAs. Spatial dependence may exist due to suburbanization processes (a population shift from central urban areas into suburbs), or urban shrinkage and counter-urbanization (a population loss by emigration in a relative short period of time).

3.4 Analysis of the co-evolution of the Chilean cities and their spatial neighbors

Table 3 presents the estimated probabilities for the joint transition of a city and its neighbors in the population distribution, that is, the LISA Markov chain corresponding to the Chilean cities and LMAs. For example, the 0.838 value corresponding to the transition HH–HH (first cell) means that the probability that a large city with large neighbors (HH) will remain in this state is 83.8%.

The test of independence of the two chains yields $\chi^2_9 = 31,225.42$ with $p<0.001$ (for the city set) and $\chi^2_9 = 498.88$ with $p<0.001$ (for the LMAs), allowing us to reject the null hypothesis and demonstrating co-dependence of the movement of a city and its neighbor cities. As in the previous cases, movements are more frequent within quadrants than between them, providing new evidence of high persistence in the system. Apart from the main diagonal, the most frequent movements are from the states HH–LH and from states HH to HL. First, the evolution from HH state (larger cities surrounded, on average, by larger cities) to LH (smaller cities surrounded, on average, by larger cities) may be interpreted as an urban shrinkage phenomenon, which is a quick population loss in response to deindustrialization or counter-urbanization processes, with people migrating massively to the city cores or rural areas, respectively. Second, movements from states HH progressing to HL could be evidence of population concentration processes (Sayas 2006). The steady-state distribution (π) shows higher probability values for LH and LL and lower values for HL and HH, indicating that only a few large cities (H state) and many small ones (L state) will exist in the long run. The results are not very significant for the LMAs, probably due to the problem of a small sample.

The Chilean urban system thus exhibits a clear and persistent pattern of agglomeration economies. The increase in the size of the intermediate cities found in the density plot demonstrates a regional concentration process. Population concentration thus occurs not only in the MR of Santiago, but also in the main important cities at regional level. This finding aligns with evidence for Latin American Countries such as Mexico (Pimentel 2000)

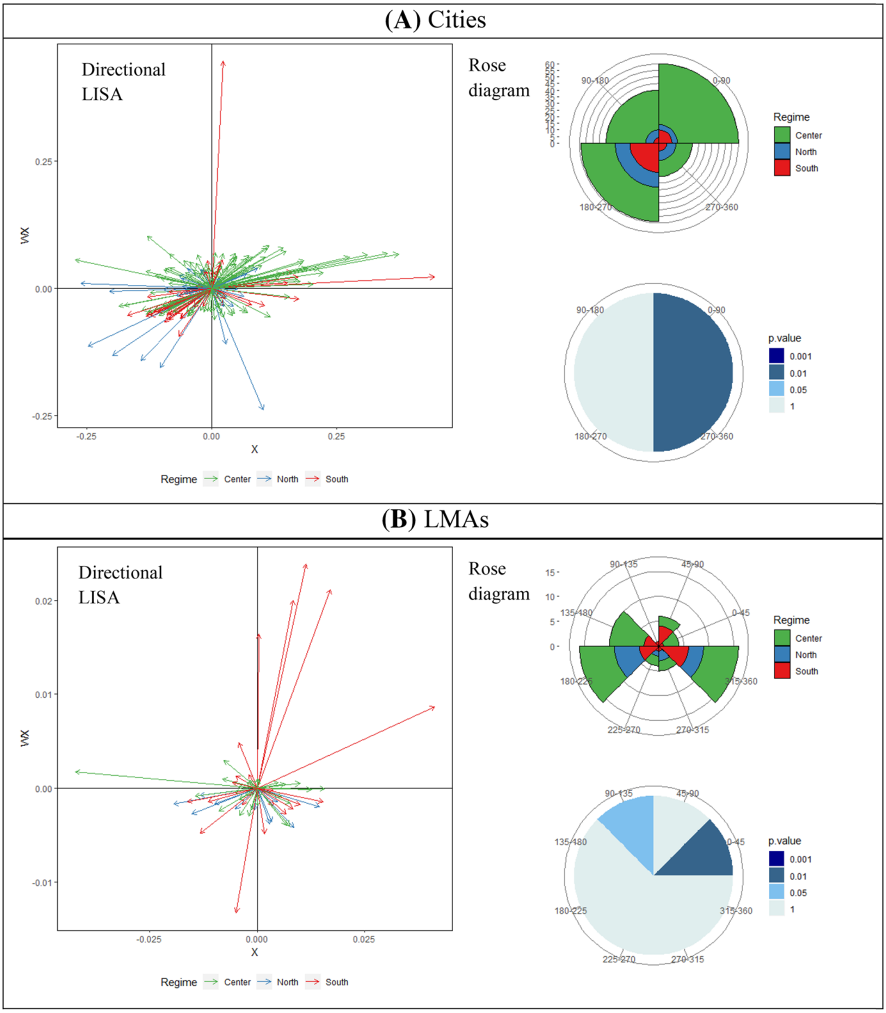

Fig. 5 Standardized directional LISA and rose diagrams of Chilean cities and LMAs by spatial regimes, 1930–2002

and Brazil (Baeninger 1997), where regional growth trends spread beyond large metropolitan areas but usually maintain a high degree of demographic concentration in large and medium-sized cities, especially in large metropolitan areas (Da Cunha 2003, 2013).

Additionally, the directional LISA approach provides a visual summary of the movements across the LISA Markov chain. This technique is also useful for identifying differences in the co-evolution of cities and neighbors across urban subsystems or spatial regimes, as is the case in Chile with the North, Central, and South urban subsystems. Figure 5 represents the standardized Directional LISA of these three Chilean urban spatial regimes. The movement vectors reflect relative changes in the LISA Markov chain between the first and last period of analysis (1930 and 2002, respectively). We have standardized the moves such that all arrows depart from the coordinate origin of the Moran scatterplot. For example, movements to the 'Southwest' part of the scatterplot indicate reduction in a city's size concurrent with reduction in its neighbors' size. Similarly, movements to the Northeast represent an increase in the size of both the city and its neighbors during the period.

To obtain a clearer view of the movement patterns, we build two rose diagrams, one representing the statistically significant co-movements of the directional LISA. The rose diagrams are divided into eight classes to visualize the usual HH, LH, LL and HL Moran scatterplot quadrants and the intermediate directions (e.g., a location in the 0°-to-45° corresponds to cities with higher population growth than its corresponding neighbors). We compute these graphs for the Chilean cities and LMAs by spatial regimes. Analysis of the inference rose diagram for the cities shows that only movements in the Northeast and Southeast areas of the Moran scatterplot are significant at 1% (Fig. 5a). If we consider this area only, the most frequent movement in the Center regime (green) is the 0°-to-90° direction in the scatterplot, indicating great growth in the different cities and their neighbors from 1930 to 2002—a clear process of population concentration in this regime. The same occurs in the Southern regime (red), though with much less intensity. The Northern regime's most significant movement occurs in the 225°-to-360° direction in the scatterplot, meaning that Northern cities experienced a predominant joint reduction in size with their neighbors (blue). This change could indicate urban migration from this regime to the others, mainly the Central Zone.

In the case of the LMAs, depicted in Fig. 5b, the only areas significant at 5% in the rose graph occur in the 0°-to-45° and 90°-to-135° directions in the Moran scatterplot. In these two areas, the only significant movements are in the Central regime, in the 0°-to-45° direction of the scatterplot, showing greater growth in the different LMAs than in their neighbors. This finding is consistent with the agglomeration forces already detected in the city group, though in this case the degree of spatial dependence is much less important. Hence, spatial dependence acts only significantly in the Center regime for the LMAs, in the same direction though with less intensity than for the city group, as shown previously.

3.5 Computation of spatial regime disparities in the co-evolution of the Chilean cities

We use the GIMA to measure the existence of disparities or inequalities in Chilean city-size from 1930 to 2002. Table 4 reports mobility and τ index decomposition, with their corresponding p values,[12] for the Chilean cities and LMAs in order to detect differences in behavior between the spatial regimes, using a regime weight matrix in the computation.

In the case of the cities, only two periods, 1952–1960 and 1970–1982, show statistically significant spatial indexes, as does the last period for the group of LMAs. This result demonstrates differences in ranking mobility in space that coincide with structural change in the 1970s in Chile. During this period, ranking mobility between neighboring regimes (M_W) was higher than mobility between non-neighboring

[12] In this paper, p values are computed with a 1000-replication process.

Table 4 Spatial Kendall indexes for Chilean city and LMAs population 1930–2002

Period	Cities					LMAs				
	M_W	M_{noW}	τ_W	τ_{noW}	p value	M_W	M_{noW}	τ_W	τ_{noW}	p value
1930–1940	0.069	0.076	0.862	0.848	0.130	0.032	0.041	0.934	0.916	0.247
1940–1952	0.079	0.080	0.842	0.841	0.446	0.044	0.057	0.911	0.885	0.180
1952–1960	**0.074**	**0.067**	**0.851**	**0.867**	**0.043**	0.034	0.057	0.931	0.884	0.040
1960–1970	0.063	0.069	0.874	0.862	0.127	0.027	0.037	0.945	0.924	0.202
1970–1982	**0.061**	**0.053**	**0.878**	**0.894**	**0.019**	**0.017**	**0.035**	**0.965**	**0.929**	**0.028**
1982–1992	0.053	0.054	0.894	0.892	0.394	0.018	0.025	0.962	0.949	0.266
1992–2002	0.056	0.055	0.888	0.889	0.445	0.018	0.028	0.962	0.943	0.173

Note: Subscripts 'W' and 'noW' indicate the neighboring and non-neighboring spatial regimes. In bold, periods showing significant spatial indexes

regimes (M_{noW}). Due to Chile's peculiar geography and the location of the spatial regimes, the neighboring regimes are North-Center and South-Center, and the non-neighboring regimes North–South.

This result implies less concordant movement of cities in the interaction between the Northern and Central regimes and between the Southern and Central ones, with respect to the interaction between the Northern and Southern regimes. A change in the rank of the cities in the Northern and Southern regimes relative to the Central regime is thus more likely to occur from the peripheral regimes to the Central regime. The asynchronous evolution of the cities in the peripheral regimes with respect to the Central regime is consistent with the reinforcement of the Chilean Central Zone due to abandonment of the ISI policy and trade liberalization. In effect, these changes in the national policy generated a strong migration trend from rural areas in the Northern and Southern regimes to intermediate and large cities in the Center, with production structures based on comparative advantage goods (Geisse 1977; Geisse and Valdivia 1978; Escolano Utrilla et al. 2007). They produced strong agglomeration forces around the Metropolitan Region of Santiago and the rest of the Central Zone.

3.6 Analysis of the cohesion among the Chilean urban spatial regimes

Table 5 reports the rank decomposition index $\Theta_{t_1-t_0}$ and its p value. As in the previous method, we used the three Chilean urban regimes for cities and LMAs. During the periods 1930–1940 and 1982–1992 (1970–2002, for the group of LMAs), the Θ index is statistically significant and takes the highest values, which implies a high degree of 'cohesion' within the spatial regimes.

In this context, cohesion implies a stronger migration process among regimes. One drawback of this index, however, is that it cannot provide information about the direction of the migration flow. As stated before, specific events in Chile's history suggest that the Central Zone was the main beneficiary of this transfer. In the period 1930–1940, the devastating consequences of the Great Depression produced bankruptcy, unemployment and rural–urban migration in Chile that benefited Central Zone growth (Geisse 1977; Geisse and Valdivia 1978; De Mattos 1999; Rodríguez and Rowe 2018). From the 1980s onward, the Central Zone underwent a selective

Table 5 Rank decomposition index Θ for Chilean urban population 1930–2002

Period	θ	p value	Period	θ	p value
1930–1940	**0.508**	**0.007**	**1930–1940**	**0.264**	**0.045**
1940–1952	0.321	0.159	1940–1952	0.163	0.201
1952–1960	0.415	0.110	1952–1960	0.169	0.218
1960–1970	0.269	0.297	1960–1970	0.130	0.382
1970–1982	0.068	0.145	**1970–1982**	**0.326**	**0.004**
1982–1992	**0.555**	**0.004**	**1982–1992**	**0.301**	**0.007**
1992–2002	0.102	0.271	**1992–2002**	**0.213**	**0.049**

Note: In bold, periods showing significant spatial indexes

agglomeration process based on higher income and better amenities, which attracted migrants with high human capital to its cities and LMAs (Cambiaso et al. 2001).

4 Conclusion

In this paper, we use a set of novel tools to enable evaluation of the influence of spatial proximity among human settlements on the evolution of the cities to detect regional differences and interactions in their spatiotemporal dynamics. Some of the statistical techniques employed in this paper have not previously applied to urban studies and are revealed to be very useful for detecting spatial dependence and spatial regimes in the evolution of an urban system. Intelligent combination of these tools, which can be computed with *estdaR* (an R package), detects different trends and spatial clusters in the development of Chilean cities over the period 1930–2002, focusing specifically on how spatial proximity affects relative sizes and rankings.

We have organized this procedure as a six-step method. The first step consists of the usual characterization of the cross-sectional distribution of the urban areas by means of standard statistical analysis and nonparametric estimations of density functions for a set of significant years. In the second and third steps, the growth process is modeled as a first-order stationary Markov chain to evaluate the effect of global and local spatial autocorrelation on the transition probabilities with a set of indices based on the spatial version of the standard Markov chain. The fourth, fifth, and sixth steps perform in-depth analysis to detect the existence and interaction of spatial regimes in the movement direction and ranking mobility of urban distribution. We use the LISA transition matrix and the directional LISA approach to capture the co-movement directions of cities and neighbors across the Moran scatterplot quadrants. We also study the existence of spatial regime differences in the ranking mobility of the Chilean urban distribution using the GIMA. Finally, we determine the ranking decomposition of city size by spatial regimes.

Application of this method to the Chilean cities and LMAs generates some interesting conclusions. First, initial exploratory data analysis shows three spatial regime clusters defined by the Northern Zone (regions I–IV and XV), Central Zone (regions V to VIII and the Metropolitan Region) and Southern Zone (regions IX–XII and XIV). These results reflect a general convergence process in the Chilean urban population at a national level, which is mainly determined by the cities of the Central Zone, which experienced periurban growth starting in 1970, conditioned by the urban core of Santiago. Additionally, the significant socioeconomic transformations in this regime operating in Chile during this period promoted the emergence of intermediate cities, particularly associated with the exploitation of natural resources (Escolano Utrilla 2012).

The cities in the Northern and Southern regimes remain more or less polarized in two groups of diverging cities, however, with a club of consolidated larger regional cities experiencing agglomeration economies and population concentration. The long-run viability of these regional capitals—in terms of their efficiency, workability, and livability—will depend on policymaking infrastructures capable of carrying out corrective intervention and regulation programs (Scott 2008).

Analysis of city-size distributional mobility demonstrates the important role played by path dependence in urban dynamics with very low inter-class mobility. Inter-class mobility of the urban system is concentrated in medium-size cities. Small and small-to-medium-size cities have a higher propensity to grow, moving upwards in the distribution, whereas medium and medium-to-large cities are more likely to lose population and move downwards, confirming the overall convergence trend of Chilean cities. Detailed examination shows evidence of a tendency toward stratification, however, due to the practical immobility—inside the population distribution—of the smallest and larges cities, which is consistent with the results for the kernel functions shown in Fig. 3.

The estimation of the spatial Markov and LISA Markov matrices enables us to conclude that the probability a city will grow increases with its neighbors' size, while large cities surrounded by smaller towns hardly experience any change in population. Spatial proximity thus matters in the urban system, usually by promoting a clear and persistent pattern of agglomeration economies, as is common in Latin American and in other less-advanced countries (Duranton 2016). The increase in the size of the intermediate cities found in the density plot demonstrates a regional concentration process. Population concentration thus occurs not only in the Central Zone, but also in the main important cities at regional level, which have been the main reservoirs of migration flows from peripheral rural areas and foreign migrants since 1970 until now.

In its urban policy, Chile must recognize three different city clubs: the Central cluster, dominated by the city of Santiago; and the other regional Northern and Southern cities, which share their own characteristics and dynamics. A more balanced regional urban system in the Northern and Southern regimes, growing at the expense of the nearby rural areas, will coexist with a highly concentrated Central regime, in which Santiago's population spills over into its satellites, moving toward (still distant) future convergence.

Funding The funding was provided by Ministerio de Economía, Industria y Competitividad, Gobierno de España (Grand No. ECO2015-65758-P).

Appendix 1

See Table 6.

Table 6 List of the Chilean cities and labor market areas (LMAs) used in this paper

Regime	Region	Labor market area	City
North	XV Arica and Parinacota	Arica	Arica
	I Tarapacá	Iquique	Iquique
			Pozo Almonte
	II Antofagasta	Antofagasta	Antofagasta
			Mejillones
			Taltal
		Calama	Calama
			Chuquicamata
		Tocopilla	María Elena
			Tocopilla
	III Atacama	Copiapo	Caldera
			Copiapó
			Tierra Amarilla
		Chanaral	Chañaral
			Diego de Almagro
		Vallenar	Huasco
			Vallenar
	IV Coquimbo	La Serena	Andacollo
			Coquimbo
			La Serena
			Vicuña
		Illapel	Illapel
			Los Vilos
			Salamanca
		Ovalle	Combarbalá
			El Palqui
			Monte Patria
			Ovalle
Center	V Valparaíso	Valparaiso	Casablanca
			Concón
			Limache
			Olmué
			Placilla de Peñuelas
			Quilpué
			Quintero
			Valparaíso
			Ventanas
			Viña del Mar
			Villa Alemana
		La Ligua	La Ligua
		Cabildo	Cabildo
		Quillota	El Melón

Table 6 (continued)

Regime	Region	Labor market area	City
			Hijuelas
			La Calera
			La Cruz
			Nogales
			Quillota
		San Antonio	Algarrobo
			Cartagena
			San Antonio
		Los Andes	Catemu
			Los Andes
			Llaillay
			Putaendo
			Rinconada
			San Esteban
			San Felipe
			Santa María
	RM metropolitan region	Santiago	Alto Jahuel
			Bajo de San Agustín
			Batuco
			Buin
			Colina
			Curacaví
			El Monte
			Hospital
			Isla de Maipo
			La Islita
			Lampa
			Maipú
			Paine
			Peñaflor
			Puente Alto
			Quilicura
			San Bernardo
			San José de Maipo
			Santiago
			Talagante
			Tiltil
		Melipilla	Melipilla
	O' Higgins	Rancagua	Codegua
			Doñihue
			Graneros
			Gultro

Table 6 (continued)

Regime	Region	Labor market area	City
			Lo Miranda
			Machalí
			Quinta de Tilcoco
			Rancagua
			Requínoa
			San Francisco de Mostazal
		Las Cabras	Las Cabras
			Peumo
		Pichidegua	Rengo
			San Vicente de Tagua Tagua
		Pichilemu	Pichilemu
		San Fernando	Chimbarongo
			San Fernando
		La Estrella	Peralillo
		Lolol	Nancagua
			Santa Cruz
	VII Maule	Talca	San Clemente
		Talca	Talca
		Constitucion	Constitución
		Curepto	Hualañé
		Cauquenes	Cauquenes
		Río Claro	Curicó
		Río Claro	Molina
		Río Claro	Teno
		Linares	Linares
		Linares	Longaví
		Parral	Parral
		San Javier	San Javier
		San Javier	Villa Alegre
	VIII Biobío	Concepción	Concepción
			Coronel
			Hualqui
			Lota
			Penco
			Santa Juana
			Talcahuano
			Tomé
		Lebu	Arauco
			Curanilahue
			Lebu
			Los Alamos
		Cañete	Cañete

Table 6 (continued)

Regime	Region	Labor market area	City
		Los Angeles	La Laja
			Los Ángeles
			Mulchén
			Nacimiento
			Santa Bárbara
		Yumbel	Bulnes
			Cabrero
			Coihueco
			Chillán
			Monte Águila
			Quillón
			San Carlos
			Yumbel
		Cobquecura	Coelemu
			Quirihue
		Tucapel	Huépil
			Yungay
South	IX Araucanía	Temuco	Cunco
			Freire
			Gorbea
			Lautaro
			Nueva Imperial
			Pitrufquén
			Temuco
		Carahue	Carahue
		Curarrehue	Loncoche
			Pucón
			Villarrica
		Angol	Angol
			Collipulli
			Renaico
		Curacautín	Curacautín
		Galvarino	Purén
			Traiguén
		Ercilla	Victoria
	XIV Los Ríos	Curaco de Velez	Paillaco
		Valdivia	Los Lagos
			Valdivia
		Lanco	Lanco
			San José de la Mariquina
	X Los Lagos	Puerto Montt	Calbuco
			Fresia

Table 6 (continued)

Regime	Region	Labor market area	City
			Llanquihue
			Puerto Montt
			Puerto Varas
		Castro	Castro
		Ancud	Ancud
		Quellon	Quellón
		Osorno	Osorno
			Río Negro
		Frutillar	Frutillar
			Purranque
		Chaitén	Chaitén
		Futrono	Futrono
			La Unión
			Rio Bueno
	XI Aysén	Coihaique	Coihaique
			Puerto Aysén
	XII Magallanes	Punta Arenas	Punta Arenas
		Natales	Puerto Natales

Appendix 2

See Fig. 6.

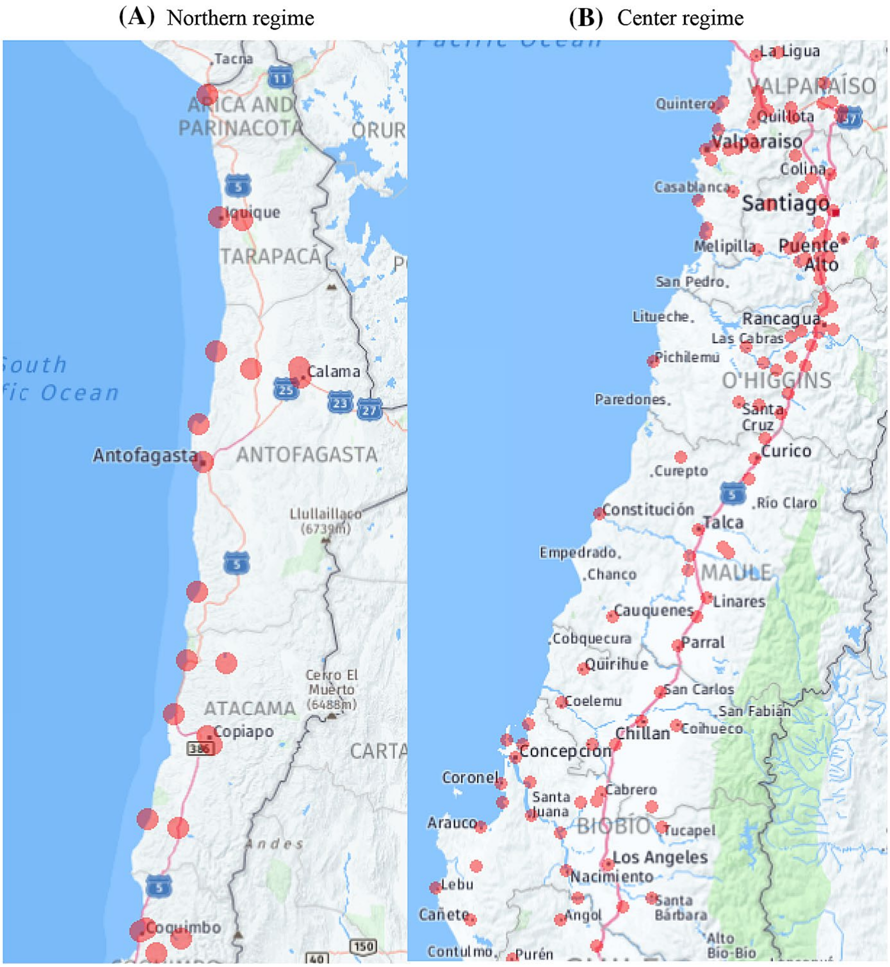

Fig. 6 Spatial regimes of urban subsystems in Chile

Southern regime

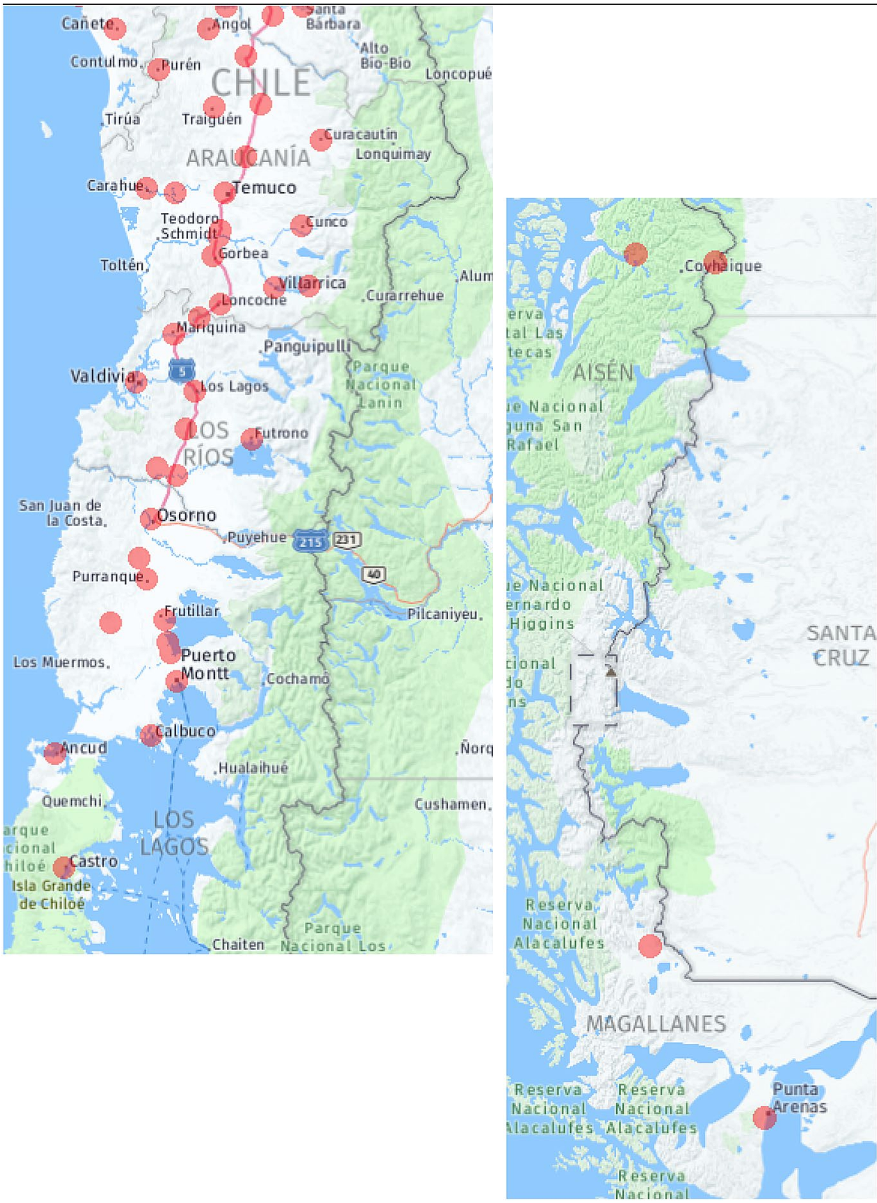

Fig. 6 (continued)

References

Agovino M (2014) Do "good neighbors" enhance regional performances in including disabled people in the labor market? A spatial Markov chain approach. Ann Reg Sci 53:93–121

Agovino M, Crociata A, Sacco PL (2016) Location matters for pro-environmental behavior: a spatial Markov Chains approach to proximity effects in differentiated waste collection. Ann Reg Sci 56:295–315

Anderson G, Ge Y (2005) The size distribution of Chinese cities. Reg Sci Urban Econ 35:756–776

Anselin L (1995) Local Indicators of Spatial Association—LISA. Geogr Anal 27:93–115

Anselin L, Rey S (2014) Modern spatial econometrics in practice: a guide to GeoDa, GeoDaSpace and PySAL. GeoDa Press, Chicago

Antrop M (2004) Landscape change and the urbanization process in Europe. Landsc Urban Plan 67:9–26

Baeninger R (1997) Redistribución espacial de la población:características y tendencias del caso brasileño. Notas de población

Bickenbach F, Bode E (2003) Evaluating the Markov Property in studies of economic convergence. Int Reg Sci Rev 26:363–392

Bustos Validiva H (2013) Historia de Isla de Maipo, Municipalidad de Isla de Maipo. Municipalidad de Isla de Maipo

Cambiaso PS, Alonso MC, Claro CF (2001) Migraciones internas hacia la Región Metropolitana de Santiago de Chile:una comparación con planteamientos teóricos. Investig Geogr 35:1

Casado-Díaz JM, Martínez-Bernabéu L, Rowe F (2017) An evolutionary approach to the delimitation of labour market areas:an empirical application for Chile. Spat Econ Anal 12(4):379–403

Da Cunha JMP (2003) Urbanización, redistribución espacial de la población y transformaciones socioeconómicas en América Latina. United Nations Publications, New York

Da Cunha JMP (2013) Questions and challenges in studies on Latin-American cities. Chapters 2013:127–152

Davis JC, Henderson JV (2003) Evidence on the political economy of the urbanization process. J Urban Econ 53:98–125

De Mattos CA (1999) Santiago de Chile, globalización y expansión metropolitana:lo que existía sigue existiendo. EURE (Santiago) 25:29–56

Delmelle E, Thill JC, Wang C (2016) Spatial dynamics of urban neighborhood quality of life. Ann Reg Sci 56:687–705

Desmet K, Henderson JV (2015) The geography of development within countries. In: Duranton G, Henderson JV, Strange W (eds) Handbook of regional and urban economics. Elsevier, Amsterdam, pp 1457–1517

Duranton G (2016) Determinants of city growth in Colombia. Pap Reg Sci 95:101–131

Escolano Utrilla S (2012) Dinámica reciente del sistema urbano chileno 1970–2002: integración a escala nacional. Bola Asoc Geógr Esp 59:129–150

Escolano Utrilla S, Ortiz Véliz J (2004) Cambios de la configuración urbana y «sintaxis del espacio» en ciudades intermedias; el caso de La Serena (Chile). Estud Geogr 65:297–320

Escolano Utrilla S, Ortiz Véliz J, Moreno Mora R (2007) Globalización y cambios funcionales recientes en las ciudades del sistema urbano chileno. Cuad Geogr 41:33–60

Geisse G (1977) Origen y evolución del sistema urbano nacional. EURE (Santiago) 5:37–46

Geisse G, Valdivia M (1978) Urbanización e industrialización en Chile. EURE (Santiago) 5:15

Gregory T, Patuelli R (2015) Demographic ageing and the polarization of regions: an exploratory space–time analysis. Environ Plan A Econ Space 47:1192–1210

Grinstead C, Snell JL (1997) Introduction to probability. American Mathematical Society, Providence

Gutiérrez MLS, Rey SJ (2013) Space-time income distribution dynamics in Mexico. Ann GIS 19:195–207

Hamilton JD (1994) Time series analysis. Princeton University Press, Princeton

Henderson JV (2005) Urbanization and growth. In: Aghion P, Durlauf S (eds), Handbook of economic growth, vol 1, no Part B, pp 1543–1591

Henderson JV, Shalizi Z, Venables AJ (2001) Geography and development. J Econ Geogr 1:81–105

Henríquez C, Azócar G, Romero H (2006) Monitoring and modeling the urban growth of two mid-sized Chilean cities. Habitat Int 30:945–964

Hoàng NH (2013) Toward an integrated ASEAN labor market. Prospects and challenges for CLMV countries. VNU J Econ Bus 29(5):34–42

Instituto Nacional de Estadísticas (2005) Chile:ciudades, pueblos, aldeas y caseríos, 2005. http://historico.ine.cl/canales/usuarios/cedoc_online/censos/pdf/censo_2002_publicado_junio_2005.pdf

Instituto Nacional de Estadísticas (2014) Auditoría Técnica a la base de datos del levantamiento censal año 2012

Ioannides Y, Overman H (2004) Spatial evolution of the US urban system. J Econ Geogr 4:131–156

Kane K, Tuccillo J, York AM et al (2014) A spatio-temporal view of historical growth in Phoenix, Arizona, USA. Landsc Urban Plan 121:70–80
Kang W, Rey S (2019) Measures of income mobility. PySAL Notbooks. http://pysal.org/notebooks/dynamics/giddy/Mobility_measures
Kendall MG (1962) Rank correlation methods. Charles Griffin & Company Limited, London
Lanaspa L, Pueyo F, Sanz F (2003) The evolution of Spanish urban structure during the twentieth century. Urban Stud 40:567–580
Le Gallo J (2004) Space-time analysis of GDP disparities among European regions: a Markov Chains approach. Int Reg Sci Rev 27:138–163
Le Gallo J, Chasco C (2008) Spatial analysis of urban growth in Spain, 1900–2001. Empir Econ 34:59–80
Mutlu S (1986) City-forming propensities in a central place hierarchy: application of Beckmann–Mcpherson model to the Turkish urban system. Ann Reg Sci 20(2):28–43
Nitsch V (2001) City growth in Europe. Duncker & Humblot, Berlin
Olave D (2005) El sistema urbano de Chile central. Desafíos actuales y medio ambientales. Scr Nova 194:69
Paci R, Usai S (2008) Agglomeration economies, spatial dependence and local industry growth. Rev d'Écon Ind 123(3):87–109
Parr JB (2012) Spatial-structure differences between urban and regional systems. Ann Reg Sci 49:293–303
Pimentel M (2000) La reestructuración de los espacios nacionales en los inicios del siglo XXI: continuidad y cambio en la distribución espacial de la población mexicana. Santiago de Chile, CELADE, mimeo
Puertas OL, Henríquez C, Meza FJ (2014) Assessing spatial dynamics of urban growth using an integrated land use model. Application in Santiago Metropolitan Area, 2010–2045. Land Use Policy 38:415–425
Quah DT (1996) Empirics for economic growth and convergence. Eur Econ Rev 40:1353–1375
Rey SJ (2001) Spatial empirics for economic growth and convergence. Geogr Anal 33:195–214
Rey S (2004) Spatial analysis of regional income inequality. In: Goodchild MF, Janelle DG (eds) Spatially integrated social science. Oxford University Press, Oxford, pp 280–299
Rey S (2015) Python Spatial Analysis Library (PySAL): an update and illustration. In: Geocomputation: a practical primer, SAGE, London, pp 233–254
Rey SJ (2016) Space–time patterns of rank concordance: local indicators of mobility association with application to spatial income inequality dynamics. Ann Am Assoc Geogr 106:788–803
Rey SJ, Janikas MV (2006) STARS: space–time analysis of regional systems. Geogr Anal 38:67–86
Rey SJ, Murray AT, Anselin L (2011) Visualizing regional income distribution dynamics. Lett Spat Resour Sci 4:81–90
Rey SJ, Mack EA, Koschinsky J (2012) Exploratory space–time analysis of burglary patterns. J Quant Criminol 28:509–531
Rey SJ, Kang W, Wolf L (2016) The properties of tests for spatial effects in discrete Markov chain models of regional income distribution dynamics. J Geogr Syst 18:377–398
Rodríguez J (2007) United Nations expert group meeting on population distribution, urbanization, internal migration and development. United Nations Secretariat New York, pp 21–23
Rodríguez J, Rowe F (2018) How is internal migration reshaping metropolitan populations in Latin America? A new method and new evidence. Popul Stud 72(2):253–273
Rodríguez J, González D, Ojeda M et al (2009) El sistema de ciudades chileno en la segunda mitad del siglo XX: entre la suburbanización y la desconcentración (The Chilean City System during the Second Half of the 20th Century: between sub-urbanization and decentralization). Estud Demogr Urbanos 24:7–48
Rowe F (2017) The Chilean internal Migration (CHIM) database: a temporally consistent spatial data framework for the analysis of human mobility. REGION 4(3):R1–R6
Rozenblat C, Pumain D (2018) Metropolization and polycentrism in the European urban system. In: Rozenblat S, Rozenblat S (eds) International and transnational perspectives on urban systems. Springer, Berlin, pp 117–142
Santiago CM, Raggi JPF, Erices LV (2016) Urban growth trends in midsize Chilean cities: the case of Temuco. urbe. Rev Bras Gestão Urbana 8:375–389
Sayas JP (2006) Urban sprawl in the periurban coastal zones of Athens. Epitheor Koinonikon Ereun 120(121):71–104

Schmidheiny K, Suedekum J (2015) The pan-European population distribution across consistently defined functional urban areas. Econ Lett 133:10–13

Scott A (2008) Inside the city: on urbanisation, public policy and planning. Urban Stud 45(4):755–772

Soo KT (2014) Zipf, Gibrat and geography: evidence from China, India and Brazil. Pap Reg Sci 93(1):159–181

Soto J, Paredes D (2016) Cities, wages, and the urban hierarchy. J Reg Sci 56:596–614

Venables AJ (2005) Spatial disparities in developing countries: cities, regions, and international trade. J Econ Geogr 5:3–21

Willis E, Garman CCB, Haggard S (1999) The politics of decentralization in Latin America. Latin Am Res Rev 34(1):7–56

Wu JX, He LY (2017) How do Chinese cities grow? A distribution dynamics approach. Phys A Stat Mech Appl 470(15):105–118

Xiufang R, Zhongwu Z, Yajie S et al (2015) Temporal and spatial evolution of the cities along China section of the. J Desert Res 35:248–252

Xu Z, Zhu N (2009) City size distribution in China: Are large cities dominant? Urban Stud 46:2159–2185

Ye X, Xie Y (2012) Re-examination of Zipf's law and urban dynamic in China: a regional approach. Ann Reg Sci 49:135–156