doi:10.1111/pirs.12088

spMorph: An exploratory space-time analysis tool for describing processes of spatial redistribution*

Juan C. Duque[1], **Xinyue Ye**[2], **David C. Folch**[3]

[1] Research in Spatial Economics (RiSE), School of Economics and Finance, Universidad EAFIT, Carrera 49 7 Sur – 50. 050022 Medellín, Colombia (email: jduquec1@eafit.edu.co)
[2] Department of Geography, Kent State University, Kent, Ohio, USA (email: xye5@kent.edu)
[3] Institute of Behavioral Science, University of Colorado at Boulder, Boulder, Colorado, USA (email: david.folch@colorado.edu)

Received: 22 January 2013/ Accepted: 1 October 2013

Abstract. This paper introduces an exploratory space-time analysis tool for determining the two components of a spatial redistribution process: (*i*) the shock, which is the moment that triggers a spatial redistribution process; for example, a new policy, a war, an earthquake, etc.; and (*ii*) the duration of the regime fade, which is the time between the shock and the moment in which a new regime emerges as a better representation of the spatial distribution of the attribute. Two examples are provided: the first uses China's provincial *per capita* GDP between 1978 and 2008, and the second uses state level housing price and unemployment rate data for the US between 2002 and 2012.

JEL classification: C23, R12

Key words: Analytical regions, spatial clustering, space-time

1 Introduction

Transitional countries are increasingly significant players in the global economy. Their unprecedented growth and multifaceted disparities have generated many fascinating issues for scholarly research, leading to a long debate about their unique characteristics and varied trajectories of development in the context of globalization. Dramatic economic growth with the partial reforms and their associated market distortions in transitional countries poses a significant challenge to existing conceptualizations of regional science, policy and practice in the Western world. As Wei and Ye (2009, pp. 59–60) argue, 'while some (scholars) maintain that transitional countries have achieved rapid economic growth without increasing regional inequality, others

* The authors thank the Cyberinfrastructure Service for High Performance Computing, 'Apolo', at EAFIT University for letting us run our computational experiments on their supercomputer. We also thank Alejandro Betancourt and Sergio Botero for their assistance in coding and editing, and the anonymous reviewers for their insightful suggestions. The usual disclaimer applies.

argue that the transition exhibits multiple forms, and is characterized by evolution, path dependency, partial reform of socialist institutions, and intensified social and spatial inequalities'. From both methodological and empirical perspectives, the above challenge has mainly been addressed by using aggregated data over time, which suffers from the ecological fallacy (EF) in a time series context (Duque et al. 2006). Using Spanish regional unemployment data, Duque et al. (2006) show that EF effects are not only observed at the cross-section level, but also in a time series framework. The empirical evidence obtained shows that analytical regions, constructed through a regionalization algorithm, are less susceptible to time effects relative to normative regional configurations, and normative configurations are an improvement over random ones.

At the same time, theoretical work on regional economic development has indicated a renewed and growing awareness of its spatial representation (Rey and Ye 2010). Many interesting constructs such as poverty traps and convergence clubs have been suggested and hotly debated (Rey and Janikas 2005; Ye and Rey 2013). Novel methods in clustering, regionalization and shape analysis have been developed, which present opportunities for both theoretical and empirical approaches to address these debates (Duque et al. 2007). Regionalization methods can also be used to spatially represent the dynamic nature of economic processes (Crone 2005; Duque et al. 2006). This paper suggests a new tool for exploratory space-time analysis, called spMorph, to help understand the changes in spatial distributions over time.

The causes of spatial redistribution are myriad, and in this paper we present two case studies that demonstrate spMorph's general approach. The first case study considers regional policy and practices in the context of globalization and transition in China; the second investigates the varied impacts of the recent economic crisis in the US. We argue that research on regional growth fortunes and the redistribution of wealth is embedded in complex economic and geographic processes and leads to multiple trajectories of development. China's mounting importance in the world economic system has been intensely debated in terms of its spectacular and uneven growth. The speed and inequality of economic restructuring and development in favour of specific regions is truly unprecedented and merits a careful examination and assessment. For instance, scholars notice that the traditional leading industrial provinces under Mao have been challenged by a rising wealthy cluster of coastal provinces. Development is also uneven within provinces, coastal regions as a whole have moved ahead of the poorer interior regions (Ye and Wei 2005; Fan and Sun 2008). In the US a housing price bubble affected the majority of US households, as they saw home values, and in turn their wealth, rapidly rise and then crash in an unprecedented fashion. However, the intensity of the housing bubble varied from place to place with many coastal and Sun Belt cities being more acutely affected than many established inland parts of the country. This paper highlights the importance of adopting new methodologies for describing processes of spatial redistribution. It is an important approach given the increasing trend of regional inequality and divergence in both transitional and developing countries such as China, and in developed countries such as the US.

To address dramatic regional development, wealth redistribution, and economic crisis, both spatial and temporal perspectives are crucial. The rapid growth in spatial economic analysis is increasingly seen as attributable to the availability of panel regional economic datasets (Ye and Rey 2013). By contrast, regional inequality studies have been slower to integrate space and time (Rey and Janikas 2005). Although spatially constrained clustering (Murtagh 1985; Gordon 1996; Duque et al. 2007) and time series clustering (Kakizawa et al. 1998; Galeano and Peña 2000; Fruhwirth-Schnatter and Kaufmann 2004; Pattarin et al. 2004; Alonso et al. 2006) have been widely studied in the literature, from the best of our knowledge, little research has been done in the area of space-time clustering on a real data (Crone 2005; Duque et al. 2006), in which a set of spatially contiguous areas is aggregated into regions such that the areas within a region show similar behaviour throughout a period of time.[1] In this paper we aim to go further with the incorporation of a scale in time, by opening the possibility of allowing for the division of the study period into subperiods that generate different, but related, spatial clusters. As there exists the concept of scale in space, which helps to decide the proper number of clusters for studying a given variable, we introduce a scale in time which indicates the proper number of sub-periods in which a period needs to be divided in order to get a better understanding of the spatio-temporal dynamics of a variable. In this paper, we demonstrate this novel approach in examining the spatio-temporal dynamics of *per capita* GDP in China, and housing prices and unemployment rates in the US.

Following the introduction, Section 2 presents the motivation and components of the proposed concepts. Section 3 then describes the methods. Section 4 discusses the case study of China, followed by the US case study in Section 5. Section 6 concludes with a summary of key findings and future research directions.

2 Elements and illustration of the concepts

spMorph is a tool that can be applied to any space-time data panel, where the attribute under analysis is measured on multiple spatial units across multiple time periods; our empirical example considers China's annual GDP *per capita* for 33 provinces between 1978 and 2008, and home price index and unemployment rate for the 48 lower US states and Washington, DC. One challenge when working with panel data is that there could be one or more meaningful breaks in the spatial pattern during the study timeframe. Not taking these breaks into consideration could muddy an empirical analysis – essentially smoothing the different spatial regimes into a single result. In this context, a 'spatial regime' is a representation that aggregates spatially contiguous areas into regions, such that the areas in each region have similar levels of the attribute under analysis and similar temporal dynamics of that attribute. spMorph seeks to summarize the spatio-temporal redistribution of the attribute by identifying both a small set of spatial regimes and the moments when the transition from one regime to another occurs.[2]

We decompose the transition between regimes into two concepts, the 'shock' and the duration of what we call 'the regime fade'. The shock is the moment that triggers a spatial redistribution process; for example, the implementation of a new policy, a war, an earthquake, etc. The duration of the regime fade is the time between the shock and the moment in which a new regime emerges as a better representation of the spatial distribution of the attribute. In other words, it is the time during which the current regime 'morphs' into the new one. The connection between these two concepts is intuitive; the type of shock determines, in part, the duration of the regime fade. One example of regime fade duration is a school board that implements a higher property tax with the goal of improving student outcomes. Over time, residents will move in and out of the school district in response to the new tradeoff between taxes and student outcomes; since residential moves are costly and time-consuming, the regime fade duration will likely be relatively long. In contrast, a massive flood could hit a region wiping out a substantial share of the housing stock in a matter of hours. Both examples imply the rise of a new spatial regime after the shock, but the duration of the regime fade in the first case is longer than in the second case.

[1] We note that there has been a substantial amount of research into space-time clustering of point data, especially in the field of epidemiology. See Kulldorff and Hjalmars (1999) for a review.

[2] In this paper there exist three types of spatial units: (i) area, which is the smallest spatial unit; (ii) region, which is the aggregation of spatially contiguous areas (an area can be a region by itself); and (iii) spatial regime is a collection of non-overlapping regions that cover the whole study area.

Duque et al. (2006) highlight the importance of designing spatial regimes (or what they call analytical regions) that are stable over time; that is, a spatial regime that is capable of representing the spatial distribution of a variable for several time periods. To design these regimes, the authors apply spatial clustering algorithms capable of dealing with time series. These algorithms take one attribute observed at multiple points in time as an input, and aggregate the areas into regions such that the statistical distance between the time series of the areas assigned to the same region is minimized. Once the spatial regime is calculated, the authors use the decomposed version of Theil's inequality index (Theil 1967, 1972) to calculate the share of total inequality accounted for by the within groups component (Iw/I).[3] This ratio is calculated for every period to evaluate the capacity of the regime to aggregate the areas into homogeneous regions in terms of the studied attribute. A good spatial regime keeps the ratio Iw/I low every period, which implies high similarity between the areas assigned to the same region. In this paper we use these two elements, spatial clustering algorithms and the ratio Iw/I, to design and evaluate the performance of the spatial regimes. See Shorrocks and Wan (2005) for a rigorous analysis of the Theil index decomposition and its applicability to other spatial inequality research.

The concepts introduced in this section are illustrated in Figure 1 using a regular lattice. The study period covers year 0 to year T; in year S a shock occurs that triggers a spatial redistribution of the variable. Spatial regime A is the result from solving the spatially constrained clustering problem using data from years 0 to S; and spatial regime B is the solution that results using data from years S + 1 to T. Both spatial regimes aggregate 63 areas into three regions. Instead of the typical regionalization approach that uses multiple attributes from a single point in time, we use multiple years' worth of data for a single attribute. Note that regimes A and B are different, which reflects a spatial redistribution of the variable. The ratio Iw/I is calculated for each regime and for each year. As expected, regime A is a better representation of the spatio-temporal distribution of the attribute between years 0 and S. When the shock occurs, regime A starts losing its capacity to represent the spatial distribution of the attribute, which causes a gradual increment of the ratio Iw/I. Two years after the shock, regime B emerges as the new regime generating the lowest ratio Iw/I every year after. The two years after the shock (S) is what we call the duration

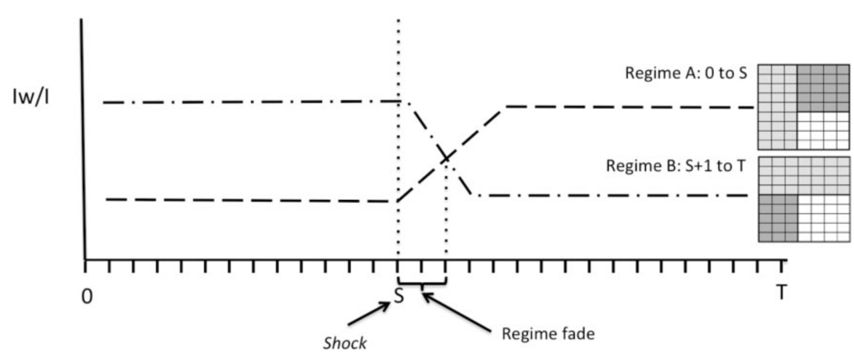

Fig. 1. Illustration of the spMorph approach

[3] The Theil inequality index is defined as, $I = I_b + I_W = \sum_{k=1}^{p} s_k \log\left(\frac{n}{n_k s_k}\right) + \sum_{k=1}^{p} s_k \sum_{i \in k} s_{i,k} \log\left(n_k s_{i,k}\right)$, where I_b is the inequality between regions and I_w is the weighted average of inequality within regions; n_k is the number of areas in region k; n is the total number of areas ($n = \sum_k n_k$) and p is the total number of regions; $s_k = \sum_{i \in k} Y_i / \sum_{i=1}^{n} Y_i$ is the share of attribute Y accounted for by region k; and $s_{i,k} = Y_{i,k} / \sum_{i=1}^{n_k} Y_{i,k}$ is the share of attribute Y in region k accounted by area i.

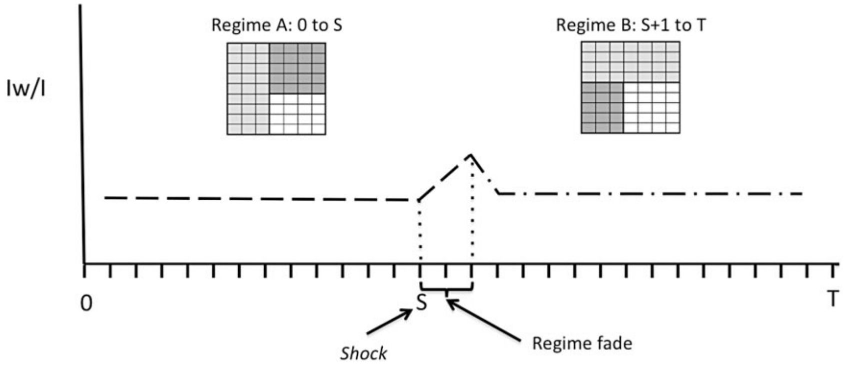

Fig. 2. The lower bound

of the regime fade. The regime fade effect means that although spatial regimes A and B are calculated using input data from periods 0 to S and S+1 to T respectively, regime A dominates from 0 to S+2 and regime B dominates from S+2 to T.

Figure 2 presents the concept of lower bound, which results from calculating, for each year, the minimum ratio Iw/I. The area below the lower bound is called the total lower bound (TLB) and it is used to compare solutions that place the shock in different years. The lower the TLB, the better the solution, because it minimizes the share of intraregional inequality over the entire study period.

All the concepts in this section were introduced using the simple case of one shock and one variable, but this procedure can be extended to more general cases. With two shocks, for example, there would be three periods: 0 to S1, S1 + 1 to S2 and S2 + 1 to T. For each period there is a spatial regime, and the three regimes represent the solution that gives the lowest TLB. The unique TLB shows the dominant period for each regime. An empirical version of the two shocks case is presented in Section 4, along with a two variable version in Section 5.

3 The algorithm

As commented above, the spatial redistribution of an attribute can be the result of more than one shock during the analysed period, and determining the moment of each shock can be a complicated task. Algorithm 1 presents the algorithmic solution for this case.[4] We call the algorithm spMorph, 'sp' for spatial and 'morph' representing the spatial transformation process that the attribute undergoes over time. In summary, spMorph takes the spatial distribution of an attribute at multiple points in time, and identifies the potential location of *s* shocks. From the algorithm's results the researcher will be able to determine (*i*) the moment when each shock occurred, (*ii*) the regime associated to each period, and (*iii*) the periods of dominance of each regime.

We illustrate the spMorph operations in Figure 3 using a stylized example. Let us suppose that we have a spatial panel Y of N = 100 areas and T = 4 years, and we want to run the spMorph algorithm for one shock (s = 1) and for spatial regimes with three regions (p = 3). Given the number of years and shocks, the resulting feasible subperiods are, $\mathbb{S}:\{1,2-4\}$, $\mathbb{S}:\{1-2,3-4\}$, $\mathbb{S}:\{1-3,4\}$; therefore, the values of $\boldsymbol{\Omega}$ are {2,3,4}. Based on these initial parameters, Figure 3 illustrates the main operations and the results of Algorithm 1.

spMorph requires the use of an algorithm for spatially constrained clustering. This algorithm designs the spatial regime for a given period of time. In this paper we apply a heuristic solution

[4] We also provide Python code in the Appendix.

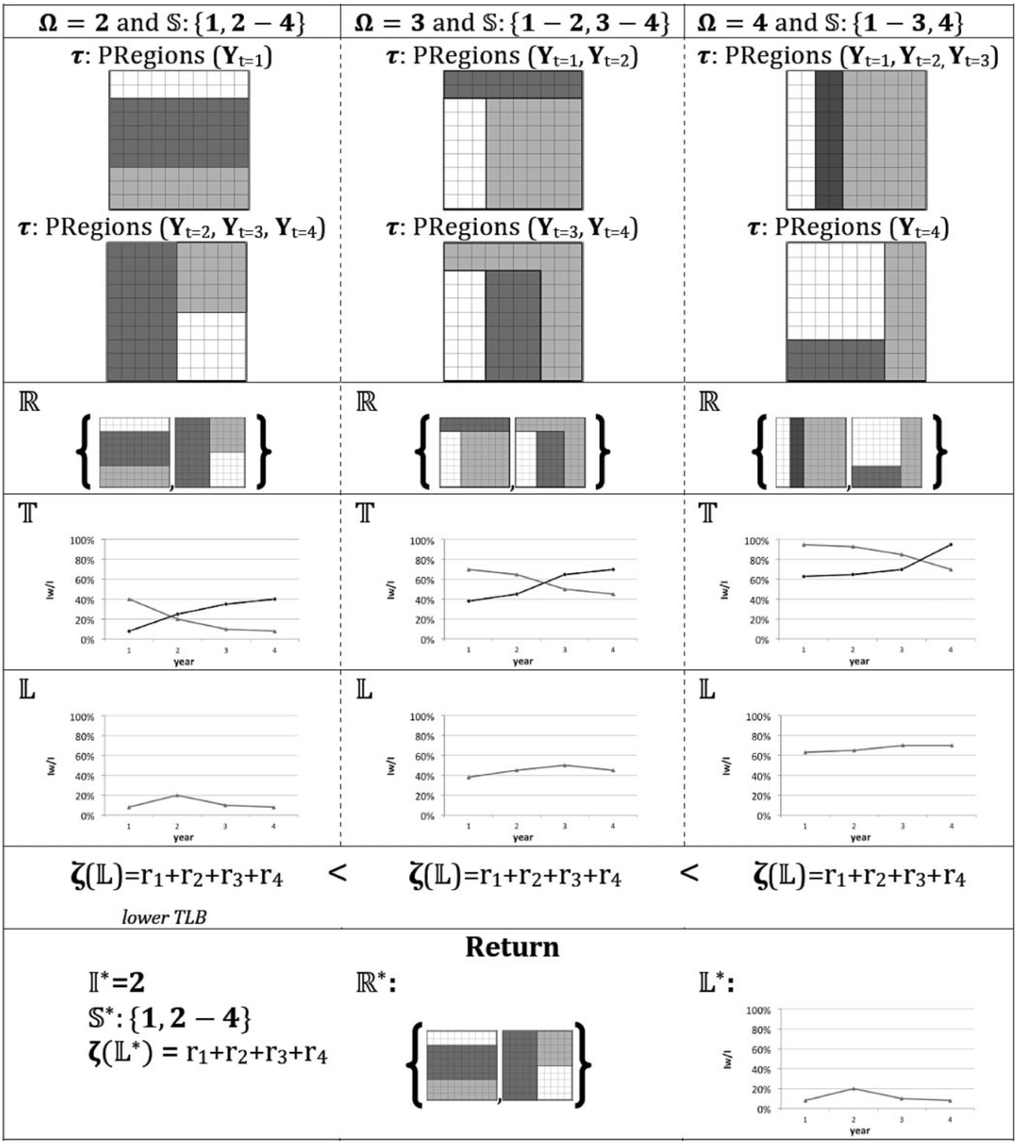

Fig. 3. Steps of spMorph for T = 4, s = 1

for the p-regions problem devised by Duque et al. (2011a). The *p*-regions problem aggregates a set of *N* areas into *p* spatially contiguous regions while minimizing the intraregional heterogeneity in terms of a given attribute. The heuristic generates feasible solutions using a scheme of 'region growth', where each region starts growing from an initial area (the seed) to which other neighbouring areas are attached (Openshaw 1977). Several feasible solutions are generated with this technique and the best solution is further improved with a local search heuristic proposed by Duque et al. (2012). This local search algorithm modifies an initial solution by moving bordering areas between regions in a way that improves the aggregation criteria while escaping from local optimal solutions. This method is just one of many ways to solve problems of spatially constrained clustering–problems that must be solved using heuristic approaches in all but trivially small cases. spMorph's reliance on heuristic approaches means that it is not guaranteed

Algorithm 1. spMorph

Algorithm: spMorph
Y: Spatial panel of **N** areas ant **T** time observations
s: Number of shocks to evaluate. $1 \leq \mathbf{s} \leq \mathbf{T}$-1
p: Number of regions

Ω: Set enumerating all possible cases of **s** shocks over **T** time observations, where each element in the set contains **s** time indices. The **s** shocks in each element of **Ω** will break the analysed period into **s** + 1 subperiods. The number of elements in **Ω** is equal to the number of **s**-combinations from **T** − 1, $\boldsymbol{C}_{\mathbf{s}}^{\boldsymbol{T}-1}$.

$\mathbb{I}^*$: ϕ, set that will contain the element in **Ω** that minimizes the function $\zeta()$.

$\mathbb{S}^*$: ϕ, set that will contain the subperiods associated to $\mathbb{I}^*$; e.g., if **T** = 10 and $\mathbb{I}^* = \{4, 7\}$, then $\mathbb{S}^* = \{1{-}3, 4{-}6, 7{-}10\}$.

$\mathbb{R}^*$: ϕ, set that will contain the **s + 1** regimes obtained by running the p-regions model for each subperiod in $\mathbb{S}^*$.

$\mathbb{L}^*$: ϕ, set that will contain the lower bound associated to $\mathbb{R}^*$.

for *i* in Ω
  $\mathbb{S}$: Set containing the subperiods associated to ***i***.
  $\mathbb{T}$: ϕ
  $\mathbb{R}$: ϕ
  for each subperiod *j* in $\mathbb{S}$
    $\boldsymbol{\tau}$: Spatial regime that results from applying the p-regions model to create **p** regions using the data from **Y** that corresponds to the subperiod **j**.
    $\mathbb{R}$: $\mathbb{R} \cup \{\boldsymbol{\tau}\}$
    $\mathbb{T}$: $\mathbb{T} \cup$. {Series that results from using the spatial regime $\boldsymbol{\tau}$ to calculate the intraregional inequality share, I_W/I, for each time observation (e.g., year, quarter, etc.) during the whole period}.
  $\mathbb{L}$: Lower bound that results from calculating, for each time observation during the whole period, the minimum value of the **s**+1 series in $\mathbb{T}$
  $\zeta(\mathbb{L})$: Total lower bound calculated as the sum of the values of the lower bound ($\mathbb{L}$)
  if $\zeta(\mathbb{L}) < \zeta(\mathbb{L}^*)$
    $\mathbb{I}^*$: *i*
    $\mathbb{S}^*$: $\mathbb{S}$
    $\mathbb{R}^*$: $\mathbb{R}$
    $\mathbb{L}^*$: $\mathbb{L}$
Return $(\mathbb{I}^*, \mathbb{S}^*, \mathbb{R}^*, \mathbb{L}^*, \zeta(\mathbb{L}^*))$

---

to find the 'true' break point(s) in a time series. However, spatial clustering algorithms have been shown to have considerable power to find true solutions under various confounding effects (Folch and Spielman 2013). Duque et al. (2007) provide a summary of regionalization methods and the tradeoffs between them.[5] Other algorithms for spatial clustering can be applied in spMorph and it becomes the researcher's decision to select the most appropriate algorithm according to their needs.

Another decision for the researcher is the number of shocks *s*. The algorithm can be run for many values of *s* and the user must determine the *s* that best fits expectations.[6] While *s* can come

[5] Some of these algorithms are available in the Python library for spatially constrained clustering called ClusterPy (Duque et al. 2011b).
[6] We acknowledge the subjective nature of these types of decisions in spMorph, and exploratory data analysis in general.

from theory or knowledge of the analysis area, it can also be guided by the results of the algorithm itself. Since the aim of spMorph is to identify general trends and summarize patterns, the best practice would be to find the fewest number of shocks that provide the most information. For a given *s*, the best solution is the one that locates the shocks in such a way as to minimize TLB; we can then use this final TLB value from spMorph for each *s* as our guide for picking an appropriate *s*. Since increasing *s* will generally lower the value of TLB,[7] the marginal gain from adding another shock must be balanced against the potential for model over-fitting. This decision can be made by plotting TLB against the number of shocks to identify the point where an increase in the number of shocks starts to produce small decreases in TLB. This decision rule will be illustrated in the next section.

4 Case study: Spatial redistribution of China's provincial *per capita* GDP between 1978 and 2008

Regional development in China exemplifies the uneven process of globalization and the significance of institutions and regions in development and change (Ye and Wei 2005). The dramatic rise of China is complicated by the diverse fortunes among its 33 provinces, as well as a general disparity between coastal provinces and inland areas. Transitional countries challenge traditional Western models, not only in the economic and political arenas, but also in theoretical inquiry (Fan and Sun 2008; Ye and Wei 2012). An exploratory approach of the type with spMorph can uncover new hypotheses on economic development that may not fit the classic models.

Regional development in China has been remarkably resilient, and has even become stronger in some cases since the onset of the recent global financial crisis in 2008. The dynamic mosaic of China's economic landscape has presented many challenges for researchers, igniting the imaginations of scholars across the social sciences. The velocity, complexity and sheer magnitude of restructuring in economic, social, political, and environmental realms can hardly be observed in the rest of the world, and merits a careful examination and assessment.

Figure 4, which presents change in *per capita* GDP (pcGDP) levels from 1978 to 2008, reveals the effects of reform on spatial association and clustering (Wei and Ye 2009). Characterized by state-owned enterprises (SOEs) and socialist institutions, the traditional leading industrial provinces in north and northeast China in 1978 were being challenged by new clusters of provinces in eastern China by 2008. Meanwhile, the coastal regions as a whole have moved ahead of the poorer interior regions, causing an intensification of the coastal-interior divide, a general rise in interregional inequality, and the formation of a poorer cluster in inland China. Development is also uneven within provinces, where new clusters and forms of regional development are emerging, from Greater Beijing in the North, to Shanghai, Jiangsu, and Zhejiang provinces in the Southeast, to Guangdong province in the South.

This redistribution can be statistically tested with the spatial version of the Cramer-von Mises two-sample test, Ψ, devised by Syrjala (1996).[8] The null hypothesis of this non-parametric test is that there is no difference between two spatial distributions. Comparing the distributions of China's provincial pcGDP in 1978 and 2008, $\Psi = 0.037$, yielding a

[7] This is because, the shorter the period, the easier it will be to find a spatial regime that aggregates the small areas into homogeneous regions, and therefore, the greater the possibility to minimize the ratio Iw/I.

[8] For more on Cramer-von Mises test see: Darling (1957) and Anderson (1962).

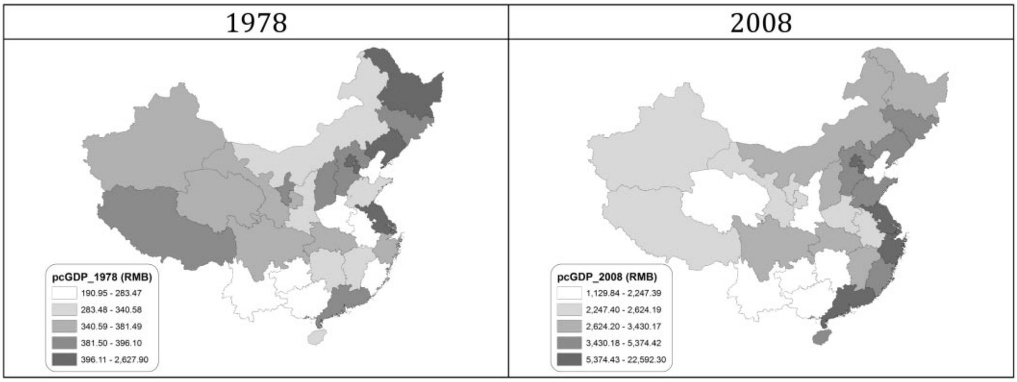

Fig. 4. Distribution of China's provincial pcGDP in 1978 and 2008

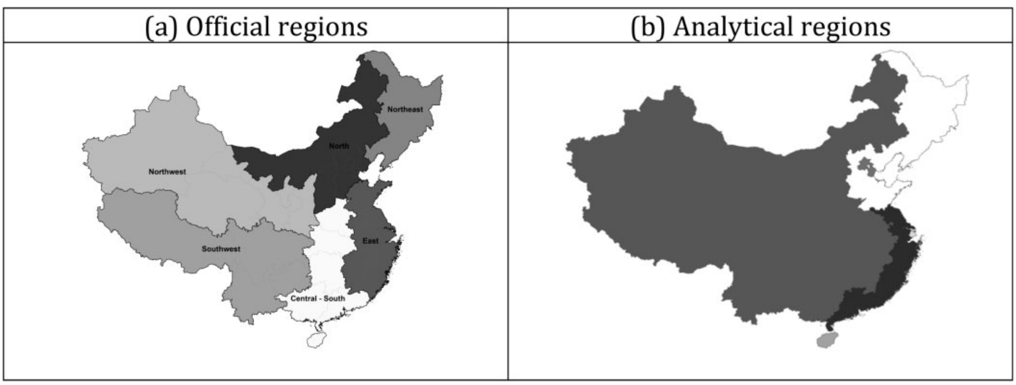

Fig. 5. Official regions vs. analytical regions for the period 1978–2008

p-value = 0.003,which indicates that the difference between the distributions in 1978 and 2008 is statistically significant.

Regional and urban policy in China is deeply rooted within a long historical tradition of separating cities from the countryside, as well as socialist practices since the establishment of the People's Republic of China (PRC) in 1949. As Wei (2007, p. 24) argues, 'the role of the state [in China] has strong geographical foundations, and is constrained by local institutions and geographies'. The regional forces, such as social, economic, and cultural agents embedded in specific physical environments, interact with national forces to shape China's 33 provinces into six major groupings of regional development (Ye and Xie 2012; Ye and Wei 2012). These six regions (we will call them 'official regions' in this paper), which are based on macroeconomic linkages and historical legacy, are East, Central-South, North, Northeast, Northwest, and Southwest (Xie and Dutt 1990). These official regions are considered to be homogenous regarding regional and urban development over time. In other words, the provinces within each region show similar dynamics. As our results will show, these commonly held beliefs over hundreds of years may no longer hold. The official regions are presented in Figure 5a.

East China and Northeast China are traditionally rich regions with advanced economies and remarkable large cities. East China is the most developed region in the country with the nation's largest city Shanghai located there. The rise of rural industries in East China contributed to a bottom-up urbanization. Northeast China carries the hallmarks of change brought about by economic development policies under Chairman Mao. As the heavy-industry base in China, it was given investment priorities before the economic reform. However, due to the introduction of a market economy in the 1980s and a decline in preferential policies, Northeast China's reliance on state owned enterprises (SOEs) has since led to slower regional growth. In North China and Northwest China, central government sponsored development is predominant. The central government is the primary financial provider in these regions, which are home to most of the country's ethnic minorities. Furthermore, a large share of the state investment in these regions has gone to several top cities thus resulting in a high urbanization and urban primacy. North China was designated as the national base of coal and energy resources and acted as the main recipient of state investment. With the national capital Beijing located there, the state force is excessively dominant. Bottom-up urbanization and fast growth of medium-sized cities in Central-South China has significantly contributed to divergent growth in this region. Southwest China has witnessed slower growth due to the lack of foreign investment. Traditional interior industrial bases and many other rural areas have fallen behind because of disadvantageous location and problematic SOEs, forming a poverty trap. Moreover, weak linkages between coastal and interior China blocked the infusion of development to poor areas. Consequently, the gaps between coastal and interior China and between rich and poor regions have widened dramatically, which directly contributes to the problem of containing provinces with various fortunes in one region.

The series labelled 'official' in Figure 6 shows the inequality within regions (Iw) as a percentage of total inequality (I) when using the official regions (Iw/I). The inequality within regions represents, on average, 70.3 per cent of total inequality. This implies that there exist important differences between provinces assigned to the same official region. For instance, East China was ignored by Mao's industrialization policy due to its coastal location and lack of natural resources. In the 1950s and 1960s, the central government invested heavily in the interior of China. In the early 1970s, China normalized its relationship with the United States and reduced its emphasis on self-reliance and national defence. With economic reforms and political decentralization, large cities in East China have grown rapidly and many township and village-owned enterprises (TVEs) were established in coastal rural areas. Hence, there is a growing divide among the fortunes of the inland and coastal provinces traditionally assigned to East

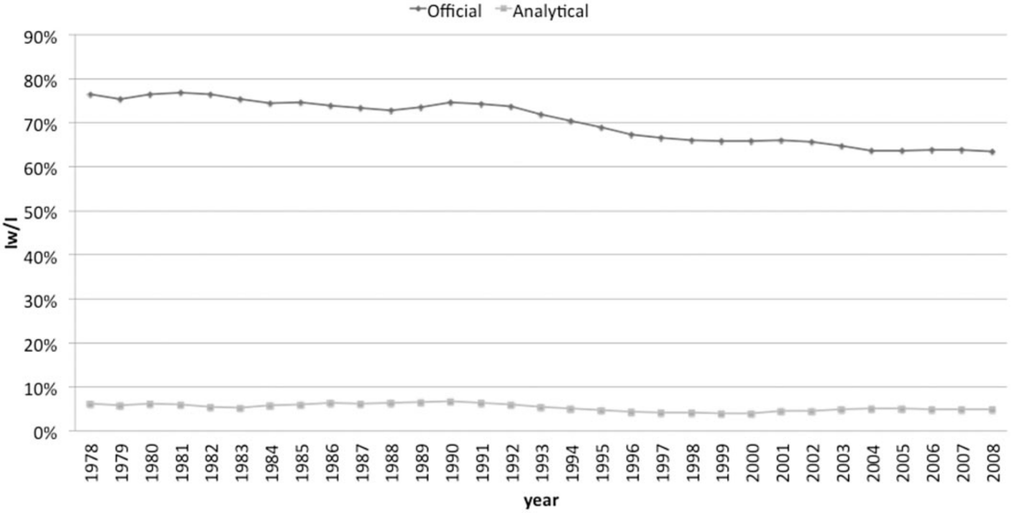

Fig. 6. Inequality within regions (Iw) as a percentage of total inequality (I)

China. A similar pattern applies to Central-South China. At the same time, the disparities between many inland provinces have declined, as revealed by Figure 5b.

While the official regions reflect traditional groupings of Chinese provinces, they do not appear to reflect the economic reality on the ground. When we apply the *p*-regions model to the 31-year panel of pcGDP, and set the number of regions to six, we find a vastly different spatial regime. Figure 5b demonstrates that all inland provinces form one analytical region, while the remaining five regions are along the coast. This map indicates that Mao's socialist ideology, Deng's pragmatism, and other factors have resulted in a great divergence in development across China.

The series labelled 'analytical' in Figure 6 shows the ratio Iw/I when using the analytical regions. Now the inequality within regions represents, on average, 5.39 per cent of total inequality. This is an important reduction compared with the share obtained for the official regions. Because of the nature of spatial data, it is not possible to predefine a threshold value below which the share of intraregional inequality can be considered low enough. The potential reduction is, at least, a function of the following aspects:

1. The spatial distribution of the variable: The spatial contiguity constraint, which requires areas assigned to the same region be geographically connected, reduces the possibility of aggregating areas with similar values. According to Duque et al. (2006), the impact of this constraint in hampering the reduction of intraregional inequality depends on the spatial distribution of the variable. Thus, when a given variable shows a high level of positive spatial autocorrelation, areas with similar values will tend to be near to each other. Such a pattern will facilitate the construction of regions containing areas with similar values, which, therefore, increases the possibility of reducing the intraregional inequality.
2. The number of regions (p): this parameter constrains the possibility of reducing the share of intraregional inequality. Thus, the higher the number of regions, the higher the possibility of reducing the share of intraregional inequality.

Because of the spatial redistribution of China's provincial pcGDP between 1978 and 2008, it is possible that a proper representation of these spatio-temporal dynamics requires more than the single regime presented in Figure 5b. To explore this possibility we apply the spMorph algorithm to determine: (*i*) the number of shocks and the year(s) they occurred; (*ii*) the time taken for each shock to generate a spatial redistribution of the variable resulting in a new spatial regime (i.e., the duration of the regime fade); and (*iii*) the regional configuration of each spatial regime. Table 1 presents the parameters we use for the algorithm. We fix the number of analytical regions at six, in line with the number of official regions. Regarding the potential number of shocks, we allow for up to four shocks during the period 1978–2008. We also calculate the solutions for 30 shocks, which although not practical from a macroeconomic evaluation perspective, provide important information about the minimum level that the total lower bound (TLB) can reach.

Figure 7 presents the lower bounds obtained for different numbers of shocks (s). The highest lower bound, labelled '0', corresponds to the case where there is no shock, which implies that there is a unique spatial regime for the 31-year period. This lower bound matches the one

Table 1. Parameters of the spMorph algorithm

Parameters	Values
Number of shocks to evaluate (s)	0, 1, 2, 3, 4, 30
Number of regions (p)	6

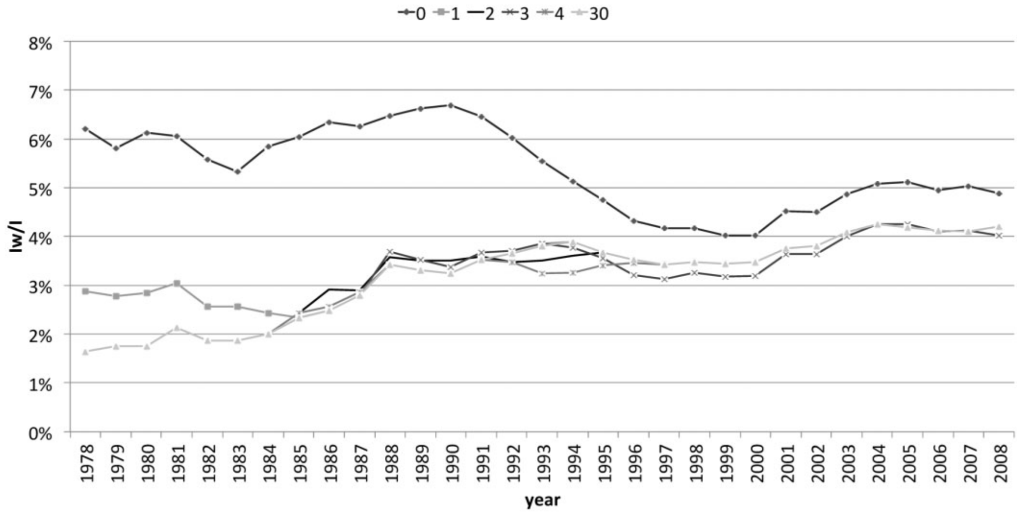

Fig. 7. Lower bound for various numbers of shocks

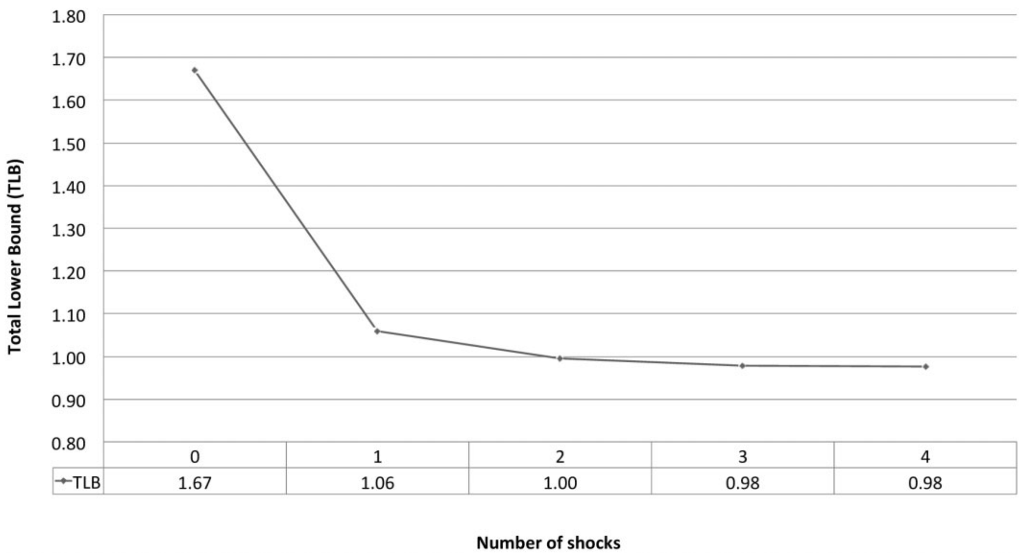

Fig. 8. Total lower bound (TLB) for different numbers of shocks

labelled 'analytical' in Figure 6. As can be seen in Figure 7, the consideration of additional shocks generates a downward movement of the lower bound. The decision now is to determine the number of shocks that best represents the spatio-temporal dynamic of China's provincial pcGDP between 1978 and 2008. While this decision is as subjective as determining the number of clusters in a conventional cluster analysis, it can be guided by the data available. It is important to keep in mind that the proposed methodology is intended to provide guidance on general trends; therefore, we recommend selecting the minimum number of shocks that generate the largest improvement in the lower bound (an 'improvement' in the lower bound is equivalent to a downward movement of the lower bound). As a possible way for determining the number of shocks, we present in Figure 8 the TLB for different numbers of shocks. From Figures 7 and

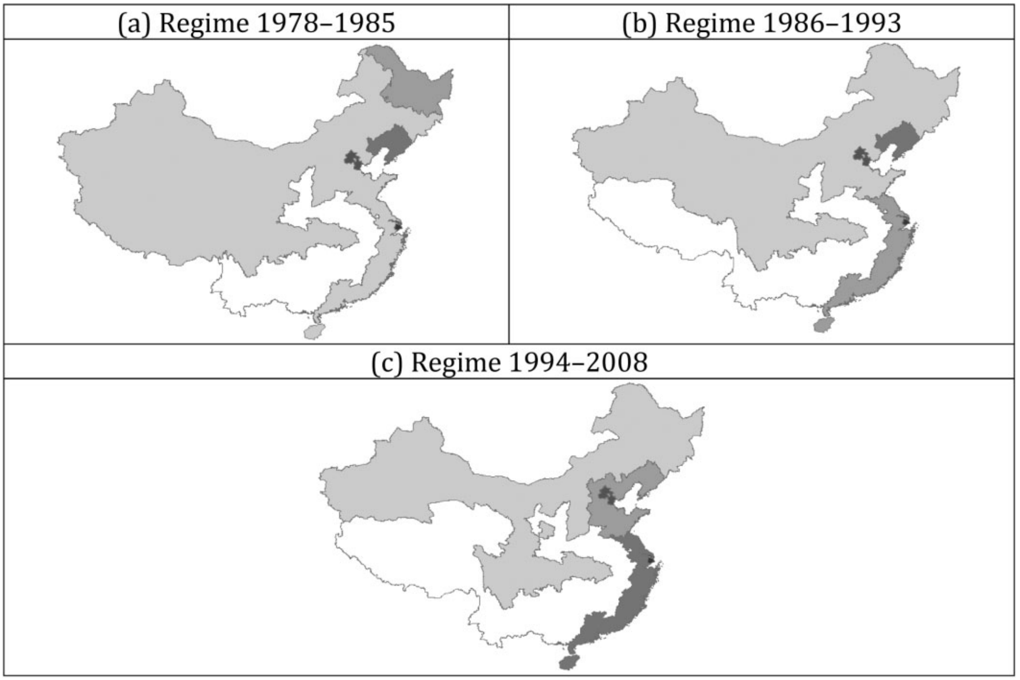

Fig. 9. Spatial regimes obtained by spMorph for the three sub-periods

8, it appears that the consideration of two shocks is a good option. Note, in both figures, that beyond two shocks, the changes in the lower bounds are minimal.

When running the algorithm for two shocks, the results indicate that the first shock occurred in 1985 and the second shock in 1993. These two shocks split the 31-year period into three sub-periods: 1978–1985, 1986–1993 and 1994–2008. These three resulting spatial regimes are presented in Figure 9. The evolution of these three regimes shows the formation and emergence of coastal clusters.

Regional inequality has long been the subject of intense academic debates on development and inequality across China's regions. In just thirty years of development, China has gradually evolved from an agricultural society to an industrial nation based on a socialist market economy, with location playing a more important role. This has resulted in dramatic variations within traditionally defined regions. It is noted that the variance reflects the difference of legacy and mechanisms of urban and regional development among regions in China, which is also imbedded in the socio-economic environment of individual regions in response to economic reforms and globalization. Since the late 1970s, China has been undergoing economic reforms–introducing market mechanisms and opening up to the outside world. The spMorph method sheds new light on the convergence and divergence debate among China scholars, and can help inform the nature, trajectory, and impacts of the reform policies, including the effects on regional inequality.

Figure 10 presents the share of intraregional inequality (Iw/I) generated for each spatial regime. The first spatial regime minimizes the ratio Iw/I for the period 1978–1985. In 1986 a new spatial regime emerges, which lasts until 1994. Although the second spatial regime was generated with data from 1986 to 1993 it predominates until 1994, when the third spatial regime begins to emerge. This is a clear example of the fade effect: it takes one year for the shock in 1993 to result in a new spatial regime.

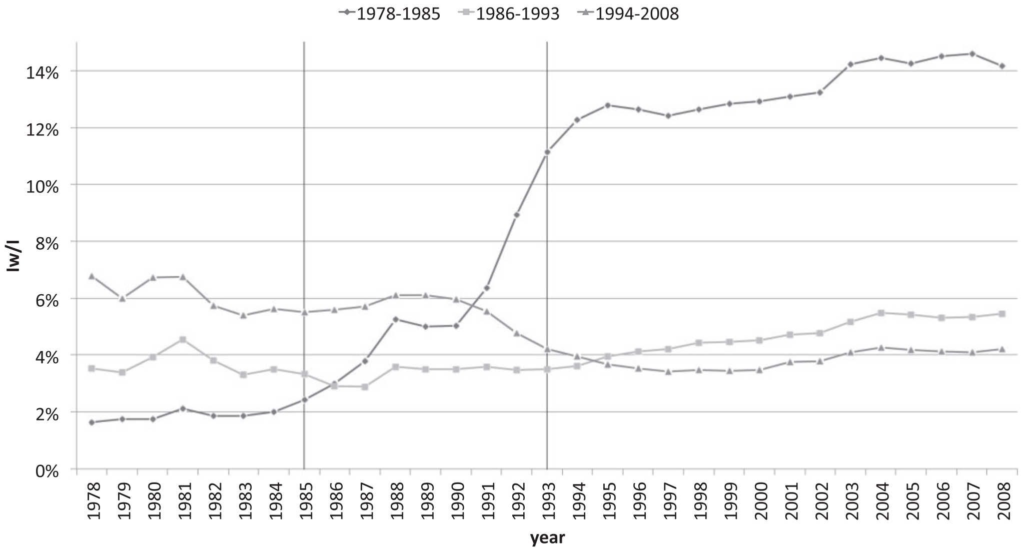

Fig. 10. Inequality within regions (Iw) as a percentage of total inequality (I)

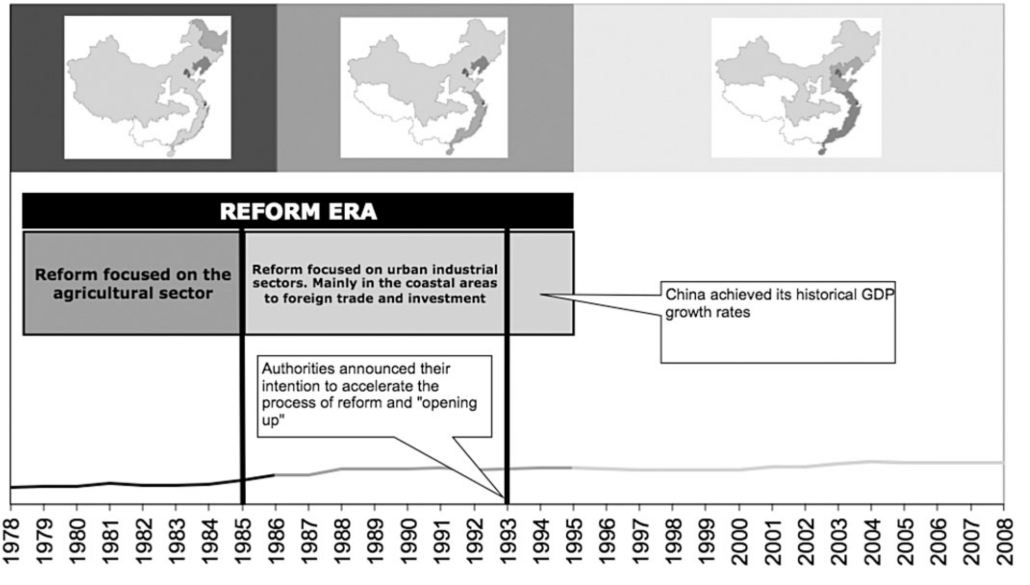

Fig. 11. Linkages between the identified shocks and major historical events

Figure 11 presents the linkages between the identified shocks and major historical events. The proposed method captures the main reform practices in China. Research on regional economic growth has shown that the complex interaction of economic wellbeing, the threat/ opportunity from globalization, and the formation of local institutional responses have driven economic development into new arenas. China's reforms can be understood as a triple transition process of decentralization, marketization, and globalization, with three dominant forces of regional development – the state, the locality, and the global investor (Wei 2007). State and locality remain the vital factors in directing and facilitating the transition of China's economy from a resource-constrained to demand-constrained one in the context of globalization and transition. Hence, identifying the spatial effects of policy is particularly important for such a large country. spMorph will help decision-makers to identify whether their regions are at risk or face opportunity should economic reorganization take place. In general, this method contributes to place-based public policy research in developing countries, especially those in transition. Practically, by tracking the evolution of the shocks over time and space, we not only monitor the process of the economic development, but can also detect irregular patterns in the process of growth as a reference for policy making.

5 Case study: Spatial redistribution of US unemployment rates and home price index between 2002 and 2012

The Chinese case study is one of significant economic transition over a period of three decades. In this section we apply the spMorph approach to the economically tumultuous period from 2002 to 2012 in the US. This is a case of a developed nation experiencing its most significant economic downturn in more than 60 years.

As the US economy emerged from the dot.com bubble and bust in 2000 and the terrorist attacks of September 11, 2001, another bubble began growing. Housing prices had been increasing steadily since the late 1990s, but by 2002 prices were clearly higher than any level seen in decades (Figure 12). These unprecedented prices continued increasing for the next four years forming an economic bubble that literally reached into the majority of US homes as home ownership rates had increased to nearly 70 per cent at the peak of the price bubble. While the housing bubble was a national phenomenon, its magnitude varied from place to place, with different metropolitan areas following different trajectories. Figure 13 presents time series of real median home values for a selection of metropolitan areas. Chicago and Boston had gradual increases in home prices, and corresponding gradual declines, while the other cities presented had much more steeper rises and falls. Furthermore, the shapes of the peaks varied with Las Vegas's dome shape indicating it sustained its high prices, while most other cities spent less time at the top of the cycle.

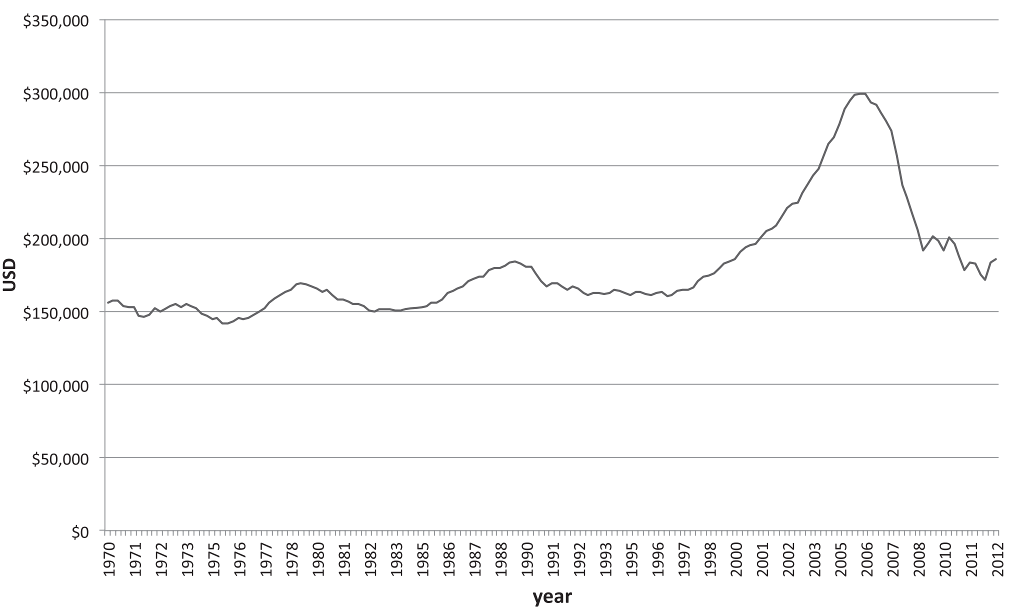

Fig. 12. Real US median home prices, 1970–2012

Source: www.jparsons.net.

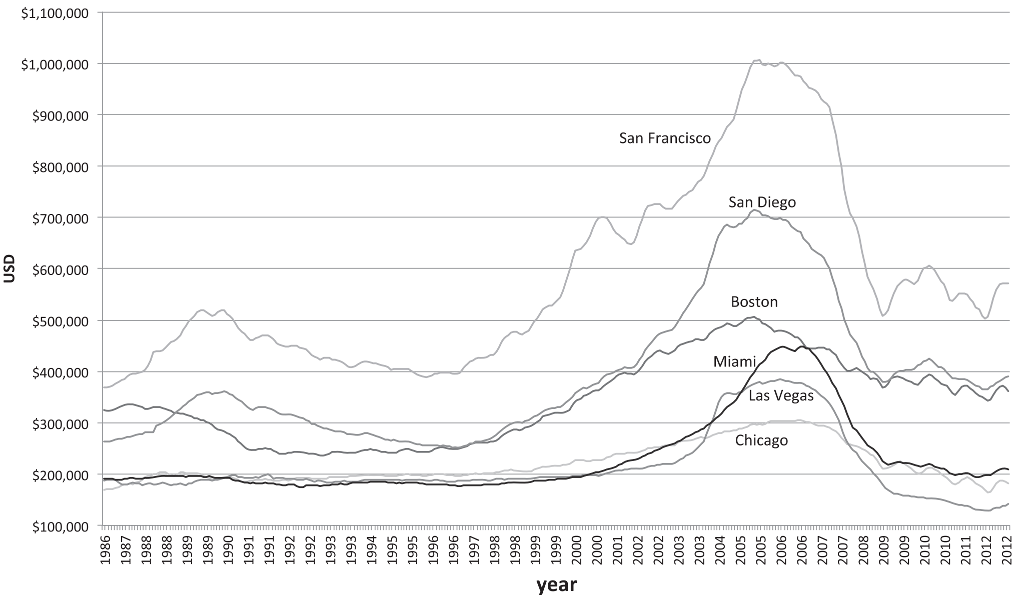

Fig. 13. Selected urban area real median home prices, 1986–2012
Source: www.jparsons.net.

Fig. 14. Distribution of home price index (HPI) in 2002 and 2012

This spatial variation in housing market patterns hints that the spMorph approach is likely a good fit for detecting structural breaks in economic time series over the past decade. Figure 14 presents the state level housing price index (HPI) for 2002 and 2012. The HPI is calculated individually for each state, with the second quarter of 2000 defined as 1.0.[9] The HPI therefore standardizes each state's housing prices, and so allows us to compare high cost of living costal states with lower cost interior states. While all but four states had higher home values in 2012 than 2002, the spatial pattern of relative change within states was quite different. In 2002, California and many other coastal states were seeing the greatest impact of the housing price increases. By 2012 however, states in the mountain west and in the mid-Atlantic were seeing the largest relative home price increases.

An economic crisis of the magnitude caused by the housing bubble was seen not only in the housing market, but rippled throughout the economy. Figure 15 presents the 2002 and 2012 unemployment rate by state. Similar to the housing market data, all but three states had higher

[9] The HPI is based on Federal Housing Finance Agency (FHFA) repeat-sales data. The data can be found at Land and Property Values in the US, Lincoln Institute of Land Policy http://www.lincolninst.edu/resources/.

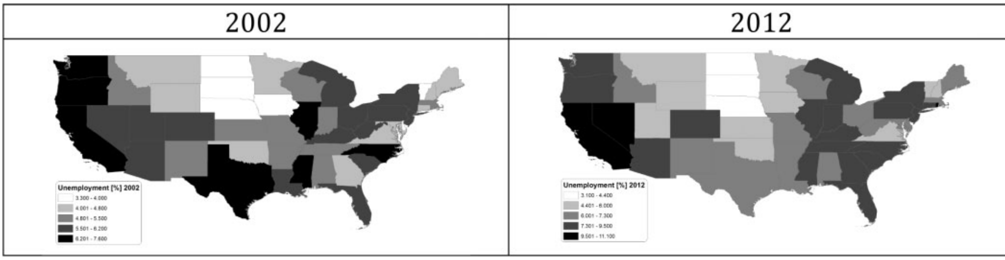

Fig. 15. Distribution of Unemployment rates for states in 2002 and 2012

unemployment rates in 2012 than 2002. In contrast to the HPI maps, the general spatial pattern in the unemployment maps looks somewhat similar. Many states shifted by only one category; most of those in the highest category in 2002, for example, dropped to a lower rank in 2012. While not as stark of a transition, there remains considerable variation in the unemployment rate before and after the crisis.

HPI and unemployment rate represent direct and indirect indicators of the housing market bubble and its aftermath. We feel comfortable assuming a single shock over the period under investigation, but its exact timing is unclear. The other key decision variable for spMorph is the number of regions.

The US has used many official and unofficial regional groupings of states in its history. These changes in definitions can largely be attributed to the expansion of the country as 'territories' transitioned to 'states', and as states evolved from their resource based origins to more diversified economies. The US government currently has three broadly used regional groupings of states: the US Census Bureau's (USCB) 'regions' (4); their 'divisions' (9); and the Bureau of Economic Analysis' (BEA) 'regions' (8). The nine census divisions have been stable since 1910, except for the additions of Hawaii and Alaska in 1959. However, these were largely influenced by groupings from the late nineteenth century, an era of rapid growth, but still predating much of the industrialization that would define the US in the decades before and after the Second World War (see Chapter 6 of US Census Bureau, 1994 for an extended history of US regions).

A 1950s era effort to update the USCB divisions using up to date social and economic data did not gain traction within the USCB. However, the BEA chose to adopt one of the proposals with a slight modification, and this formed the eight BEA regions still in use today (Crone 2005). Crone (2005), in an effort to modernize the BEA regions, performed a regionalization using an approach with multiple attributes from multiple years. We use the BEA regions (Figure 16a) and Crone regions (Figure 16b), which both split the country into eight spatially contiguous clusters, as benchmarks for spMorph approach.

For this case, as in China, we define analytical regions using the *p*-regions approach (Duque et al. 2011a). Following the BEA and Crone regionalizations, we identify eight regions; and in this case we use the full 11-year panels for HPI and unemployment rate. Therefore, the algorithm seeks to design a unique spatial regime that minimizes the intraregional heterogeneity for both variables over for the whole period (Figure 16c). The three maps in Figure 16 show considerable similarity, with many pairs of states grouped together in all the approaches. We note that the Crone regions are also analytic regions, and the BEA regions, while not finding their provenance in computer driven algorithms, could also be considered analytic regions due to their grounding in socio-economic data of the day.

Next, we evaluate the capacity of these three analytical regions to summarize the spatio-temporal dynamics of both variables over the whole period. Figure 17 shows the Iw/I ratio for HPI. In this case there are not significant differences between the three analytical regions. None

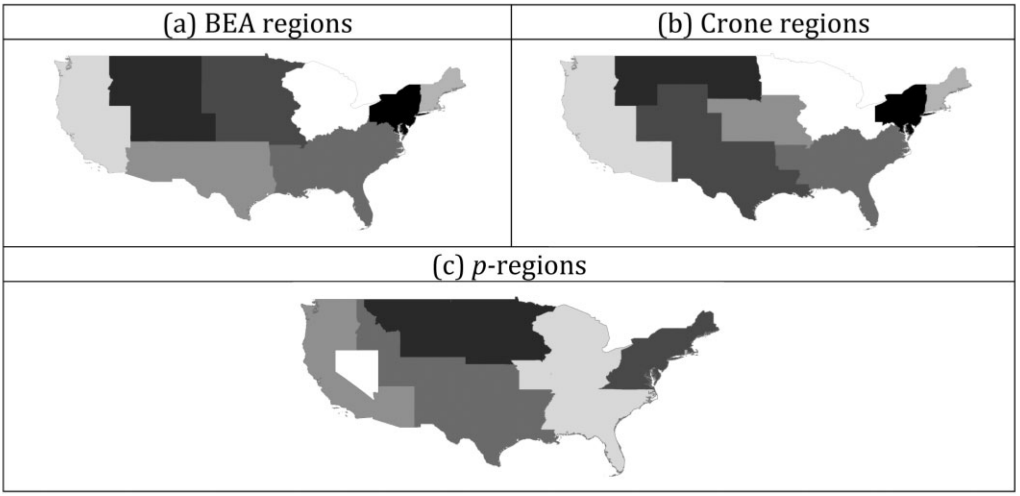

Fig. 16. BEA regions, Crone regions and analytical regions for the period 2002–2012
Notes: There are three singletons in the spatial regime 16c: Nevada, Rhode Island and Washington, DC.

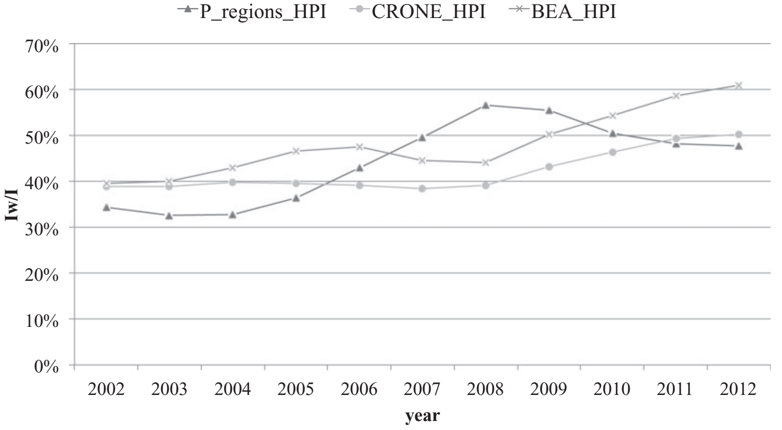

Fig. 17. Inequality within regions (Iw) as a percentage of total inequality (I): home price index

of the spatial regimes consistently outperform the other two options, and the average intraregional inequality share (Iw/I) is around 45per cent. Figure 18 shows the Iw/I ratio for unemployment. The *p*-regions outperform the BEA and Crone regions in all 11 years tested, but its capacity to capture the spatio-temporal dynamics of the variable is poor for the period 2002–2007. The average intraregional inequality share for BEA and Crone regions is 57 per cent, which suggests that neither are good spatial regimes to summarize unemployment rates between 2002 and 2012.

Up to this point we know that none of these spatial regimes is capable of summarizing the spatio-temporal dynamics of HPI and unemployment rate during the period 2002–2012. We also know that there was an important economic shock in 2007 and we want to see what spMorph can tell us about the changes in the spatio-temporal dynamics of these two variables. For this, we run spMorph using one shock (i.e.; $\mathbb{S}$: $\{2002-2007, 2008-2012\}$; $\mathbf{\Omega} = \{2008\}$), eight regions ($p$ = 8), and two variables (HPI and unemployment). With these parameters, spMorph generates two

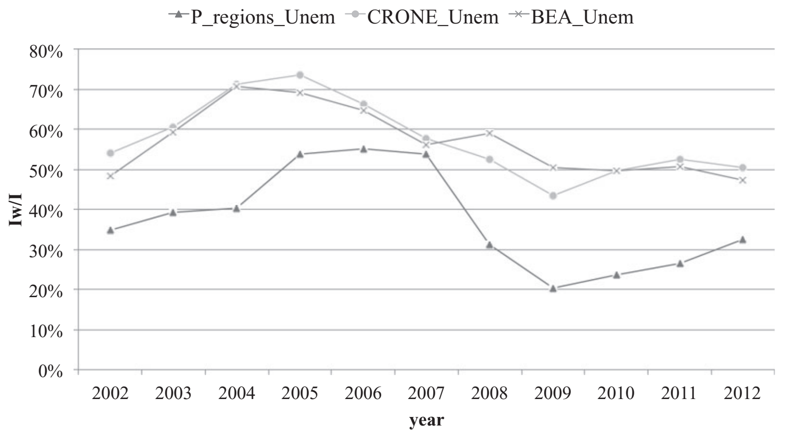

Fig. 18. Inequality within regions (Iw) as a percentage of total inequality (I): unemployment rate

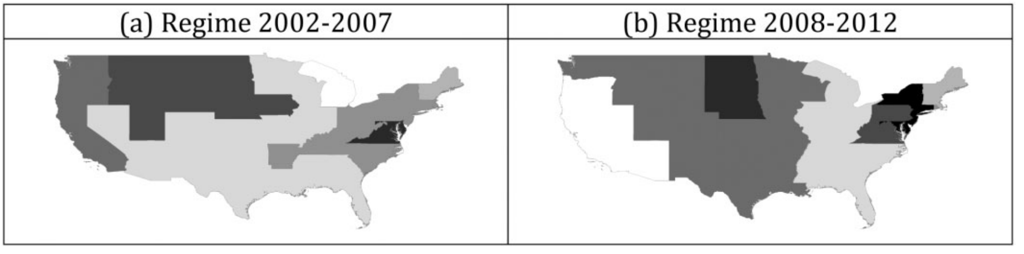

Fig. 19. Spatial regimes obtained by spMorph for the two sub-periods
Notes: The eighth region in 19a is Washington, DC, and the eighth region in 19b is Rhode Island.

spatial regimes: the first one uses HPI and unemployment from 2002 to 2007, and the second spatial regime results from using HPI and unemployment from 2008 to 2012. These two regimes are presented in Figure 19. To some extent the grounding for these two regimes can be seen in the two years of input data presented in Figures 14 and 15. The 2002 input data maps appear slightly more random than their 2012 counterparts, and this is reflected in the two regimes. The 2008–2012 regime is more compact reflecting more similarity between nearby states, while in the 2002–2007 regime there is one region that spans nearly the entire country. While the shapes of the regimes may not reflect conventional notions of US regions, they do reflect the HPI and unemployment rate used to create them.

Each one of these two regimes is used to calculate the Iw/I ratio from 2002 to 2012. This ratio is calculated for each variable separately. This means that we end up with two TLBs and two regime fades, one for each variable. This provides us information about which variable reacted first to the shock as well as the capacity of each regime to summarize the spatio-temporal dynamics of both variables during the analysed period. Figure 20 shows that for HPI, the regime fade is nearly non-existent as there is little distance between 2007 and the line cross. This result is consistent with the notion of a bubble bursting and rapid change in the economic landscape. Figure 21 presents the same information, but for unemployment rate. Here we find unemployment rate taking nearly one year to resettle into a new regime. Again this is consistent with the interlinked nature of the US economy, where the end of the housing bubble led to a broad recession that spread to most sectors of the US economy.

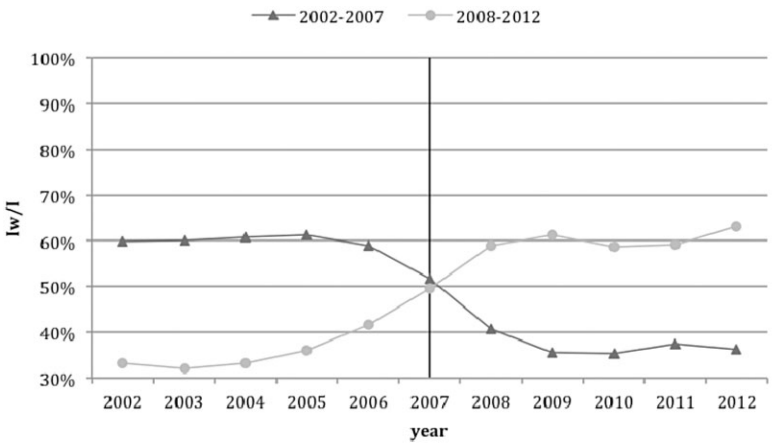

Fig. 20. Inequality within regions (Iw) as a percentage of total inequality (I): home price index

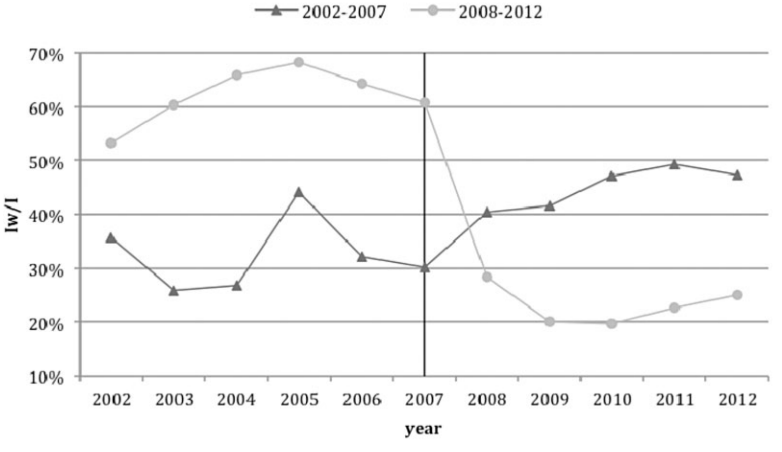

Fig. 21. Inequality within regions (Iw) as a percentage of total inequality (I): unemployment

6 Conclusions

The applications in this paper presented two complex cases: one in which the researcher is exploring the potential location of the shock that triggers the spatial redistribution of the attribute; and another where the researcher is investigating how two different attributes react to the same shock. However, this methodology can be applied to other cases. For example, how does the spatial distribution of unemployment rate react within different countries of the Euro-zone as a consequence of a European policy? This is a case of one variable and different spaces – how do different countries react to the same shock?

Another perspective is that we leverage the tools of spatial clustering solely to uncover the temporal break(s) in the variable. From this perspective, spMorph can be used in concert with unit root tests and other structural break identification tools (see Glynn et al. 2007 for a summary of some approaches in the econometric field). It is likely that spMorph can be used as an alternative tool to identify, in an endogenous way, the number and position of structural breaks in time series. We see this as fruitful line for further research, but one that requires a thorough treatment of the theoretical issues in an econometric setting.

Several overarching dimensions can be identified in follow-up research: (a) an emphasis on knowledge discovery in a multi-mechanism real world with various shocks and transitions, including comparisons of multi- and across space-time scales might shed more light on the topic; (b) a focus on inequality as the central research theme; (c) a recognition of the need to develop and implement novel space-time analytical methods for analysing the complexities of regional development; and (d) a concern for the social and policy implications of economic activities in the context of globalization and reform.

Appendix

A stripped-down version of the actual Python code used to obtain the presented results. Procedures for handling input and output data were removed and also caching functions used for performance purposes.

```
def spMorph(self, variables, minShocks, maxShocks, inequalityIndex, \
                          outFile, clusterAlgorithm, nClusters, **kargs):
    bestSolutions = {}
    '''
    Prepare output files.
    '''
    for nElements in range(minShocks, maxShocks + 1):
        bestSol = [] # (t, tb, tw, tw/t, loweTw/t ,comb ,objectiveFunction)

        '''
        combinations built-in function:
        Returns subsequences of variables[1:] of length nElements
        '''
        for comb in combinations(variables[1:], nElements):
            comb = list(comb)
            comb.sort()
            comb = tuple(comb)
            t, tb, tw, tw_t, lowerTw_t, a2r = inequalityShock(variables, \
                                  comb, inequalityIndex, clusterAlgorithm,\
                                                        nClusters, **kargs)
            if bestSol = = [] or sum(lowerTw_t) < = (bestSol['of']):
                bestSol = {'t' : t,
                           'tb' : tb,
                           'tw' : tw,
                           'tw_t' : tw_t,
                           'lowerTw_t' : lowerTw_t,
                           'comb' : comb,
                           'of' : sum(lowerTw_t),
                           'a2r' : a2r}
        '''
        Write best solution (bestSol) to files.
        '''
def inequalityShock(self, variables, shokVariables, inequalityIndex, \
                                 clusterAlgorithm, nClusters, **kargs):
    tempSet = [], area2regions = {}
    area2regionsList = [], tempSetOrdered = []

    '''
    layer = importArcData( region data )
    The layer object holds information about the region and has access to
    the clustering algorithms.
    '''
```

```
for var in variables:
    if var in shokVariables:
        if tempSet = = []:
            raise NameError('First period could not be a shock period')
        else:
            tempSet.sort()
            tempSet = tuple(tempSet)
            clusterArgs = (clusterAlgorithm, tempSet, nClusters)
            layer.cluster(*clusterArgs, **kargs)
            area2regions[tempSet] = layer.region2areas
            tempSetOrdered.append(tempSet)
            tempSet = [var]
    else:
        tempSet.append(var)
tempSet.sort()
tempSet = tuple(tempSet)
clusterArgs = (clusterAlgorithm, tempSet, nClusters)
layer.cluster(*clusterArgs, **kargs)
area2regions[tempSet] = layer.region2areas
tempSetOrdered.append(tempSet)

'''
getVars function:
Getting subsets of data
This function allows the user to extract a subset of variables from a
layer object.
'''
Y = layer.getVars(*variables)
t = [], tb = [], tw = [], tw_t = []
for a2r in tempSetOrdered:

    '''
    inequalityMultivar function:
    Inequality index for multiple variables.
    This function calculates a given inequality index for multiple
    variables.
    '''

    t2, tb2, tw2, tw_t2 = inequalityMultivar(Y, area2regions[a2r], \
                                             inequalityIndex)
    a2rc = area2regions[a2r]
    t.append(t2)
    tb.append(tb2)
    tw.append(tw2)
    tw_t.append(tw_t2)
    area2regionsList.append(a2rc)
lowerTw_t = [min(x) for x in zip(*tw_t)]
return t, tb, tw, tw_t, lowerTw_t, area2regionsList
```

References

Alonso AM, Berrendero JR, Hernández A, Justelb A (2006) Time series clustering based on forecast densities. *Computational Statistics & Data Analysis* 51: 762–776

Anderson TW (1962) On the distribution of the two-sample Cramer-von Mises criterion. *The Annals of Mathematical Statistics* 33: 1148–1159

Crone TM (2005) An alternative definition of economic regions in the United States based on similarities in state business cycles. *The Review of Economics and Statistics* 87: 617–626

Darling DA (1957) The Kolmogorov-Smirnov, Cramer-von Mises tests. *The Annals of Mathematical Statistics* 28: 823–838

Duque JC, Artís M, Ramos R (2006) The ecological fallacy in a time series context: Evidence from Spanish regional unemployment rates. *Journal of Geographical Systems* 8: 391–410

Duque JC, Church RL, Middleton RS (2011a) The p-regions problem. *Geographical Analysis* 43: 104–126

Duque JC, Dev B, Betancourt A, Franco, JL (2011b) ClusterPy: Library of spatially constrained clustering algorithms. Version 0.9.9. RiSE-group (Research in Spatial Economics), EAFIT University

Duque JC, Ramos R, Surinach J (2007) Supervised regionalization methods: A survey. *International Regional Science Review* 30: 195–220

Duque JC, Rey SJ, Anselin L (2012) The max-p-regions problem. *Journal of Regional Science* 52: 397–419

Fan CC, Sun MJ (2008) Regional inequality in China, 1978–2006. *Eurasian Geography and Economics* 48: 1–20

Folch DC, Spielman SE (2013) Identifying regions based on flexible user defined constraints. *International Journal of Geographical Information Systems*. doi:10.1080/13658816.2013.848986

Fruhwirth-Schnatter S, Kaufmann S (2004) Model-based clustering of multiple time series. CEPR Discussion Paper No. 4650

Galeano P, Peña D (2000) Multivariate analysis in vector time series. *Resenhas* 4: 383–403

Glynn J, Perera N, Verma R (2007) Unit root tests and structural breaks: A survey with applications. *Revista de Métodos Cuantitativos para la Economía y la Empresa* 3: 63–79

Gordon AD (1996) A survey of constrained classification. *Computational Statistics & Data Analysis* 21: 17–29

Kakizawa Y, Shumway RH, Taniguchi M (1998) Discrimination and clustering for multivariate time series. *Journal of the American Statistical Association* 93: 328–340

Kulldorff M, Hjalmars U (1999) The Knox method and other tests for space-time interaction. *Biometrics* 55: 544–552

Murtagh F (1985) A survey of algorithms for contiguity-constrained clustering and related problems. *The Computer Journal* 28: 82–88

Openshaw S (1977) A geographical solution to scale and aggregation problems in region-building, partitioning and spatial modeling. *Transaction of the Institute of British Geographers* 2: 459–472

Pattarin F, Paterlini S, Minerva T (2004) Clustering financial time series: An application to mutual funds style analysis. *Computational Statistics & Data Analysis* 47: 353–372

Rey SJ, Janikas MV (2005) Regional convergence, inequality, and space. *Journal of Economic Geography* 5: 155–176

Rey SJ, Ye X (2010) Comparative spatial dynamics of regional systems. In: Páez A, Le Gallo J, Buliung R, Dall'Erba S (eds) *Progress in spatial analysis: Theory, computation, and thematic applications*. Springer, Heidelberg Dordrecht London New York

Shorrocks A, Wan G (2005) Spatial decomposition of inequality. *Journal of Economic Geography* 5: 59–81

Syrjala SE (1996) A statistical test for a difference between the spatial distributions of two populations. *Ecology* 77: 75–80

Theil H (1967) *Economics and information theory*. Amsterdam, North Holland

Theil H (1972) *Statistical decomposition analysis*. Amsterdam, North Holland

US Census Bureau (1994) Geographic areas reference manual. Washington, DC, Department of Commerce

Wei Y (2007) Regional development in China: Transitional institutions, embedded globalization, and hybrid economies. *Eurasian Geography and Economics* 48: 16–36

Wei YD, Ye X (2009) Beyond convergence: Space, scale, and regional inequality in China. *Tijdschrift voor Economische en Sociale Geografie* 100: 59–80

Xie Y, Dutt AK (1990) Regional investment effectiveness and development levels in China. *GeoJournal* 20: 393–407

Ye X, Rey SJ (2013) A framework for exploratory space-time Analysis of Economic Data. *Annals of Regional Science* 50: 315–339

Ye X, Xie Y (2012) Re-examination of Zipf's law and urban dynamic in China: A regional approach. *Annals of Regional Science* 49: 135–156

Ye X, Wei YD (2005) Geospatial analysis of regional development in China: The case of Zhejiang province and the Wenzhou model. *Eurasian Geography and Economics* 46: 342–361

Ye X, Wei YD (2012) Regional development, disparities and polices in globalizing Asia. *Regional Science Policy & Practice* 4: 179–182

doi:10.1111/pirs.12088

Resumen. Este artículo presenta una herramienta preliminar de análisis espacio-temporal para determinar los dos componentes de un proceso de redistribución espacial: (*i*) la perturbación, que es el momento que desencadena un proceso de redistribución espacial; por ejemplo, una nueva política, una guerra, un terremoto, etc.; y (*ii*) la duración del desvanecimiento del régimen, que es el tiempo entre la perturbación y el momento en que emerge un nuevo régimen, como una mejor representación de la distribución espacial del atributo. Se ofrecen dos ejemplos: el primero utiliza el PIB *per cápita* provincial de China entre 1978 y 2008, y el segundo utiliza precios de la vivienda a nivel estatal y datos de la tasa de desempleo de los EE.UU. entre 2002 y 2012.

要約: 本稿では、空間の再配分のプロセスの重要な要因を2つ決定するのに、探索的な空間時間分析ツールを導入する。すなわち、（i）新しい政策、戦争、地震など空間の再配分のきっかけとなるショック、（ii）レジームが喪失するまでの時間、すなわち、ショックから属性の空間的分布をよりよく示すものとして、新しいレジームが現れるまでの時間である。ここでは、2つの事例を示す。最初の事例では、1978年から2008年にかけての中国の省における一人当たりGDPを使用する。次の事例では、2002年から2012年までの米国の州別の住宅価格と失業率データを使う。