Dynamics of Regional Business Cycles in Colombian Cities: Insights into Growth Rates, Recession Diffusion, and Cluster Behavior

JUAN DAVID VALENCIA BELTRÁN

Universidad del Valle
Facultad de Ciencias Sociales y Económicas
Maestría en Economía Aplicada
Santiago De Cali, Colombia
2024

Dynamics of Regional Business Cycles in Colombian Cities: Insights into Growth Rates, Recession Diffusion, and Cluster Behavior

JUAN DAVID VALENCIA BELTRÁN

Trabajo de investigación presentado como requisito para optar al título de:

Magíster en Economía Aplicada

Director:

Julián Fernández Mejía

Profesor Asistente de la Facultad de Ciencias Económicas y Administrativas

Pontificia Universidad Javeriana, Bogotá D.C. - Colombia

Universidad del Valle
Facultad de Ciencias Sociales y Económicas
Maestría en Economía Aplicada
Santiago De Cali, Colombia
2024

Dynamics of Regional Business Cycles in Colombian Cities: Identification of Growth Rates, Recession Propagation, and Cluster Analysis

Juan David Valencia Beltrán[1]

Abstract: Using a Markov switching panel, this research addresses the missing literature on the business cycle dynamics in 23 cities and metropolitan areas in Colombia. First, we seek to identify growth rates during periods of expansion and recession using a model based on employment growth, and economic covariables defined a priori for each city and metropolitan area, finding heterogeneous values in these rates. Secondly, cluster analysis identifies cities with shared economic patterns, finding some factors that help explain these clusters. Finally, conducting an analysis reveals how historical recessions propagate in the country based on the identified clusters. The results highlight the diversity in growth rates and the presence of common factors during recessions in different cities, providing a comprehensive understanding of the complexities in urban economic dynamics in Colombia. This approach contributes to more effective policy formulation by considering regional variability in response to economic cycles.

JEL Classification: E32, R11, R12, C55, E31

Keywords: Business Cycle, Regional dynamics, Cluster analysis, Economic recession

Resumen: Utilizando un panel de conmutación de Márkov, esta investigación aborda la literatura ausente sobre la dinámica del ciclo económico en 23 ciudades y áreas metropolitanas de Colombia. En primer lugar, se busca identificar las tasas de crecimiento durante los periodos de expansión y recesión utilizando un modelo basado en el crecimiento del empleo, y variables económicas definidas a priori para cada ciudad y área metropolitana, encontrando valores heterogéneos en estas tasas. En segundo lugar, el análisis de clústeres identifica ciudades con patrones económicos compartidos, encontrando algunos factores que ayudan a explicar estos clústeres. Por último, la realización de un análisis revela cómo se propagan las recesiones históricas en el país en función de los clústeres identificados. Los resultados resaltan la diversidad en las tasas de crecimiento y la presencia de factores comunes durante las recesiones en diferentes ciudades, proporcionando una comprensión integral de las complejidades en las dinámicas económicas urbanas en Colombia. Este enfoque contribuye a una formulación de políticas más efectiva al considerar la variabilidad regional en respuesta a los ciclos económicos.

Clasificación JEL: E32; R11; R12

Palabras clave: Ciclo de negocios, Dinámicas regionales, Análisis de clústeres, Recesión económica

---

[1] Email address: juan.d.valencia@correounivalle.edu.co

Table of Contents

1. Introduction

Compared to other countries with similar landmass and population, Colombia exhibits a stunning array of cultural and geographic variations (Hudson & Library of Congress, 2010). This diversity likely translates into a series of regional economies with differentiated economic structures. As a result, these economies may exhibit varying responses to economic shocks, rather than a uniform national pattern.

Despite existing literature on identifying regional clusters in Colombia, these studies remain inconclusive. This is likely due to the variation in methodologies, geographic units of analysis, and data employed across the studies. Besides, a significant portion of research on the behavior of regional economies in Colombia lacks a proper methodology for cluster identification. These studies often rely on traditional definitions of regions based on shared geographical characteristics, which can introduce selection bias into the results by failing to account for economic factors during the regional clustering process. Furthermore, research in Colombia that simultaneously focuses on the endogenous characterization of regional clusters and the behavior of the regional business economic cycle is scarce. This is despite the significant development of regional business cycle research in recent decades across the globe, with the irruption of supranational entities such as the European Union, or in some cases, in countries where their economy is approached as a set of interconnected regional economies, as in the case of the United States (Hamilton & Owyang, 2012), and the United Kingdom (Fingleton et al., 2012).

This study aims to conduct a comprehensive analysis of the economic cycles in Colombian regions, focusing on cities and metropolitan areas. Unlike previous research, it takes a novel approach by first endogenously identifying clusters of urban entities. This clustering is based on analyzing changes in employment growth rates (a key dependent variable of the economic cycle) alongside other measurable characteristics of each city and metropolitan area. This study is the first, to our knowledge, to allow for the possibility that regions may belong to multiple clusters. This approach provides a more accurate representation of economic complexity and the interconnectedness between these units of analysis. Furthermore, the study will establish the average growth rates during both contraction and expansion phases of the economic cycle for each city and metropolitan area. Additionally, it will analyze the inflow and outflow behavior within these clusters during recessionary periods. These analyses aim to reveal the underlying patterns observed in certain cities. To achieve this, we leverage labor occupation growth data from the National Administrative Department of Statistics of Colombia spanning from the moving quarter of January-March 2007 to April-June 2023. This dataset covers 23 cities and metropolitan areas and serves as input for the Markov-switched panel approach developed by Hamilton and Owyang (2012). Employing Bayesian methodologies, this approach incorporates both cross-sectional and longitudinal time-series dimensions for comprehensive analysis.

Our analysis using the chosen methodology revealed three key findings. First, cities where logistics and accommodation sectors employ a significant portion of the workforce tend to be more resilient during economic downturns. Conversely, cities and metropolitan areas with a high concentration of industry exhibit lower volatility in both growth and recessionary periods. These characteristics explain the observed heterogeneity in growth rates across different economic cycles in various cities and metropolitan areas.

Second, the model identified four clusters, revealing that most cities and metropolitan areas belong to multiple clusters with varying probabilities. This suggests that most cities and metropolitan areas tend to follow a similar business cycle, though not exclusively.

Third, the data show that cities and metropolitan areas experience recessions more frequently than national recessions. Additionally, the Great Recession of 2008 initially impacted all cities and metropolitan areas, with some recovering faster than others. In contrast, the COVID-19 pandemic caused a recession in all cities and metropolitan areas, followed by a synchronized recovery. In both cases, the data suggest that contractions in cities and metropolitan areas began before the national recession dates recognized by economic literature.

The findings of this study bear significant implications for regional economic policy. As Beraja et al. (2024) point out, it is essential to consider both regional and aggregate fluctuations to comprehensively understand national business cycles, given the notable disparities in economic conditions among regions. In Colombia, as observed in their analysis, divergent responses are noted at the regional level compared to the national economy in terms of the recurrence and duration of business cycles. Highlighting the role of employment types in cities in shaping local economic cycles, examining these varied regional responses becomes crucial. In this regard, it is important to consider the specific characteristics of each region's business cycle behavior when formulating economic policies. Doing so would enable governments to implement differentiated strategies of countercyclical measures during times of recession in specific sets of cities. As a basis for this approach, it is proposed to begin with factors that allow some cities to be more resilient or to recover more swiftly from recessionary periods. Simultaneously, efforts should be made to enhance these characteristics in cities that possess them.

The rest of this document is structured as follows: Section 2 discusses regional business cycles and presents clustering studies in Colombia. Section 3 delves into statistical methods for data analysis. Section 4 analyzes the data, considering the impact of the COVID-19 pandemic. Section 5 scrutinizes the obtained results, while Section 6 concludes these findings.

2. Literature Review

2.1. Regional business cycles

In recent decades, the study of regional cycles has gained increasing significance due to the ongoing formation of supranational organizations (e.g., the European Union) and the emerging perspective that countries are part of interconnected regional economies (Hamilton & Owyang, 2012). A central basis concerning the treatment of regional economies was postulated by Barro et al. (1991) based on the neoclassical growth model. They sought to understand regional economic convergence, indicating that poorer economies grow at faster rates than richer ones, thus reducing regional disparities over time. Factors such as labor and capital mobility can influence the speed of convergence. Their findings indicate that achieving gradual convergence may require a significant amount of time.

After studying regional economic cycles in the United States, Carlino and Sill (1997) highlighted how economic theory approaches regional fluctuations. They identified two distinct lenses through which these fluctuations are viewed. The traditional perspective suggests that regional economic changes are temporary, while an alternative viewpoint posits that the duration of these changes can be extended, depending on the region and the nature of encountered shocks. Their study involved distinguishing cycles from trends across eight regions and identifying a core group characterized by closely correlated economic cycles. They concluded that the observed regional cyclical variations have significant implications for regional economic policies due to varying levels of volatility among the regions and their differing reliance on economic cycles.

According to Cárdenas et al. (2015), studies in the New Economic Geography (NEG) highlight the presence of economies of scale in specific regions, promoting economic development. This development subsequently impacts employment levels. Furthermore, according to equilibrium and disequilibrium theories, differences in unemployment rates are attributed to constraints on labor mobility and the presence of incentives specific to particular areas These perspectives emphasize the role of labor mobility constraints and regional incentives in shaping disparities in unemployment rates.

Empirical results support the existence of this phenomenon at various scales; for example, Wall (2007) employed a Bayesian approach to identify turning points in Japan's business cycles between 1976 and 2005, revealing regional differences in growth rates during periods of expansion and recession. The study, based on the Index of Industrial Production, highlights a decline in both average growth and rate variations. This suggests the influence of regional industrial composition. The regional economies approach also applies to cities; illustrating this, Owyang et al. (2008) treat them as cross-sections with similar characteristics. The analysis highlights differences in economic characteristics that influence employment growth during high and low phases. Hamilton and Owyang (2012), across multiple studies, found that business cycles in the United States found notable variation in expansion and recession phases, with similar findings for U.S. states, particularly in regions heavily reliant on oil production and agriculture, which were more susceptible to economic downturns compared to others. In the U.K., Fingleton et al. (2012) found differences in the resilience of 12 U.K. regions from 1971 to 2010. For this, they used a vector error correction model (VECM), finding that the interlinkages of these regional economies are revealed in how they react to and recover from recessionary shocks. In the case of the European Union, it is important to analyze not only at the national level but also at the subnational level. For instance, Gadea-Rivas et al. (2018) researched 213 European Union regions, focusing on sectoral employment to understand the regional composition, and utilizing the percentage of manufacturing employment as a proxy for industrialization. They also considered the number of establishments and population in their assessment of economic policies. Conversely, Wang et al. (2023) used the business cycle of Chinese provinces using industrial linkage as a reference to validate their resilience in the economic cycle, finding differences in the timing of recession between them and the national economy.

In the Latin American context, a study by Haddad et al. (2002) in Brazil used a general equilibrium analysis approach to examine the economic interconnections between regions. The research highlighted the imbalances in international trade among regions, which could be due to the implementation of regional trade agreements. These agreements may also worsen the differences between states over time. Kondo (2022) also investigated regional cycles in Mexico by studying the indirect spatial effects through a Markov switching model for a data panel. The study focused mainly on Mexican states that showed an increased synchronization with the country's overall economic cycle.

2.2. The Link between Economic Cycles and Employment

Multiple studies have linked some employment rates used as a dependent variable in the economic cycle (Zuccardi, 2002; Bohórquez et al., 2009; Hamilton & Owyang, 2012; Gadea-Rivas, 2018). This relationship was pointed out by Rissman (1999), who highlighted the importance of investigating how fluctuations in regional employment affect the economic cycle, developing a model that uses a Kalman filter to infer a common factor which differentially affects different regions. She concluded that there is a significant relationship between employment at the regional level and the business cycle in the United States.

Bohórquez et al. (2009) assert that, within the business cycle theory, researchers have identified various relationships among variables like consumption, investment, and employment across the temporal dynamics of GDP. These relationships take shape within concepts such as rational expectations and Walrasian equilibrium. Beraja et al. (2019) clarify the connection between employment and the business cycle, treating employment as a sensitive indicator to economic fluctuations. Their study, spanning both aggregate and regional data analysis, confirms that changes in domestic demand in the United States significantly impact national-level employment. The employment growth rate proves to be a practical tool for examining regional business cycles. Utilizing a Neo-Keynesian model, the researchers identify ordinary equations between regional and aggregate economies during the Great Recession, emphasizing the equality of wage Phillips curves under specific assumptions. Incorporating regional data into the overall model enhances the understanding of wage dynamics, uncovering notable variations in the flexibility of wages within the context of employment and the business cycle.

2.3. Studies on regional economy based on the analysis of clusters in Colombia.

A review of academic literature on regional economic analysis in Colombia, focusing on cluster identification among entities, reveals a lack of convergence in defining cluster regions. This lack of convergence stems from variations in methodological approaches, the regions studied, and the characteristics of the data employed. One of the earliest studies was conducted by Barón (2002), who utilized the level of deposits in the financial system as input data. This data was collected quarterly for 24 departments, employing a process called *VARCLUS*, available in the *SAS* statistical software, to group the variables into hierarchical clusters.[2] However, it's important to note that this method was exploratory and descriptive rather than inferential. Zuccardi (2002) also attempted to characterize a set of clusters applied to cities. To do so, he used the labor occupation growth rates of seven metropolitan areas in Colombia between 1986 and 2000, evaluating their average annual growth level. He employed a cohesion force that captured the degree of co-movement between national economies, finding that the areas did not converge in an equilibrium relationship.

Bohórquez et al. (2009) investigated how employment rates link cities and the national economy. Using unrestricted VAR models, they analyzed employment cycles in 13 metropolitan areas. This revealed two key findings. First, they identified cities with more adaptable employment trends, categorized as either less influenced by national factors or more synchronized with other cities. Second, the study explored how geographic proximity and population exchange influence regional interdependencies. They discovered that larger cities often precede or influence the employment cycles of smaller, nearby cities. In some cases, the cycles of larger cities could even accurately predict employment patterns in the smaller cities they impact.

Quintero Otero and González (2012) investigated the impact of monetary policy in the Caribbean, Central Andean, and Pacific regions of Colombia, using geographical criteria to define these regions beforehand. They used panel data of quarterly GDP at the national level and the annual growth of the cyclical component

[2] Their methodology identified 7 clusters, but when two departments were identified as independent clusters, the author finally decided to add them to the already established clusters.

of regional GDP between 2002 and 2010. Their findings show that monetary policy shocks affect the economic cycle in these regions in a non-uniform manner, with the Caribbean region being the most affected. They found that monetary policy has a differential impact in regions with greater export openness and those with greater importance in the mining sector. However, no differential effect was observed in departments with high industrial production. The methodology used overcame the limitation of the annual series of regional data by employing direct projections instead of the standard structural vector autoregression approach.

Cárdenas et al. (2015) analyzed structural factors impacting unemployment in 23 cities and metropolitan areas using Multiple Factor Analysis for Multiple Contingency Tables (MFACT). They employed 185 variables to identify underlying factors influencing employment differences, focusing on age structure and labor participation incentives. Their approach revealed five axes via Ward's criteria, showing that cities with high unemployment did not consistently share the same structure, contrary to prior studies. However, the study lacked assumptions about the multivariate distribution under the MFACT model, limiting statistical inferences. Six distinct clusters emerged, each with unique characteristics, highlighting the diversity in unemployment determinants. These clusters did not exhibit a shared geographic distribution, except for a cluster comprising major cities along the Colombian Caribbean coast, which had low unemployment rates, high labor participation, and large household sizes.

Meanwhile, Gómez Sánchez et al. (2022) focus on analyzing Colombia's regional cycles, dividing the country into four geographically continuous regions in their view, with a fifth region covering areas with the highest economic activity (Bogotá D.C., Valle del Cauca, and Antioquia). Their objective is to highlight the importance of the relationship between regional business cycles and manufacturing productivity in Colombia. Using a methodological approach that combines econometric models such as the system of general methods of moments, the Kalman filter, and VARX models, they conclude that there is a pro-cyclical and non-contemporaneous causal relationship from productivity to GDP in the more developed regions, while in other regions, the relationship is counter-cyclical and non-contemporary, indicating that economic policy should support the manufacturing sector with subsidies, access to capital and promotion of international trade to improve competitiveness.

3. Methodology

The methodology developed by Hamilton and Owyang (2012), which has gained recognition for its effectiveness in analyzing panel data with covariates, utilizes an exploratory method. This approach has been applied in various studies characterizing regional business cycles (Hernández-Murillo et al., 2017;

Kondo, 2022; Wang et al., 2023). Unlike predefined groups, geographic units are clustered based on their own economic characteristics. This allows researchers to identify shared features and analyze regional recessions specific to each resulting cluster. This approach has revealed significant variations in the expansionary and recessionary phases of regional economic cycles across US states.

Hamilton and Owyang (2012) defined co-movement in the commonality of regions based on regional clusters derived from employment growth rates. They delineated the business cycle as consisting of discrete transitions between expansion and recession phases. The methodology involves a probabilistic framework based on time series models and stochastic processes, emphasizing Bayesian inference. The practical application utilizes Gibbs sampling to sequentially generate values for variables of interest, such as latent states, transition matrices, and structural parameters. A cross-validation approach is employed to evaluate model performance, addressing steps such as regime generation and parameter updating using Markov chain Monte Carlo. This methodology facilitates data grouping into regional clusters, considering the correlation between employment growth rates and covariates. The probability of belonging to each cluster is modeled as a logistic variable, and inference is then drawn based on the observed regional correlations in employment growth.

3.1. Markov model and Bayesian analysis.

This section presents the Hamilton and Owyang (2012) approach for identifying clustered business cycles and determining endogenous divisions within clusters.

Let y_{tn} denote the employment growth rate in region n at time t, and y_t represents an $(N \times 1)$ dimensional vector comprising all employment observations for the N geographic regions at time t. Additionally, let s_t be an $(N \times 1)$ dimensional vector following a Markov chain containing recession indicators, where $s_{tn} = 1$ when region n is in recession and $s_{tn} = 0$ when it is in expansion. Consider Eq. 1:

$$y_t = \mu_0 + \mu_1 \odot s_t + \varepsilon_t \tag{1}$$

During expansion phases, μ_0 signifies the mean employment growth in region n on average, and $\mu_0 + \mu_1$ represents the mean growth experienced during recessions phases. The operator $\odot$ denotes the Hadamard product, a mathematical operation entailing the pairwise multiplication of corresponding elements. According to the model, recessions phases are proposed as the main driver of changes in job creation within regional units. The error term ε_t is modeled to conform to a normal distribution with mean zero and variance Ω as outlined in Eq. 1. Another assumption is that the Ω is diagonal, which reduces the number of variance parameters. This is necessary for the algorithms used in the analysis.

$$\Omega = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma_N^2 \end{bmatrix} \tag{1}$$

The model suggests that simultaneous economic downturns drive correlated employment growth across regions. While assertive, this perspective offers insight into how the business cycle spreads regionally. Addressing the model's complexity, arising from the substantial number of states and intricate relationships, requires additional simplifications; hence, a cluster-based approach is employed. The premise posits that the recession process can be grouped into a small number of clusters $K << 2^N$, alongside an aggregate indicator $z_t \in \{1,2,\dots,K\}$ that indicates if a recession is happening in cluster k at a given time t. Determining these cluster patterns involves considering a h_{nk} cluster index, where the nth component equals 1 if region n is affiliated to cluster k, and 0 otherwise. All the regions belonging to cluster k are in recession when $z_t = k$. Regional dynamics are delineated through Eq. 1:

$$y_t \mid z_t = k \sim N(m_k, \Omega) \tag{1}$$

Describing m_k, as in Eq. *2*:

$$m_k = \mu_0 + \mu_1 \odot h_k. \tag{2}$$

In a Markov-switching model, states $h_1, \dots, h_K$ are crucial for inference but estimating them from data is challenging. We set priors: $h_K = 0$ (expansion) and $h_{K-1} = 1$ (recession). Other states (idiosyncratic) are grouped as $\kappa = K - 2$. The values of $h_1, \dots, h_\kappa$ are unobserved but influence the probability distribution of the observed data $\{y_t\}_{t=1}^T$. Subsequently, the existence of a vector x_{nk} with dimensions $(P_k \, x \, 1)$ of covariates influencing whether a region n experiences a recession $z_k = k$ is treated through a logistic model in Eq. 3:

$$p(h_{nk}) = \begin{cases} 1/[1 + exp\,(x'_{nk}\beta_k)] & if\, h_{nk} = 0 \\ exp\,(x'_{nk}\beta_k)/[1 + exp\,(x'_{nk}\beta_k)] & if\, h_{nk} = 1 \end{cases} \tag{3}$$

The parameter β_k is interpreted as a logistic clustering parameter for the entire population. Realizing Eq. 3 is impossible at this stage because the values h_k are unavailable. Therefore, latent variables ξ_{nk} and ψ_{nk}, are introduced, defined as follows: $\xi_{nk} = x'_{nk}\beta_k + 2\psi_{nk}e_{nk}$, where $\psi_{nk} \sim KS$ and $e_{nk} \sim N(0,1)$. The variable ξ_{nk} takes on a logistic distribution with a mean determined by $x'_{nk}\beta_k$. The cumulative distribution function of ξ_{nk} is given in Eq. 4:

$$Pr(\xi_{nk} \le z) = \frac{1}{1 + exp(x'_{nk}\beta_k - z)} \tag{4}$$

Therefore, it is found Eq. 5:

$$Pr\ (\xi_{nk} > 0) = \frac{exp\ (x'_{nk}\beta_k)}{1 + exp\ (x'_{nk}\beta_k)} \tag{5}$$

In simpler terms, if consider ξ_{nk} generated from $N(x'_{nk}\beta_k, \lambda_{nk})$ distribution where $\lambda_{nk} = 4\psi_{nk}^2$ for $\psi_{nk} \sim KS$ and set h_{nk} to 1 when $\xi_{nk} > 0$, it is like saying that h_{nk} was determined based on the probability specified in Eq. Error! Reference source not found.. In the realm of data analysis, the objective lies in d eriving Bayesian posterior inferences for both population parameters and unobserved latent variables. Within this study, these variables are categorized into distinct groups. The regional economic growth is represented by the indicator $\theta = \{\mu_0, \mu_1, \Omega\}$ and a transition probability matrix $\boldsymbol{P}$. There are also two types of latent variables exist: a vector $z = (z_1, \ldots, z_T)'$ indicating recession clusters at each date, and $h = \{h_1, \ldots, h_\kappa\}$ indicating the affiliation of each geographic region, where $h_k = (h_{1k}, \ldots, h_{Nk})'$ characterizes regions in cluster k. Variables $\xi_k = (\xi_{1k}, \ldots, \xi_{N\kappa})'$ and $\lambda_k = \{\lambda_{1k}, \ldots, \lambda_{N\kappa}\}'$ are auxiliary variables determining h_κ, as per Eq. 6:.

$$h_{nk} = \begin{cases} 1 & \text{if } \xi_{nk} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

$$\xi_{nk} \mid \beta_k, \lambda_{nk} \sim N(x'_{nk}\beta_k, \lambda_{nk})$$

Where $\xi_{nk} \mid \beta_k, \lambda_{nk} \sim N(x'_{nk}\beta_k, \lambda_{nk})$, $\lambda_{nk} = 4\psi_{nk}^2$ and $\psi_{nk} \sim KS$. All unobserved variables related to cluster memberships are represented by $H = \{h, \xi, \lambda\}$, in this context $\xi = \{\xi_1, \ldots, \xi_\kappa\}$, $\lambda = \{\lambda_1, \ldots, \lambda_\kappa\}$, and $\beta = \{\beta_1, \ldots, \beta_\kappa\}$ signifies the collection of logistic coefficient vectors. Subsequently, Bayesian inference is employed for parameter estimation. To initiate this process, a series of initial assumptions are required by the model, as detailed in Table 1:

Table 1. Foundations for Estimating Parameters.

	Parameter	Prior Distribution	Hyperparameter	
Average employment growth	$[\mu_{0n}, \mu_{1n}]'$	$N(m, \sigma^2 M)$	$m = [0.5, -1]'; M = I_2$	$\forall n$
Variance of ε_t	σ_n^{-2}	$\Gamma\left(\frac{v}{2}, \frac{\delta}{2}\right)$	$v = 0; \delta = 0$	$\forall n$
Transition probability matrix	P	$D(\alpha)$	$\alpha_i = 0$	$\forall i$
Logistic clustering parameter	β_k	$N(b_k, B_k)$	$b = 0_p; B = \frac{1}{2} I_p$	$\forall k$

Notation: The number of covariates is denoted by p denotes the number of covariates, denoted as x; $n \in \{1, \ldots, N\}$ corresponds to the geographic region; $k \in \{1, \ldots, K\}$ signifies the quantity of clusters; normal priors are established for both μ and β, inverse gamma priors for σ_n^2 , and a Dirichlet distribution serves as the prior for P.

The simulation's growth assumptions involve an average annual employment growth rate of 0.5% during an expansion, with a probable range from -0.5% to 1.5%. Conversely, the average employment growth in a recession is expected to range from -2% to +1%. These figures approximate Colombia's labor participation growth rate between 1983 and 2014, as summarized by Arango et al. (2016).

The following steps generate the joint distribution of variables, laying the foundation for modeling. Subsequently, the discussion delves into generating random samples, providing detailed explanations on obtaining the necessary values. This process encompasses generating fundamental parameter samples, transition matrices, and latent states of the model, addressing essential issues like identifiability, and handling uniqueness problems in parameter estimations and transition matrices. The estimation of the model depends on a predetermined number of clusters, denoted as K, posing a typical model selection challenge. This strategy utilizes k-fold cross-validation and the marginal likelihood function to establish K. The dataset Y_t is split into $R = 2$ folds for cross-validation. Evaluation utilizes a score measuring the discordance between observations and predictions, employing an entropy loss function. The score is computed for each block, aiming to minimize it by weighing the difference with the inverse of the covariance matrix. The selection of the cluster count aims to minimize this overarching score across all cross-validation iterations. We will set the range of clusters from $K = 2$ *to* $K = 5$ to be the minimum and maximum number of clusters tested. For each value of K, the Gibbs sampler will be run for 2,500 iterations. An initial set of 25,000 iterations will be discarded as they represent a warmup phase.

4. Data

While other works have explored the utilization of GDP as a more suitable dataset to validate growth or decline in the economic cycle, as evidenced in several research studies (Barón, 2002; Hamilton & Owyang, 2012; Kondo, 2022; Wang et al., 2023), challenges arise in obtaining this data, often due to a lack of information or limitations in the number of observations. In Colombia, the National Administrative Department of Statistics (DANE, abbreviated in Spanish) only publishes annual departmental GDP values since 2005, with coverage excluding some departments. Consequently, the employment growth rate emerges as the alternative proxy, aligning with literature indicating a robust correlation between employment growth and the economic cycle. However, the availability of departmental employment growth data is limited to annual records starting from 2001, making it challenging to study due to a scarcity of observations. In light of this limitation, we use data from 23 cities and metropolitan areas, offering advantages with its monthly provision corresponding to a moving quarter, spanning January-March 2007 to June-August 2023. This approach yields more observations than the limited annual records and encompasses 23 departments out of 32, ensuring a comprehensive dataset derived from the Integrated Household Survey conducted by DANE (2023).[3]

The panel data set used was the rate of employment growth in employment defined by ratio 2:

$$Employment\ rate = \frac{Employed\ people}{Labor\ Force} \tag{2}$$

This rate considers the number of employed persons divided by the labor force, composed of those of working age who are employed or seeking employment.[4] The employment growth rate is given by ratio:

$$Employment\ growth\ rate = \frac{Employment\ rate\ year\ 1}{Employment\ rate\ year\ 0} - 1 \tag{3}$$

The model also requires a set of covariates associated with each city and metropolitan area that characterize the pre-event probability of belonging to a given cluster. For this purpose, we identify four specific covariates: the percentages of individuals employed in manufacturing industries, financial and insurance activities, logistics and accommodation activities, and the registered companies per 1000 inhabitants.[5] The selection of the manufacturing employment share is based on its role as a primary indicator of a city's productive capacity. In contrast, opting for personnel engaged in financial and insurance activities reflects a locality's economic dynamism concerning financial investment, which is crucial for resource allocation and assessing the city's ability to mobilize capital. These two factors align with the choices made by Hamilton and Owyang (2012). Furthermore, the employment proportion in logistics and accommodation, which covers the *transportation and storage* sector as well as the *accommodation and food services* sector, contributes to better classification and clustering since these sectors often concentrate in specific cities. Finally, the number of companies per inhabitant is a relevant metric as it diversifies the economic structure, reduces dependence on specific sectors, and strengthens the city's ability to adapt to economic changes.

---

[3] Technical information on the Integrated Household Survey (GEIH for its abbreviation in Spanish) is available in Spanish at https://t.ly/tsXRT. The cities and metropolitan areas covered include Bogotá, Cali M.A. (Cali and Yumbo), Medellín M.A. (Barbosa, Bello, Caldas, Copacabana, Envigado, Girardota, Itagüí, La Estrella, Medellín and Sabaneta), Barranquilla M.A. (Barranquilla and Soledad), Bucaramanga M.A. (Bucaramanga, Floridablanca, Girón and Piedecuesta), Manizales M.A. (Manizales and Villamaría), Pasto, Pereira M.A. (Dosquebradas, La Virginia and Pereira), Cúcuta M.A. (Cúcuta, Villa del Rosario, Los Patios, Puerto Santander and El Zulia), Villavicencio, Ibagué, Montería and Cartagena de Indias.

[4] DANE defines the employment rate or "*Tasa de Ocupación* (TO)" as the ratio of employed people to the working-age population (15 years or older). However, this definition includes working-age individuals not actively seeking employment, leading to a distorted view of actual employment. To address this, the use of the ratio 3 is adopted, focusing specifically on those actively engaged in employment or actively seeking work. This aligns with DANE's calculation of the unemployed population (Unemployed / Labor Force). Note that the employment rate serves as a measure of the demand for employment.

[5] Percentages of employed population in each sector were extracted from the Integrated Household Survey (DANE, 2023), depicting monthly averages from the moving quarter (January-March 2015 to April-June 2023). Calculating the number of companies per 1,000 inhabitants involved dividing registered companies in each city/metropolitan area from the Business Statistical Directory Bulletin 2019-2021 (retrieved from https://t.ly/L8Pbp) by the 2021 population forecast from the country's latest census (retrieved from https://t.ly/4Y9-B) and then multiplying the result by 1,000.

4.1. Seasonal adjustment of the data

The data provided by DANE are not seasonally adjusted, requiring the removal of seasonal effects. We apply one of the latest methods, X-13ARIMA-SEATS, a statistical technique for seasonal adjustment and time series analysis created by the U.S. Census Bureau — the government agency responsible for generating demographic and economic data for the United States.[6] This method decomposes the time series into trend, seasonality, and irregularity components, using iterative techniques of centered moving averages, providing a seasonally adjusted series with quality tests and statistics. For each of the 23 series, a comprehensive validation of the seasonally adjusted time series was conducted using various statistical tests. The normality of the data was assessed with the Doornik-Hansen test, analyzing the distribution of residuals. Additionally, the linearity of residuals was examined with the Ljung-Box test on squared residuals, and the independence of seasonal residuals was evaluated through the Ljung-Box test on the autocorrelation of these residuals. These tests, addressing critical aspects such as normality, linearity, and independence, ensured the robustness and reliability of the seasonally adjusted time series—essential for subsequent analysis.

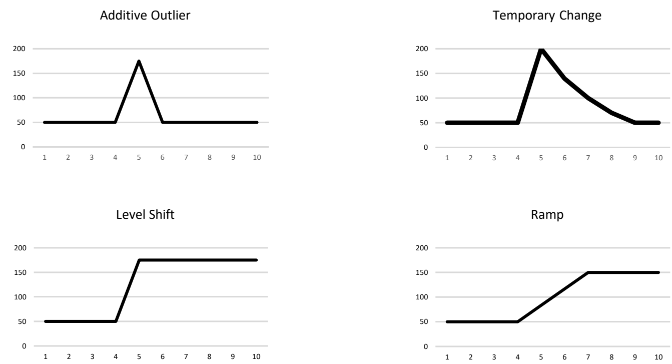
4.1.1. Covid-19 seasonal effects

The COVID-19 pandemic has significantly altered the behavior of various economic indicators. Consequently, addressing its impact on data seasonal adjustment becomes imperative. To accomplish this, we will adopt the intervention analysis methodology proposed by Foley (2021). This method decomposes the time series into trend, seasonality, and irregularity components, using iterative techniques of centered moving averages, providing a seasonally adjusted series with quality tests and statistics. For each of the 23 series, a comprehensive validation of the seasonally adjusted time series was conducted using various statistical tests. The normality of the data was assessed with the Doornik-Hansen test, analyzing the distribution of residuals. Additionally, the linearity of residuals was examined with the Ljung-Box test on squared residuals, and the independence of seasonal residuals was evaluated through the Ljung-Box test on the autocorrelation of these residuals. These tests, addressing critical aspects such as normality, linearity, and independence, ensured the robustness and reliability of the seasonally adjusted time series—essential for subsequent analysis.

This method is designed to address distinct anomalies in series data and is employed to mitigate the effects of non-seasonal events on the seasonal pattern of economic data. Figure 1 shows the four most common anomalies or outliers in a series.

[6] More information about the X-13ARIMA-SEATS method is available at https://t.ly/YyXsN.

Figure 1. Main types of outliers

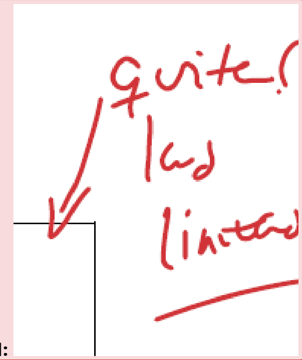


Source: Own elaboration based on Foley (2021).

Additive Outliers (AO) represent abnormal values that deviate significantly from the typical series pattern, mainly occurring at isolated points or the final point of a time series. Temporary Changes (TC) indicate a transient shift in the series level, characterized by an immediate change followed by a swift return to the original state, often influenced by factors like holidays, or one-time events. Level Shifts (LS) suggest a sudden and sustained alteration in the baseline level, often stemming from structural changes such as policy adjustments or economic shifts. Lastly, linear ramps (RP), commonly employed for modeling gradual trends or shifts in data over time, resemble LS but facilitate a smoother transition between levels over a specified period (Foley, 2021).

For the seasonal adjustment of data and the detection of outlier anomalies, we utilize the open-source seasonal adjustment software *JDemetra,* developed by the European Union Statistics Agency - EUROSTAT - and the package developed for *R* software*, RJDemetra.*[7] While the statistical package automatically identifies outliers, Foley (2021) and statistical agencies stress the importance of manual identification during the COVID-19 pandemic.[8] The software frequently overlooks these effects. Consequently, all COVID-19 phase outliers underwent manual analysis for each series, focusing on outlier characteristics, and were then incorporated into the seasonal adjustment model based on the most suitable outlier type determined through series analysis. Figure 2 shows an example of a manual adjustment made to the time series of Santa Marta during the pandemic. Four outliers corresponding to four observation periods and one

Commented [JV1]:



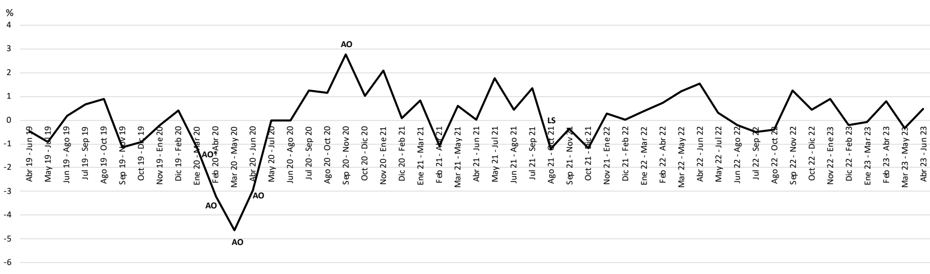
Commented [JV2R1]: Corregido.

[7] Additional information about the software *JDemetra+* is available at https://t.ly/e9IoJ . Package information for the *R* software is hosted at https://t.ly/6gQkL.

[8] Particularly helpful in understanding seasonal management in the context of the COVID-19 pandemic were the lectures given by Eurostat statistical agencies, available at https://t.ly/qCRDV , and Latin American national statistical agencies, available at https://t.ly/lZFbH.

level change were identified manually, while the system automatically detected only one outlier in the illustrated period. It is essential to highlight that retaining data from the pandemic is grounded in the premise that its influence was widespread across all studied areas, justifying its consideration in the research analysis.

Figure 2. Patterns of outliers identified in the time series of the city of Santa Marta.



* Automatically detected by the software

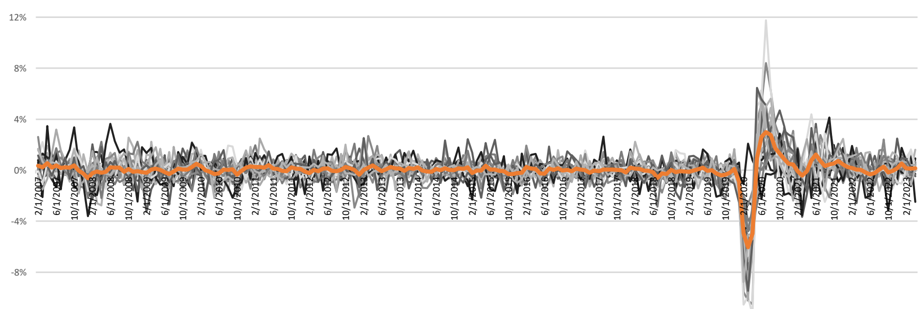
Note: Only the period between the moving quarter April-June 2019 and April-June 2023 is presented. Source: Own elaboration.

5. Results

5.1. Employment growth by city and metropolitan area

Figure 3 depicts seasonally adjusted average employment growth across 23 cities and metropolitan areas. Notably, the pre-pandemic period exhibited stable growth, maintaining an average close to zero. Growth and decline variations were generally modest, rarely exceeding three percentage points, except for the city of Quibdó, which notably emerged as the most volatile during this period.

Figure 3. Average employment growth of 23 cities and metropolitan areas of Colombia



Note: The average of the 23 series is highlighted in orange - source: Own elaboration.

The impact of COVID-19 on data emerged notably from January-March 2020, peaking during March-May 2020 coinciding with Colombia's national quarantine.[9] During the quarantine, the most significant quarterly decrease in city averages was -6.05% in March-May 2023, while the highest growth of 2.99% occurred in July-September 2023. Neiva displayed both the lowest (-11.35%) and highest (11.75%) growth values, showing the highest volatility among cities. Quibdó showed countercyclical behavior, growing during March-May 2020 while others declined, but decreasing in June-August 2023 amid general growth, typical of cities with high labor informality rates (Gómez, 2013). A gradual return to pre-pandemic levels began in December 2021-February 2022, three months after the quarantine ended in Colombia.

5.2. Growth coefficients in periods of expansion and recession.

According to the employed methodology, the iterations yield average growth rates for expansion periods denoted as $(\boldsymbol{\mu}_0)$ and recession periods denoted as $(\boldsymbol{\mu}_0 + \boldsymbol{\mu}_1)$ in each city. Additionally, the methodology calculates the rate of residual variance $(\boldsymbol{\sigma}^2)$ in the growth rates. Figure 4 shows these coefficients.

Figure 4. Average growth coefficients of 23 cities and metropolitan areas of Colombia

Average growth coefficient in the expansion | **Average growth ratio in recession** | **Average residual coefficient in growth rates**

Source: Own elaboration. [10]

Concerning the growth rate during the expansion, no typical geographic pattern is observed, with an average of 1.14% growth among all entities. The city of Barranquilla exhibits the lowest growth value at 0.13%,

---

[9] The national mandatory quarantine in Colombia started on March 25, 2020, according to *Decreto 457 de 2020* (retrieved from https://shorturl.at/cpruC), and concluded on August 30, 2021, as declared by *Resolución 844 de 2020* (retrieved from https://t.ly/Toii4).

[10] The map delineates the territorial boundaries of each city or metropolitan area in red. The same color is applied to the department it belongs to. For Bogotá DC, the urban extension is outlined, in the same color for the capital district extension.

closely followed by Quibdó with 0.16% and Manizales with 0.22%. The noteworthy underperformance of Barranquilla is significant, especially considering its prominent role as an industrial pole in the country (Otero-Cortes et al., 2023). However, it is crucial to note that this city has consistently been characterized by an informality rate exceeding the national average. Additionally, despite the observed dynamism in the city's industrial sector, the jobs generated exhibit significant deficiencies in job security, wages, and benefits (Otero-Cortes et al., 2023); this may be attributed to the fact that the labor supply in Barranquilla has experienced higher growth than the demand for employment (Galvis-Aponte et al., 2019).

In Quibdó, the low growth rate is possibly explained by persistent unemployment rates and the reception of a displaced population in an already precarious labor market (Robledo-Caicedo, 2019). In Manizales, as per the analysis conducted by Guarín (2017), it is evident that, from a productivity-oriented perspective of the industrial system, the city has witnessed a decline in the vitality of coffee plantation activity. Historically, this sector served as the engine of its economy. This decline has resulted in a lack of diversification and sophistication in the productive apparatus, limiting the generation of quality jobs during periods of economic growth.

Among the cities displaying greater dynamism during expansion periods, Florencia leads with a remarkable growth rate of 2.62%. Regrettably, there is a lack of academic studies supporting the city’s boom during expansion periods or analyzing the dynamics of employment within. However, some studies highlight its high unemployment rate at the national level, surpassing the average of the 23 major cities and exhibiting a general upward trend (Robledo-Caicedo, 2020). Notably, Florencia serves as the capital of one of the departments most affected by the internal conflict in Colombia. Thus, it is plausible to theorize that the city’s economy experiences remarkable growth concurrently with the materialization of peace agreements reached by the government with the insurgency (Rivera & Echeverri, 2020).

Next, the city of Armenia boasts a commendable growth rate of 2.22%. Despite limited literature addressing Armenia’s dynamic growth, Barrantes (2021) emphasizes the persistent issue of unemployment in the city over the years. It is crucial to note that Armenia and five other jurisdictions enjoy tax benefits in income tax for new companies engaged in specific activities, facilitated through a special regime known as *Zona Económica y Social Especial* (Special Economic and Social Zone).[11] However, precise figures regarding its impact on employment generation or its contribution to the increased formation of companies in this city remain elusive (Barrantes, 2021). In addition to the regions benefiting from this scheme, Armenia’s central geographical position within the golden “triangle of the country” (a term coined to represent an imaginary

[11] In addition to Armenia, the cities with this condition are Quibdó, Buenaventura, and Barrancabermeja, as well as the departments of Guajira, Norte de Santander, and Arauca. Retrieved from https://t.ly/O80F6 and https://t.ly/o0-kE.

triangle among the three largest cities in Colombia) is noteworthy.[12] Cali ranks third in average growth in expansion cycles, with 2.07%. As noted by González and Mora (2011), the city has distinguished itself by recurrent cycles of economic acceleration and deceleration, incorporating into its labor force not only residents of its metropolitan area but also residents of surrounding municipalities in the region. According to Ortiz and Uribe (2007), Cali has historically been characterized as labor-intensive; therefore, it can more easily benefit from the ample supply of available jobs during periods of economic growth. A discernible pattern is generally observed in these cities, aligning with studies correlating higher unemployment rates and increased employment generation due to demand shocks (Bartik, 2015).

When evaluating the average coefficients during recessionary periods, an overall average growth coefficient of -0.14 is identified across all entities. Despite this, five cities exhibit positive growth (Popayán, Pereira M.A., Riohacha, Monteria, and Santa Marta), albeit all very close to 0% (less than 0.03%). Among the remaining 16 cities, 13 fall within the range of [-0.1%, 0%), with only 5 exceeding a growth coefficient of 0. Ibagué stands out with the lowest growth coefficient of -0.96%, attributed to its elevated unemployment rates compared to the national average, along with its higher overall participation rate and low economic growth (Ayala-García, 2014). Manizales follow this with -0.60% and Quibdó with -0.58%. Furthermore, a positive correlation of 0.31 is observed between the proportion of people employed in logistics and accommodation services and the growth coefficient during recessionary periods; this suggests that cities with a higher focus on tourism may exhibit greater resilience during these times compared to other cities.

The growth rates mentioned earlier should be assessed in conjunction with the residual variability rates of the coefficients. Among the cities, only Quibdó (1.58%) and Ibagué (1.16%) surpass the 1% variability assumption. Considering Quibdó's historically volatile growth rates, the mentioned figure does not appear notably abnormal. However, it sheds light on the nuanced low growth of Ibagué during the recessionary periods. Conversely, Bogotá (0.27%), Barranquilla M.A. (0.30%), and Medellín M.A. (0.36%) exhibit the lowest volatility. Notably, Bogotá and Medellín stand out with employment rate growth during expansion periods (0.72% and 0.79%, respectively). Furthermore, their growth coefficients are close to 0 in recession periods (-0.02% and 0.00%, respectively), indicating a more stable long-term employment growth than other cities. Historically, Bogotá maintained economic growth rates above national averages until the mid-2000s, and while its growth rate has decreased, it has followed a stable trajectory (Guzman et al., 2017).

In the case of Medellín, despite not having lower unemployment rates compared to other cities, studies indicate higher-quality employment and lower informality rates (Sanchez, 2013). These factors, coupled with the city's high industrial density and market potential (Cárdenas et al., 2015), may contribute to a

[12] Definition of the Colombian "golden triangle" - obtained from U.S. Department of Commerce (2021). Santiago de Cali Profile. Retrieved from https://t.ly/YB6MX.

lower loss of labor occupation during economic downturns, underscoring Medellín's strengths. Supported by a negative correlation of -0.50, cities with higher proportions of people employed in manufacturing industries exhibit lower variability in growth rates, indicating greater stability in the growth rate of these cities.

5.3. Identified Clusters

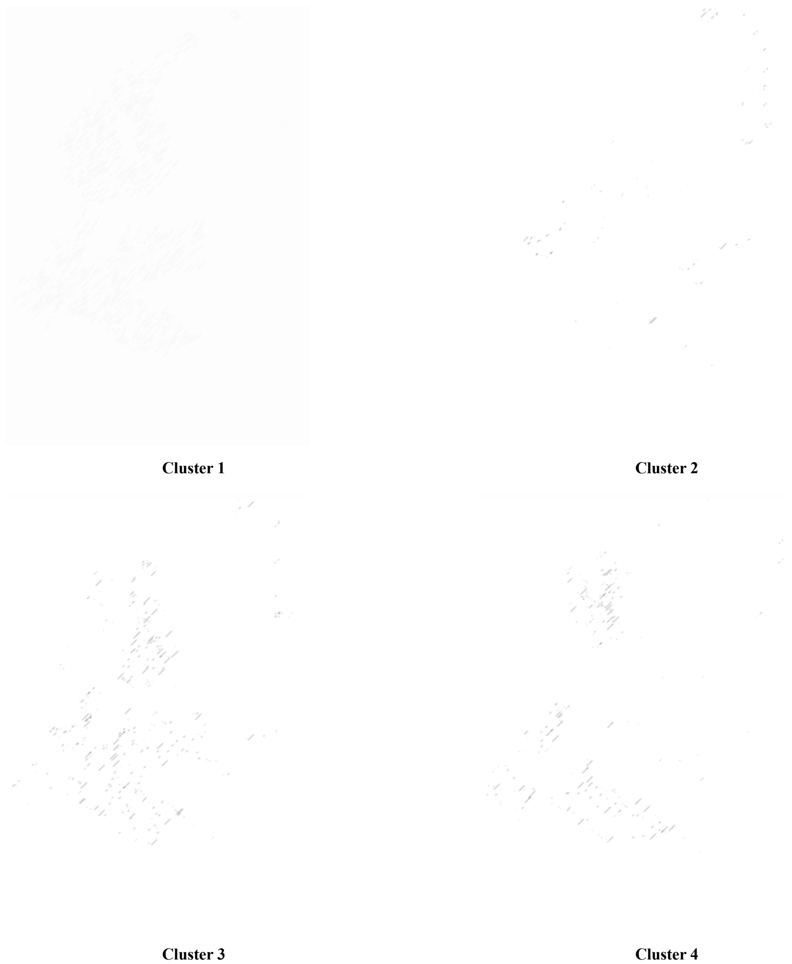
The modeling process identifies clusters of cities and metropolitan areas that share a tendency to synchronize their economic cycles. Four clusters, which are not mutually exclusive, are delineated based on this shared behavior. It's important to note that this likelihood varies across clusters. In our analysis, a city is categorized as belonging to a cluster if its affiliation probability exceeds 0.5.

Cluster 1 encompasses 18 cities with a probability of belonging that exceeds 80%. These cities include Bogotá D.C., Medellín M.A., Cali M.A., Bucaramanga M.A., Pasto, Pereira M.A., Cúcuta M.A., Monterio, Villavicencio, Florencia, Popayán, Valledupar, Neiva, Riohacha, Santa Marta, Armenia, Cartagena, and Sincelejo. Additionally, Tunja has a probability exceeding 50%. Notably, 19 out of the 23 cities analyzed are present within this cluster, indicating a high level of inclusion and suggesting a generalized behavior across most cities. As for Cluster 2, only seven cities surpass 80%: Cali M.A., Bucaramanga M.A., Florencia, Valledupar, Popayán, Bogotá D.C., and Pereira M.A. Furthermore, three cities—Medellín M.A., Quibdó, and Cartagena—exceed 50%. Notably, within this cluster, it is observed that, except for Cartagena, several cities in the Caribbean region and some cities in the country's south deviate from the economic trends.

Cluster 3 exhibits uniform behavior with all cities having a probability of participation in this cluster above 80%. Cluster 4 identifies 15 cities with a probability of belonging that surpasses 80%, including Cali M.A., Florencia, Valledupar, Armenia, Cartagena, Pasto, Santa Marta, Monteria, Pereira M.A., Villavicencio, Popayán, Riohacha, Cúcuta M.A., Neiva, and Sincelejo. It is important to note that this cluster does not encompass several cities in the central region of the country or the city of Barranquilla.

Figure 5 visually represents these clusters on a map, highlighting the probability that each city or metropolitan area is associated with a specific cluster.

Figure 5. Cluster affiliation probabilities identified in 23 cities and metropolitan areas.



Notes: A city or M.A. is assigned to a cluster if it has a posteriori affiliation probability greater than 0.5 for that cluster.[13] - Source: Own elaboration.

The first and fourth clusters exhibit a more national character, as most cities either belong to the same cluster or have a high probability of belonging. Therefore, extracting correlations with covariates in these cases is not feasible. In Cluster 2, a correlation of 0.41 is observed with the proportion of people employed

[13] The map delineates the territorial boundaries of each city or metropolitan area in red. The same color is applied to the department it belongs to. For Bogotá DC, the urban extension is outlined, in the same color for the capital district extension.

in financial and insurance activities and -0.39 with the proportion of people employed in logistics and accommodation activities; this aligns with the low probability of including cities in the Caribbean region, whose economies are heavily reliant on tourism. For Cluster 4, a correlation of -0.58 is found with the proportion of people employed in financial and insurance activities and -0.54 with the probability of belonging to logistics and accommodation. In broad terms, this cluster hosts cities with low levels of financial liquidity, dependence on sectors related to tourism and logistics, and a lower concentration of firms (with an average of 18.78 compared to 24.99 in cities not included in this cluster).

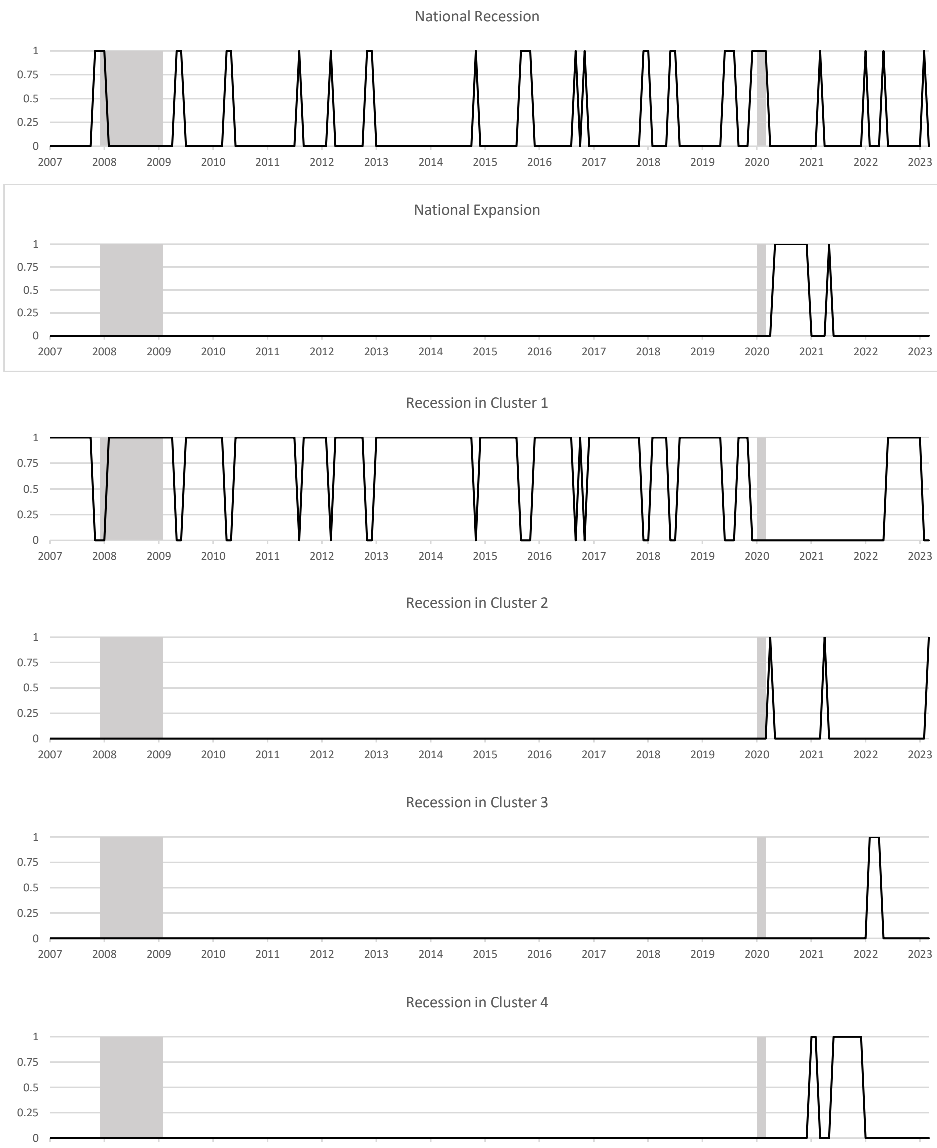
Despite these differences, 19 cities share at least two clusters, and five cities simultaneously belong to all clusters. Therefore, interpreting the results of the cluster groupings should be done cautiously, as there might be some uniformity among some cities composing them, with variations in the probability of belonging to a specific cluster. While Clusters 2 and 3 show some geographical continuity, our investigation focuses on the city and metropolitan area level. Here, economic disparities and city-specific characteristics may differ from those of higher-ranking subnational entities. Additionally, as highlighted by Duranton (2015), the country's road traffic conditions significantly impact trade in cities, which determines that the characteristics of the economy are more crucial than the commercial flow between them. Therefore, this continuity cannot be directly compared with previous studies on subnational administrative regions. Given the considerations above, the results of our study differ from the more exclusionary cluster patterns observed in the United States by Hamilton and Owyang (2012) in their state-level study. Instead, our findings exhibit more similarities to the province-level results in China as reported by Wang et al. (2023).

5.4. Clustered business cycles.

Figure 6 presents the mean probabilities of national expansion, national recession, and recession occurrence for each identified cluster, estimated using Markov panel switching with the data incorporated into the model. It is crucial to emphasize that "national expansion" and "national recession" denote shared states experienced by all cities. Two distinct recession events unfolded in Colombia throughout the study period, each associated with pivotal economic occurrences: the Great Recession from the subprime mortgage crisis and the COVID-19 crisis from 02/01/2020 to 04/01/2020 (Arango et al., 2016; Fedesarrollo, 2020; Posada, 2020). [14]

[14] Alfonso et al. (2011) defined the Great Recession dates, utilizing a weighted diffusion index across various sectors. The beginning of the COVID-19-induced recession was determined by Fedesarrollo (2020) and its conclusion by Posada (2020), referencing the end of the U.S. cycle recession. This aligns with findings by Hernández (2023), utilizing the Bry-Boschan algorithm with Colombia's quarterly real GDP cycle.

Figure 6. Posterior Probabilities of aggregate regimes.



Notes: Posterior Probabilities are monthly, reflecting the moving quarter. Shaded areas align with the periods identified as national recessions according to the NBER methodology. Source: Own elaboration.

Cities in cluster 1 consistently undergo prolonged and recurrent recessionary periods. However, the onset of the COVID-19 pandemic has altered the established similarities in the behavior of the country's cities. Before this event, only four cities—Quibdó, Manizales M.A., Barranquilla M.A., and Ibagué—did not follow the same recessive economic cycle as the rest. Post-COVID-19, the model identifies three additional clusters. Cluster 2 consists of cities that undergo an additional moving quarter of the COVID-19 recession compared to other cities, while nearly half of the cities emerge from this recession. Notably, several cities in the Colombian Caribbean are excluded from this cluster. [15] Cluster 3, demonstrating behavior akin to a national recession, with all cities in this cluster having a probability of belonging exceeding 90%, persists for three consecutive moving quarters starting in early 2022. This occurrence can be viewed as an extension of the national recession that began in February–March 2022. In Cluster 4, two distinct recessionary periods emerge. However, cities such as Quibdó, Medellín M.A., Bucaramanga M.A., Barranquilla M.A., Bogotá D.C., Tunja, Ibagué, and Manizales M.A. deviate from the recessive behavior associated with this cluster, spanning nine moving quarters.

Examining the cluster's behavior during national crisis events reveals that the impact of the Great Recession is evident in all cities during the moving quarter from December 2007 to February 2008, sustaining until February 2008 to April 2008. However, cities in cluster 1 continued to be in a recessionary state until the moving quarter of May-July 2009. In the COVID-19 crisis, despite cluster 1 cities already being in a recession, it is theorized that this extension is unrelated to the pandemic's nature, given its unexpected nature. The widespread impact of this crisis on all cities was detected from January 2020 to March 2020, continuing until April-June 2020. Only cluster 2 cities experienced an additional quarter of recession. Subsequently, a prolonged expansion period for all cities unfolded in the following moving quarter from July-September 2020 to January-March 2021.

5.5. Regime Transition Probabilities and Logistic Coefficients Analysis

Table 2 presents the regime transition probabilities denoted as p_{ij}, illustrating the likelihood of the economy transitioning from one regime at time t, to another at $t+1$. The transition probabilities between different clusters are not feasible. For instance, suppose $z_t = 1$ in quarter t; this signifies that only cities or metropolitan areas (M.A.) included in cluster 1 would be in a recession. It has been predetermined that these cities and M.A. cannot emerge from the recession, and a distinct subset of cities and M.A. begins a recession at $t+1$. To achieve this, the constraint $P_{ij} = 0$ is enforced, effectively eliminating the possibility of such transitions.

[15] The cities not belonging to Cluster 2, and thus not following the recessionary trend in the moving quarter of May–July 2020, include Monteria, Pasto, Riohacha, Santa Marta, Cúcuta, M.A. Barranquilla, M.A. Armenia, Ibagué, Sincelejo, Villavicencio, and Neiva.

Table 2. Estimated transition probabilities.

	From National Expansion	From National Recession	From Recession in Cluster 1	From Recession in Cluster 2	From Recession in Cluster 3	From Recession in Cluster 4
To National Expansion	0.78	0.00	0.00	1.00	0.00	0.00
To National Contraction	0.00	0.44	0.11	0.00	0.34	0.23
To Recession in Cluster 1	0.00	0.44	0.89	**0.00**	**0.00**	**0.00**
To Recession in Cluster 2	0.00	0.09	**0.00**	0.00	**0.00**	**0.00**
To Recession in Cluster 3	0.00	0.03	**0.00**	**0.00**	0.66	**0.00**
To Recession in Cluster 4	0.22	0.00	**0.00**	**0.00**	**0.00**	0.77

Notes: Boldface indicates the restrictions on transition probabilities that $P_{ij} = 0$ where $i, j \in \{1, \ldots, l\}$ and $i \neq j$.
Source: Own elaboration

Based on the findings, it is observed that the probability of transitioning from a state of national recession to a state of expansion, or vice versa, is zero. Consequently, national expansions and contractions do not unfold sequentially; instead, they typically originate in specific cities and metropolitan areas before gradually extending to encompass all the examined cities and metropolitan areas. Cluster 1 displays the highest likelihood of staying in the same regime. Cluster 2, having only one expansion quarter, has an absolute probability of transitioning into the expansion regimen. Clusters 3 and 4 show an increased probability of shifting into national contractions, with Cluster 4 being especially susceptible to entering a recession following a national expansion (though, given only one expansion period, this is somewhat evident). Furthermore, our findings suggest that contractions will likely persist with a high probability in the subsequent period.

Table 3 presents the posterior means of the logistic coefficients β_k associated with each idiosyncratic cluster $k = (1,2,3)$, with highlighted entries indicating instances where only 68% of the calculations yielded consistent directionality compared to the inferred mean. The results of the β_k analysis for the three clusters identify patterns in the probability of membership. The interpretation of the logistic coefficients β_k is that for holding all other variables constant, a one-unit increase in the corresponding covariable associated with β_k is linked to an increase in the logarithm of the odds regarding covariable membership.

Table 3. Estimated logistic coefficients.

	Cluster 1 β_1	Cluster 2 β_2	Cluster 3 β_3	Cluster 4 β_4
Constant	-0.09	0.08	0.03	0.01
Number of companies per capita	**-0.10**	0.06	**0.41**	**-0.42**
Proportion of employees in manufacturing industries	0.02	0.07	**0.42**	-0.09
Proportion of employees in logistics and accommodation	**0.23**	**-0.16**	**0.43**	**0.69**
Proportion of employees in financial and insurance activities	0.31	**0.47**	0.09	-0.13

Source: Own elaboration.

In the examination conducted, it's highlighted that the number of companies per capita has a negative impact on the likelihood of belonging to Cluster 1, whereas a higher proportion of employees in logistics and accommodation sectors positively affects the probability of affiliation with this cluster. Conversely, within Cluster 2, a higher concentration of employees in logistics and accommodation sectors results in a negative impact, while a greater presence of employees in financial and insurance activities shows a positive association. Regarding Cluster 3, the number of companies per capita, alongside the proportion of employees in manufacturing industries and logistics and accommodation, positively influences the probability of belonging to this cluster. In Cluster 4, company density has a negative influence, while the proportion of employees in logistics and accommodation sectors exerts a stronger influence compared to Clusters 1 and 3.

6. Conclusions

By examining results from a Markov panel switching model applied to the growth rates of 23 cities and metropolitan areas in Colombia, this study presents the anticipated average growth rates for both recession and expansion cycles. We provide an empirical characterization of the growth rates observed in select cities, followed by the introduction of identified clusters and the associated predicted probabilities within the domains of expansion, recession, and cluster recessions.

The analysis of these elements reveals different patterns. Regarding growth rates during these phases, we observe notable heterogeneity in values; the highest rates during expansion peak at 2.6%, while the lowest rates during recession periods drop to -0.96%. Cities with high proportions of the population employed in logistics and accommodation activities display more resilient behavior during economic contractions. In contrast, cities with elevated employment rates in manufacturing industries exhibit less volatile growth rates throughout recession and contraction periods. Four distinct clusters are identified, with only one existing before the COVID-19 pandemic, while three different clusters emerge afterward. Although geographical continuity is observed in two of the identified clusters, it is hypothesized that this may be a coincidence, given that these clusters pertain to cities and not subnational units. Throughout the study period, the behavior of regimes surrounding the previously identified national crises, namely the Great Recession of 2008 and the COVID-19 pandemic, was analyzed. In both instances, a national recession is observed, followed by a subset of cities emerging first from this recession. Significantly, it is only after the COVID-19 pandemic that a period of national expansion is noted.

The obtained results serve as the groundwork for future investigations, aiming to explore the specific characteristics of cities that exhibit greater dynamism during economic growth, and those that demonstrate greater resilience during periods of contraction. We briefly describe some of these characteristics, but it is essential to acknowledge that other aspects may be involved. This comprehensive understanding could enable identification of characteristics or strategies employed by these cities, allowing governments to replicate them in applicable cities and sectors. Furthermore, certain cities can more effectively recover from national recessions than others. The limited literature on regional spreads emphasizes the necessity of investigating the spread of recessions with additional covariates; this highlights the importance of further research to comprehend the dynamics of economic fluctuations at a regional level.

References

Alfonso, V., Arango, L., Arias, F., & Pulido, J. (2011). Ciclos de negocios en Colombia: 1975-2011. doi:https://doi.org/10.32468/be.651

Arango, L. E., Parra-Escobar, F. F., & Pinzón-Giraldo, Á. J. (2016). El ciclo económico y el mercado de trabajo en Colombia: 1984-2014. *Ensayos Sobre Política Económica, 34*(81), 206-228. doi:https://doi.org/10.32468/be.911

Ayala-García, J. (2014). Crecimiento económico y empleo en Ibagué. *Documentos de Trabajo Sobre Economía Regional y Urbana*(210). doi:https://doi.org/10.32468/dtseru.210

Barón, J. D. (2002). Las regiones económicas de Colombia: Un análisis de clusters. *Documentos de Trabajo Sobre Economía Regional*(23). doi:https://doi.org/10.32468/dtseru.23

Barrantes, D. K. (2021). Análisis sobre el impacto de los beneficios tributarios, la dinámica empresarial y la generación de empleo para la ciudad de Armenia. *Tesis/Trabajo de grado*. Obtido de http://hdl.handle.net/10654/39840

Barro, R. J., Sala-I-Martin, X., Blanchard, O. J., & Hall, R. E. (1991). Convergence Across States and Regions. *Brookings Papers on Economic Activity*(1), 107-182. doi:https://doi.org/10.2307/2534639

Bartik, T. (2015). How effects of local labor demand shocks vary with the initial local unemployment rate. *Growth and Change, 46*(4), 529-557. doi:https://doi.org/10.1111/grow.12101

Beraja, M., Hurst, E., & Ospina, J. (2019). THE AGGREGATE IMPLICATIONS OF REGIONAL BUSINESS CYCLES. *Econometrica, 87*(06), 1789-1833. doi:https://doi.org/10.3982/ECTA14243

Bohórquez, L. C., Godoy, J. M., & Téllez, D. F. (2009). Los ciclos del empleo regional en Colombia: dependencia interregional 2001-2006. *Ensayos sobre Política Económica, 27*(58), 16-55. doi:https://doi.org/10.32468/Espe.5801

Cárdenas, C. A., Hernández, M. A., & Torres, J. E. (2015). A Statistical Analysis of Heterogeneity on Labour Markets and Unemployment Rates in Colombia. *Desarrollo y Sociedad*(75), 153-196. doi:https://doi.org/10.13043/dys.75.4

Carlino, G., & Sill, K. (1997). Regional economies: separating trends from cycles. *Business Review, Federal Reserve Bank of Philadelphia*(May/June), 19-31.

DANE. (2023). *DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA - DANE.* Obtido de Información histórica del Mercado laboral. Agosto 2023.: https://www.dane.gov.co/index.php/estadisticas-por-tema/mercado-laboral/empleo-y-desempleo/mercado-laboral-historicos

Duranton, G. (2015). Roads and trade in Colombia. *Economics of Transportation, 4*(1-2), 16-36. doi:https://doi.org/10.1016/j.ecotra.2014.11.003

Fedesarrollo. (2020). Fedesarrollo establece que el primer mes de recesión económica sucedió en marzo de 2020. *COMUNICADO DE PRENSA*. Obtido de https://img.lalr.co/cms/2020/11/24173644/2020-11-24-Comunicado-de-Prensa-Ciclos-Econ%C3%B3micos-y-Recesi%C3%B3n-2020.pdf

Fingleton, B., Garretsen, H., & Martin, R. (2012). Recessionary shocks and regional employment: evidence on the resilience of UK regions. *Journal of regional science*, 109-133. doi:https://doi.org/10.1111/j.1467-9787.2011.00755.x

Foley, P. (2021). Seasonal adjustment of Irish official statistics. *Statistical Journal of the IAOS*(37), 57-66. doi:10.3233/SJI-200778

Gadea-Rivas, M. D., Gómez-Loscos, A., & Leiva-Leon, D. (2018). Increasing linkages among European regions. The role of sectoral composition. *Economic Modelling, 80*, 222-243. doi:https://doi.org/10.1016/j.econmod.2018.11.009

Galvis-Aponte, L. A., Rodríguez-Puello, G. O., & Ovallos-Bencardino, S. (2019). Desempleo y calidad de vida laboral en las áreas metropolitanas de Barranquilla, Cartagena y Santa Marta. *Documentos de Trabajo Sobre Economía Regional y Urbana*(279). doi:https://doi.org/10.32468/dtseru.279

Gomez Sanchez, A. M., Sarmiento-Castillo, J. I., & Fajardo-Hoyos, C. L. (2022). Regional business cycles and manufacturing productivity: empirical evidence in Colombia. *EconomiA, 23*(1), 62-87. doi:https://doi.org/10.1108/ECON-05-2022-0020

Gómez, M. (2013). Análisis del ciclo económico en una economía con rigideces nominales y un amplio sector informal. *Ensayos Sobre Política Económica, 72*(31), 51-66. doi:https://doi.org/10.1016/S0120-4483(13)70004-6

González, C. G., & Mora, J. J. (2011). Políticas activas de empleo para cali-colombia. *Estudios Gerenciales, 27*(118), 13-41. doi:https://doi.org/10.1016/S0123-5923(11)70144-9

Guarín, A. G. (2017). Recomendaciones a la productividad en Manizales a través de mejores condiciones de empleo. *Questionar: Investigación Específica, 5*(1), 9-19. doi:https://doi.org/10.29097/23461098.97

Guzman, L. A., Oviedo, D., & Bocarejo, J. P. (2017). City profile: The Bogotá Metropolitan Area that never was. *Cities, 60*, 202-215. doi:https://doi.org/10.1016/j.cities.2016.09.004

Haddad, E. A., Domingues, E. P., & Perobelli, F. S. (2002). Regional effects of economic integration: the case of Brazil. *Journal of Policy Modeling,, 5*(24), 453-482. doi:https://doi.org/10.1016/S0161-8938(02)00125-4

Hamilton, J. D., & Owyang, M. T. (2012). The Propagation of Regional Recessions. *The Review of Economics and Statistics, 94*(4), 935-947. doi:https://doi.org/10.1162/REST_a_00197

Hernández, R. R. (2023). El ciclo económico en Colombia: una aplicación del método de regresión de Hamilton al Índice de Seguimiento a la Economía. *Tesis de Maestría en Ciencias Económicas*. Obtido de https://repositorio.unal.edu.co/handle/unal/84070

Hernández-Murillo, R., Owyang, M. T., & Rubio, M. (2017). Clustered housing cycles. *Regional Science and Urban Economics, 66*, 185-197. doi:https://doi.org/10.1016/j.regsciurbeco.2017.06.003

Hudson, R. A.; Library Of Congress. (2010). *Colombia: A Country Study.* Washington, D.C: Federal Research Division. Obtido de https://www.loc.gov/item/2010009203/

Kondo, K. (2022). Spatial dependence in regional business cycles: evidence from Mexican states. *Journal of Spatial Econometrics, 3*(1). doi:https://doi.org/10.1007/s43071-021-00018-z

Ortiz, C. H., & Uribe, J. I. (2007). Hacia un modelo de desarrollo incluyente para el Valle del Cauca Estudios Gerenciales. *Estudios Gerenciales, 23*(102), 13-62. doi:https://doi.org/10.1016/S0123-5923(07)70001-3

Otero-Cortes, A., Alvarez, O., & Acosta, K. (2023). Mercado Laboral y Pobreza en Barranquilla. *Documento sobre economía regional y urbana*(323). doi:https://doi.org/10.32468/dtseru.323

Owyang, M. T., Piger, J. M., Wall, H. J., & Wheeler, C. H. (2008). The economic performance of cities: A Markov-switching approach. *Journal of Urban Economics, 64*(3), 538-550. doi:https://doi.org/10.1016/j.jue.2008.05.006

Posada, C. (2020). Crecimiento y ciclos de la economía colombiana (2005–2019). *Documentos de trabajo*. Obtido de http://hdl.handle.net/10784/17086

Quintero Otero, J. D., & González, A. (2012). Política monetaria y ciclos económicos regionales en Colombia. *CEDE*(42). doi:http://hdl.handle.net/1992/8359

Rissman, E. R. (1999). Regional employment growth and the business cycle. *ECONOMIC PERSPECTIVES-FEDERAL RESERVE BANK OF CHICAGO, 23*(4), 21-41.

Rivera, Y., & Echeverri, A. (2020). Turismo y Paz: una estrategia de desarrollo local. *NOVUM, revista de Ciencias Sociales Aplicadas, 1*(10), 42-60. Obtido de https://revistas.unal.edu.co/index.php/novum/article/view/79790

Robledo-Caicedo, J. (2019). La pobreza en Quibdó: Norte de carencias. *Documentos de Trabajo Sobre Economía Regional y Urbana*(277). doi:https://doi.org/10.32468/dtseru.277

Robledo-Caicedo, J. (2020). La pobreza en Florencia: Un análisis de sus factores, consecuencias y posibles soluciones. Obtido de https://doi.org/10.32468/dtseru.287

Sanchez, A. (2013). La reinvención de Medellín. *Lecturas de Economía*(78), 185-227. doi:https://doi.org/10.17533/udea.le.n78a15768

Wall, H. J. (2007). Regional business cycle phases in Japan,. *Federal Reserve Bank of St. Louis Review, 89*(1), 61-76. doi:https://doi.org/10.20955/wp.2006.053

Wang, X., Sun, Y., & Peng, B. (2023). Industrial linkage and clustered regional business cycles in China. *International Review of Economics and Finance, 85*, 59-72. doi:https://doi.org/10.1016/j.iref.2023.01.002

Zuccardi, I. E. (2002). Los ciclos económicos regionales en Colombia, 1986 - 2000. *Documentos de Trabajo Sobre Economía Regional y Urbana*(25). doi:https://doi.org/10.32468/dtseru.25

Appendices

Table 4. Average growth coefficients and cluster membership probabilities for each city.

City	μ_0	μ_1	$\mu_0 + \mu_1$	σ^2	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Bogotá D.C.	0.4285	-0.4570	-0.0285	0.2822	1.0000	0.9708	0.9460	0.0004
Medellín M.A.	1.0489	-1.0216	0.0273	0.3673	1.0000	0.7952	0.9984	0.0320
Cali M.A.	1.6967	-1.7230	-0.0263	0.5244	1.0000	1.0000	1.0000	1.0000
Barranquilla M.A.	0.3297	-0.4943	-0.1646	0.3267	0.0000	0.0144	0.9924	0.0100
Bucaramanga M.A.	1.2013	-1.1744	0.0270	0.5393	1.0000	0.9996	1.0000	0.0204
Manizales M.A.	0.8891	-0.8705	0.0186	0.5556	0.0000	0.9448	0.9748	0.0000
Pasto	1.1262	-1.1316	-0.0054	0.4908	1.0000	0.3108	0.9836	1.0000
Pereira M.A.	1.3900	-1.3666	0.0234	0.5636	1.0000	0.9388	1.0000	0.9988
Cúcuta M.A.	1.5169	-1.5311	-0.0143	0.6329	1.0000	0.0144	1.0000	0.9868
Ibagué	2.7672	-2.8196	-0.0525	1.0825	0.0000	0.0072	0.9972	0.0000
Montería	0.7203	-0.8659	-0.1456	0.4973	1.0000	0.4600	0.9928	0.9992
Cartagena	0.0910	-1.5557	-1.4646	0.6123	0.9992	0.6024	0.9964	1.0000
Villavicencio	1.8947	-1.9073	-0.0126	0.4516	1.0000	0.0004	1.0000	0.9988
Tunja	0.1227	-1.9899	-1.8673	0.6185	0.5416	0.9448	0.9328	0.0004
Florencia	2.2860	-2.3633	-0.0773	0.9468	1.0000	0.9972	1.0000	1.0000
Popayán	0.1612	-1.8247	-1.6635	0.7492	1.0000	0.9888	0.9940	0.9976
Valledupar	0.5686	-1.2380	-0.6694	0.5070	1.0000	0.9904	1.0000	1.0000
Quibdó	0.0741	-0.4754	-0.4012	1.6643	0.0536	0.7248	0.9752	0.3124
Neiva	0.1400	-2.6520	-2.5120	0.6887	1.0000	0.0004	1.0000	0.9788
Riohacha	0.1264	-1.9671	-1.8407	0.7115	1.0000	0.0920	0.9980	0.9896
Santa Marta	0.1037	-1.4112	-1.3075	0.4510	1.0000	0.0684	0.9228	1.0000
Armenia	2.5149	-2.5744	-0.0595	0.8719	1.0000	0.0100	1.0000	1.0000
Sincelejo	1.3882	-1.3853	0.0029	0.5179	1.0000	0.0064	0.9984	0.9744

Source: Own elaboration.