

RESEARCH NOTE

A Monthly Regional Indicator of Economic Activity: An Application for Latin America

Pavel Vidal*, Lya Paola Sierra*, Johana Sanabria† and Jaime Andrés Collazos†

* Pontificia Universidad Javeriana Cali, CO

† Banco de la República, CO

Corresponding author: Pavel Vidal (pavel@javerianacali.edu.co)

The study developed by Stock and Watson (1991) has served as a catalyst for the construction of monthly indicators of economic activity. However, its application has been focused on national statistics. In this article we develop a new empirical strategy of three steps to build regional economic indicators on a monthly basis that can be replicated in provinces or municipalities in Latin America. We apply this empirical strategy to develop the Monthly Indicator of Economic Activity (in Spanish, IMAE) for one of the main regions of Colombia: Valle del Cauca.

El estudio desarrollado por Stock y Watson (1991) ha servido de catalizador para la construcción de indicadores mensuales de la actividad económica. Sin embargo, su aplicación se ha centrado en las estadísticas nacionales. En este artículo desarrollamos una estrategia empírica de tres pasos para la construcción de indicadores de actividad económica mensuales regionales que pueden ser replicados en provincias o municipios de América Latina. Dicha estrategia empírica la empleamos en la construcción del Indicador Mensual de Actividad Económica (IMAE) en una de las principales regiones de Colombia: Valle del Cauca.

At the international level, the study developed by Stock and Watson (1991) has served as a catalyst for the construction of monthly indicators of economic activity. Such indicators enable the monitoring of general economic conditions before gross domestic product (GDP) data are published because statistical offices typically record GDP on a quarterly basis and publish the first results after several weeks.

However, the application of this methodology has been focused on national statistics. There are very few examples of monthly indicators of economic activity for a region within a country. We did not find any regional indicators that were constructed using Stock and Watson's (1991) methodology in Latin America. In international literature, we identified only two applications, both in the United States: the work by Megna and Xu (2003, 701–713) for the state of New York and by Crone (2005) for all fifty states.¹

This article applies a version of this methodology that is adjusted on the basis of the characteristics of the data that are usually present at regional level. The contributions of this study are twofold; first, we develop a three-step empirical strategy to build regional economic indicators on a monthly basis that can be replicated in provinces or municipalities in Latin America; that is, we add two more steps to the conventional methodology of Stock and Watson (1991). Second, we apply this empirical strategy to develop the Monthly Indicator of Economic Activity (in Spanish, IMAE) for one of the main regions of Colombia: Valle del Cauca. However, the interest of the study goes beyond the specific case of the Valle del Cauca, which is presented as an example of the viability and applicability of the empirical strategy that we propose.

Valle del Cauca occupies a strategic geographic position for Colombia, because it is the gateway to trade with Asia and the entire Pacific Coast of Canada, the United States, and Latin America. However, as in most of

¹ The literature on regional indicators of economic activity in Latin America is scarce and recent. The few existing regional exercises have mostly used the methodology of the National Bureau of Economic Research. For the case of Argentina, see, for instance, Jorrat (2005, 43–100) and Rivero (2007, 31–73). For the case of Mexico, the Bank of Mexico summarized the information of economic cycles of thirty-two states of Mexico using diffusion indexes developed by Guerrero and Martínez-Ovando (2014, 34–36).

the provinces in Latin America, to plan finances and make spending and investment decisions, the regional government and the companies operating in the department monitor on monthly bases the evolution of dissimilar economic variables and receive data on the region's annual GDP with a several-month lag. In this article, we show how we constructed an indicator that unifies all this information in a single variable to estimate the general state of the department's economy on a monthly basis. Because this problem is not particular to the government and companies of Valle del Cauca, we present a methodology of interest, which can be easily replicated in other regions in Latin America.

The IMAE is already in use in the Regional Economic Bulletin of the central bank of Colombia (Banco de la República 2015). Each month, the IMAE indicates whether the economic activity of Valle del Cauca is accelerating or decelerating, and whether the estimated economic growth is above or below the historical average. This real-time information about the economic cycle falls within what is internationally known as “nowcasting” because it provides an estimate of the state of the economy at that time (in each month), according to the information available to date.

Indicators of economic activity are expected to contribute to better fiscal management of the regional governments in Latin America, because the government will have greater ability to project its income and plan its spending budget. Meanwhile, companies can make a better assessment of the income effect on demand for their goods and services. In general, all economic actors will have information in “real time” on the general state of the economy, which is key to early decision making and efficient resource allocation.

The first step of this empirical strategy begins with the estimation of a dynamic factor model (DFM) with the Kalman filter. The DFM assumes that there is an unobserved or latent variable common to a group of various observed variables. By using series related to economic activity, the latent variable approximates the general state of the economy. The DFM seeks to identify common repetitive sequences in the series, that is, the co-movements. This pattern in the common dynamic provides a signal about the evolution of the economic cycle. For regional indicators, an initial study has to be conducted to search for and select key variables for the economic cycle according to the economic structure of the province or municipality and the characteristics of the information available.

The second step includes the temporal disaggregation of the department's annual GDP using the Litterman method (1983, 169–173). A monthly GDP series is obtained to include this information in the coincident activity index. The methodology applied in this study is therefore unusual because it combines information on an annual basis and a monthly basis in a single indicator. The use of the Litterman method allows us to take into consideration the fact that regions do not have a quarterly GDP, only an annual one.

Finally, the third step incorporates the application of the univariate structural time-series model initially developed by Harvey (1989) to smooth the estimated common factor and get a better signal of the business cycle. The inclusion of Harvey (1989) methodology allows for the consideration that regional available series have high volatility, so it is necessary to smooth the indicator's trajectory for a more reliable estimate of the economic cycle.

The most common methods of DFM estimation are the principal components method and the Kalman filter. The former determines the least number of principal components that explain the largest proportion of variability in the original data set, with the first principal component being a linear combination of the original variables in the direction of maximum variability. Some studies in this regard include those by Gupta and Kabundi (2011, 1076–1088); Forni, Hallin, Lippi, and Reichlin (2009, 62–85); Schumacher (2007, 271–302); and Boivin and Ng (2006, 169–194). One of the characteristics of principal components estimation is that it requires a large number of variables, which becomes a disadvantage for regional application.

The origin of the filter can be found in a paper by Kalman (1960, 35–45). The Kalman filter is the main algorithm for estimating dynamic systems represented in state-space form. Among the applications performed with this method to construct indicators for economic activity are those by Angelini, Banbura, and Rünstler (2008, 1–22) and Camacho and Doménech (2012, 475–497), who applied a DFM, mixing monthly and quarterly data, to predict economic activity in the Eurozone and Spain, respectively. Additionally, Aruoba, Diebold, and Scotti (2009, 417–427) performed mixtures of frequencies with few series for the United States: with only five variables, they produced an index of economic activity that can be observed weekly. Other models focused on the Eurozone are those by Camacho, Pérez-Quirós, and Poncela (2013) and Camacho and Pérez-Quirós (2010, 663–694). There are also other DFM applications, such as the case

study by Poncela, Senra, and Sierra (2014, 3724–3735) that evaluates common movement in the prices of different raw materials.²

In Colombian national literature, the development of economic indicators at the macroeconomic level using DFMs has also had an extended presence. The following studies are noted: Melo and Nieto (2001, 1–32); Melo and colleagues (2001, 1–66); Melo, Nieto, and Ramos (2003, 1–57); Castro (2003, 1–26); Kamil, Pulido, and Torres (2010, 1–40); and Marcillo (2013, 1–41).

The sections that follow describe the three steps of the empirical strategy used to build a regional economic indicator, explain the selection and transformation process of the variables, report the three steps of the methodology, and present conclusions.

Methodology

Step 1: The Dynamic Factor Model and the Kalman Filter

The methodology used in this study uses Stock and Watson (1991) as one of its principal references. In general, the dynamic factor model assumes that a vector Y_t , with N observed variables of economic activity, can be represented as the sum of two unobservable components that are mutually independent: a component common to all variables (F_t) that represents the state of the economy and an idiosyncratic component (μ_t) that represents the dynamics of each series. Accordingly, the equation for the observed series Y_t is as follows:

$$Y_t = P F_t + \mu_t, \quad (1)$$

where Y_t is a vector ($N \times 1$) of the monthly variables of economic activity; F_t is the common factor or co-movement of the series; P is a factor loading matrix ($N \times 1$); and μ_t is a vector ($N \times 1$) of specific or idiosyncratic components. The dynamics of the factors are given by the following:

$$\Phi(B)F_t = a_t, \quad (2)$$

where a_t is normal multivariate white noise with a variance and covariance matrix Σ_a . $\Phi(B) = I - \Phi_1 B - \dots - \Phi_p B^p$ and B are the lag operator. We also consider that the idiosyncratic components may have a dynamic structure in the following form:

$$D(B)\mu_t = e_t, \quad (3)$$

where $D(B) = \text{diag}(D_i(B))$ is a diagonal matrix that includes the specific dynamics of each idiosyncratic disturbance; and $D_i(B) = 1 - d_{i1}B - \dots - d_{ip_i}B^{p_i}$, $i = 1, 2, \dots, m$, corresponds to the autoregressive structure of the component in each series represented with the lag operator. In addition, e_t is zero mean white noise with a diagonal covariance matrix.

The estimate of the DFM is performed with maximum likelihood by applying the Kalman filter. To that end, the model is written in state space:

$$Y_t = H z_t + \epsilon_t \quad (4)$$

$$z_t = G z_{t-1} + v_t \quad (5)$$

The measurement equation, equation 4, describes the relationship between the observed variables (data) and a vector of state variables. The state equation, equation 5, describes the dynamics of the state variables (unobserved variables) over time, where H is a matrix ($N \times m$) and ϵ_t is ($N \times 1$), with $\epsilon_t \sim \text{iid } N(0, R)$. G is a transition matrix $m \times m$, v_t is $m \times 1$, and $v_t \sim \text{iid } N(0, Q)$. It is assumed that the errors in the measurement equation ϵ_t are independent from the errors in the transition equation v_t , such that:

$$E(\epsilon_t v_s) = 0. \quad (6)$$

² The Kalman filter also has applications in estimating models with coefficients that change over time. For example, Kim and Nelson (1989, 433–440) studied variations in the reaction function of the Federal Reserve; Haldane and Hall (1991, 436–443) investigated the changing relationship between the British pound and the Euro and the dollar; and Vidal and Fundora (2008, 97–116) estimated a model that dynamically relates growth with balance of payment restrictions in Cuba.

The state vector contains the most important information of the system at each point in time.³ The basic Kalman filter is based on an algorithm of predicting and updating vector z_t , which is repeated for each observation from the beginning to the end of the sample using the initial values of the system parameters (matrices: H, G, R, and Q). The algorithm minimizes the average squared prediction errors recursively: in each observation, z_t is updated with the new information contained in the prediction error (Kalman gain).⁴

Thus, in estimating the vector of the state variables z_t and in matrix H, the most relevant information for this study is found: the common factor or co-movement (F_t) of the variables used and the matrix of loads or weights (P) that they each contribute to calculating the factor.

Step 2: The Temporal Disaggregation of GDP

Step 1 is used to estimate a common factor of the monthly variables available. However, to efficiently use the available information, we worked to incorporate the annual GDP data available for the department since 2000 in the indicator. The goal is for the indicator cycles to be more consistent with the annual GDP growth rates for the province or municipality in the period in which this information is available.

In an extension of the work of Stock and Watson (1991), Angelini, Banbura, and Rünstler (2008, 1–22) and Camacho and Doménech (2012, 475–497) developed a methodology of dynamic factor models with the possibility of combining data on quarterly GDP with higher frequency series. This methodology is not applicable to Valle del Cauca because no quarterly GDP information is available. In general, very few regions within countries have quarterly GDP statistics in Latin America.

In this article, we apply a methodology that is consistent with the characteristics of the information that is typically available at the regional level. A methodology is proposed in which monthly and annual data are combined in the indicator after temporarily disaggregating GDP with the common factor previously estimated with the Kalman filter in the DFM. Specifically, the common factor is used to measure GDP on a monthly basis, following the Litterman method (1983, 169–173).

Temporal disaggregation consists of generating a high-frequency series starting from low-frequency data. Some of these methods use one or more additional related variables with a high frequency (for a detailed review of the different methods of temporal disaggregation, see Quilis 2001; Hurtado and Melo 2010, 1–36). In this article, a factor (F_t) that was previously extracted as co-movement from the monthly series is used as the related variable.

The temporal disaggregation technique for economic series proposed by Litterman (1983) generalizes the approach described by Fernández (1981, 471–478) and assumes the existence of a linear relationship between the estimating series (h) and the high-frequency p related series (w):

$$h = w\beta + u, \quad (7)$$

where $h = [h_1, \dots, h_N]$ is a vector $N \times 1$ with unobserved high-frequency values of the series after temporal disaggregation; $w = [w^1, \dots, w^p]$ is vector $N \times p$ with p high-frequency related variables (for this study, $p = 1$); β is a vector $p \times 1$ of coefficients; and u has the dimension $N \times 1$ with zero mean and covariance matrix V .

Litterman (1983, 169–173), unlike Chow and Lin (1971, 372–375), argues that u_t follows a process I(1) whose innovation is in turn a stationary AR(1). Formally:

$$u_t = u_{t-1} + \varepsilon_t \quad (8)$$

$$\varepsilon_t = \alpha \varepsilon_{t-1} + a_t \quad |\alpha| < 1 \quad (9)$$

$$a_t \sim iid N(0, \sigma_a^2),$$

where a_t is a white noise process with variance σ^2 .

Step 3: The Univariate Structural Time-Series Model

The common factor resulting from the estimate with the DFM sometimes contains too much noise, which makes it difficult to have a clear reading of the state of the economy. To have an indicator that can be interpreted and used by economic actors, it is necessary to extract a stronger signal, that is, to estimate the most permanent trajectory of the variable.

³ The presence of common and idiosyncratic factor dynamics enables their projection over time. In the estimate for the construction of the IMAE, both components are specified as AR(1) processes.

⁴ In Hamilton (1994) and Koopman, Shephard, and Doornik (1999, 113–166), there is a complete description of the Kalman filter.

To that end, there are different methods of signal extraction. There are empiricist methods, which are characterized by implementing decomposition based on linear filters whose structure and parameters are not dependent on the nature of the data but instead have preset values. These include moving averages, exponential smoothing, X-11 ARIMA (autoregressive integrated moving average) (and its most recent versions), and the Hodrick-Prescott filter.

In this article, we use the univariate structural time-series model initially developed by Harvey (1989). This method, unlike the empiricist methods, take into account the particular characteristics of each time series when estimating their different unobservable components (trend, cycle, seasonality, and irregular component). It allowed each of the unobservable components to have a stochastic nature and is distinguished by independently estimating the trend of the cyclical component, and they can estimate up to three cycles of a different period.

Its broadest specification, considering all components of a time-series y_t is as follows:

$$y_t = \mu_t + \psi_t + \gamma_t + \varepsilon_t \tag{10}$$

$$\mu_t = \mu_{t-1} + \beta_t + \eta_t \tag{11}$$

$$\beta_t = \beta_{t-1} + \xi_t \tag{12}$$

$$\sum_{j=0}^{s-1} \gamma = \omega_t \tag{13}$$

where the trend is broken down into changes of level μ_t and slope β_t ; γ_t is the seasonal component; and ε_t is the irregular component, which can consist of white noise or follow an autoregressive process. In turn, η_t , ξ_t , and ω_t are white noise with variance σ^2_η , σ^2_ξ , and σ^2_ω ; they are called hyperparameters and allow components to evolve stochastically if they are not equal to 0. ψ_t represents the cyclical component, which is modeled with periodic sine and cosine functions:

$$\psi_t = \cos\lambda_c \psi_{t-1} + \text{sen}\lambda_c \psi^*_{t-1} + \kappa_t \tag{14}$$

$$\psi^*_t = -\text{sen}\lambda_c \psi_{t-1} + \cos\lambda_c \psi^*_{t-1} + \kappa^*_t \tag{15}$$

where κ_t and κ^*_t are white noises that are not correlated with each other or with any other disturbances in the model and with common variance σ^2_κ ; and the parameter λ_c is the frequency measured in radians; that is, it represents the number of times the cycle is repeated in a time period with length 2π . To estimate the model, we used the program Structural Time Series Analyser, Modeller and Predictor (STAMP) by Koopman, Harvey, Doornik, and Shephard (2009).

Data

The initial stage in building a regional indicator is variable selection. For Valle del Cauca, initially, twenty-seven monthly variables related to the economic activity of the department were analyzed (in Table A1 of the Appendix, the entire group of variables considered is shown).

The selection of the series to calculate the indicator is based on the following criteria: (1) series with a monthly frequency that were published during the period of study; (2) variables with the largest annual correlation with departmental GDP; (3) key variables that represent different sectors and or components of demand in the region's economy; and (4) variables with less lag in monthly publication.

The seven variables chosen that met these criteria and were eventually used to estimate the IMAE were the following:

1. Ground sugarcane (CAN)
2. Cement shipments (CEM)
3. Nonresidential energy consumption (ENER)
4. New vehicle sales (VEH)
5. Exports at constant prices (X)
6. The Regional Industrial Production Index (Índice de Producción Industrial Regional, or IPIR)⁵
7. Imports at constant prices (M)

⁵ We used data obtained by Escobar (2015) at the Banco de la República, which consisted of calculating a Regional Industrial Production Index (in Spanish, IPIR) by applying the share of value added in the Annual Manufacturing Survey (in Spanish, EAM) from DANE to the indices of the Monthly Manufacturing Survey (in Spanish, EMM).

The selected variables contain direct and/or indirect information about different key activities of the regional economy on either the supply side or the demand side. Importantly, each variable contains information beyond the sector where it is measured; this is the case for nonresidential energy demand and the data on exports and imports. All the variables provide indirect information on the general state of business activity in the region. Similarly, new vehicle sales and imports are linked to consumption and the economic situation of households. In this manner, estimating the state of the department's economy is facilitated with few variables. The fact that the indicator is constructed using a few variables makes it feasible to repeat the methodology in other regions.

The sectors that are directly considered in the indicator are agriculture, construction, energy, industry, and trade. Compared to the sectors that are typically considered indicators of economic activity internationally, the inclusion of ground sugarcane in the IMAE is noteworthy. According to the Ministry of Agriculture and Rural Development, sugarcane accounts for 94 percent of agricultural production in Valle del Cauca and 49 percent of the sown area, making it the main agricultural crop of the department. Figures from the Asocaña business organization show that 77.3 percent of the production of sugar mills in Colombia is concentrated in Valle del Cauca. Sugarcane is the raw material for many important networks in other industries in the department, such as beverages, food, paper, cardboard, and pharmaceuticals; additionally, it is used in ethanol production and electricity cogeneration (Escobar, Moreno, and Collazos 2013, 3–10).

Furthermore, it is not usual to include external trade data in indicators of economic activity. According to figures from the Annual Report on Transportation in 2013, in Colombia, 96.7 percent of foreign trade is moved by maritime shipping. Valle del Cauca has a comparative advantage in having the only seaport on the Colombian Pacific Coast. The port of Buenaventura moved 46.2 percent of the total burden of foreign trade in 2013. Thus, higher shipments of goods to and from abroad indicate increased economic activity in the department by promoting transportation, logistics, employment, and industry in the region.

The variables that comprise the IMAE are the type used internationally in coincident economic activity indicators.⁶ They contain contemporary information on the supply side about economic sectors or effective demand in the economy. Therefore, it is assumed that the turning points of the IMAE coincide with the dates of turning points in the economic cycle, unlike leading indicators in which turning points anticipate the economic cycle. To construct the latter type of indicators, data would be needed with regard to expectations, interest rates, monetary aggregates, and other variables that contain advance information on economic activity.

Stock and Watson (1991) used the following macroeconomic variables in their estimations: industrial production, household income, sales, and employment. The variables considered for the IMAE directly or indirectly contain information about these categories, with the exception of employment. The series of total employment and unemployment rate available for the department had many outliers and were not consistent with the other variables used in the IMAE. Other regions in Latin America might have better data on employment so are worth taking into account.

The following analysis and treatment of the series were performed. First, the augmented Dickey-Fuller (ADF), Phillips-Peron (PP), and Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) unit root tests were performed for each of the series to evaluate the presence of unit roots. Except for sugarcane, which is stationary, all selected series were integrated of order one. In addition, seasonality and outliers in the series were eliminated using the TRAMO-SEATS program of Gómez and Maravall (1996). A logarithmic first difference was applied to $I(1)$ variables, which was expressed in monthly variation rates. Last, all variables were standardized (see Table A2 in the Appendix for details on the treatment performed on each series).

Figure 1 shows the transformed series. All series have high variance; thus, it is expected that the co-movement of the series also show high variability.

Results

Step 1: The Estimation of the Common Factor

Figure 2 shows the results for the common factor estimated by the Kalman filter. In **Table 1**, the weights that each variable contributes are shown (normalized to sum to 1). All series have positive correlations with the factor; that is, all are procyclical.

Table 1 shows that 46 percent of the common factor is composed of industry and the new vehicle sales variable (which approximates conditions regarding consumption and the economic situation of households). Next in importance are variables related to the external sector: exports with 16 percent and imports with 15

⁶ The variables used are classified by Marcillo (2013, 1–41) as coincident.

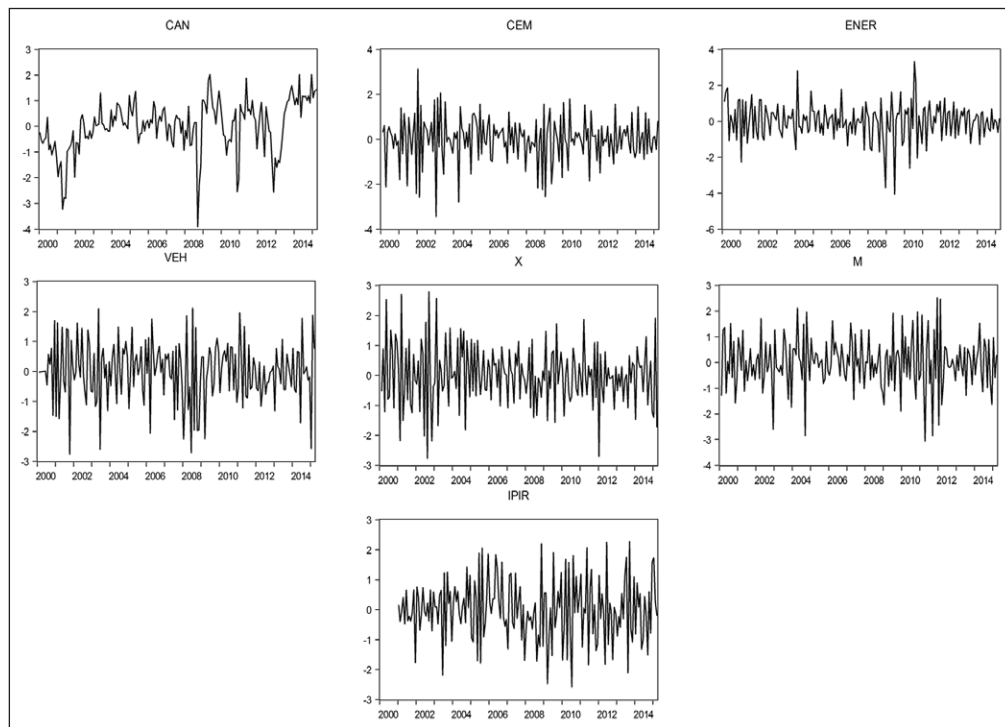


Figure 1: Monthly series used to estimate the IMAE, January 2000–March 2015 (Series in their stationary transformation, in logarithm, corrected for outliers and standardized).

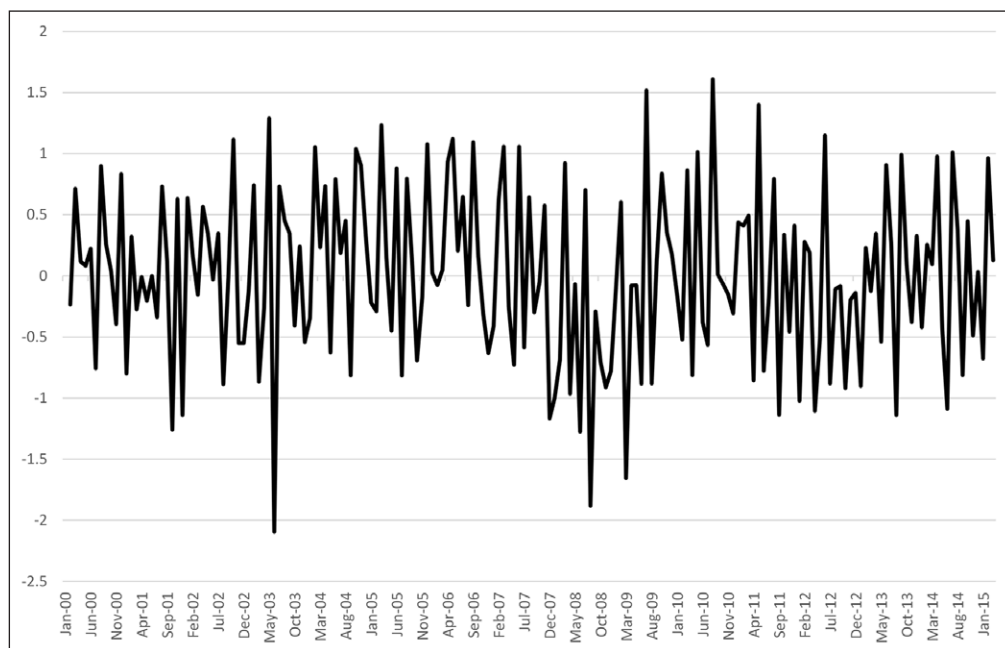


Figure 2: Common factor (F_t) estimated with the Kalman filter. Rate of monthly variation (%), January 2000–March 2015.

percent. Energy consumption has a weight of 12 percent, and the ground sugarcane and cement shipments each have a weight of 6 percent.

Step 2: Common Factor Corrected with GDP Information

As explained already in the methodology section, the indicator incorporates information on the department's annual GDP data published by the National Administrative Department of Statistics (in Spanish, DANE). Thus, the aim is for indicator cycles to be more consistent with the annual growth rates of GDP in the period in which this information is available.

Table 1: Weights (W) that each variable contributes to the common factor.

1. Ground sugarcane	0.06
2. Cement shipments	0.06
3. Energy consumption	0.12
4. New vehicle sales	0.24
5. Exports	0.16
6. Regional IPI	0.22
7. Imports	0.15

Table 2 shows the estimated monthly GDP data for the period 2000–2013 using the common factor (F_t) as related variable in the Litterman method. It is shown at constant prices (given that the annual data were in constant prices) and without seasonality (given that the common factor was constructed with seasonally adjusted variables).

Figure 3 shows the common factor corrected with monthly GDP information. The new series is composed from January 2000 to December 2013, using the rates of change of monthly GDP, and from January 2014 to March 2015, using the common factor previously estimated by the Kalman filter. The fact that GDP is disaggregated with the common factor ensures consistency throughout the series.

Importantly, the routine for the systematic use of the IMAE involves updating the factor by taking into account new annual GDP data whenever DANE publishes new information. However, there will never be a complete updated series for the IMAE corrected with GDP data, given the lag in publication. Therefore, monthly GDP does not replace the IMAE. The real-time assessment of the economy in the last period of the series (nowcasting) would always have been made with the “pure” data from the common factor. In this sense, the main benefit of measuring GDP monthly as part of the methodology to construct the IMAE is that doing so allows for a better assessment of the past, which improves the estimation of economic cycles.

Step 3: The Economic Cycle

In the trajectory of the factor in **Figure 2**, a cyclical component can be observed but with turning points that are hidden by the “noise” of the series. With this series, it is difficult to obtain a clear reading on the state of the economy in the region. Therefore, it is necessary to smooth the series by applying the univariate structural time-series model.

In the specification of the univariate structural model, neither seasonality components nor slope were included because both were eliminated in the factor estimation. The results of the estimation by the univariate structural model show that the series has the following components:⁷

- A fixed level equal to 0.0031, which indicates the average monthly growth rate of the series (annual would be 3.8 percent).
- An irregular AR(1) component. This type of component coincides with the specification of both the DFM and the Litterman method.
- A cycle period of 4 years and 10 months. It is noteworthy that this length of time is an average period of a stochastic cycle, based on growth rates (growth cycle). For the average period, it is the cycle known in the literature as the “business cycle.” The estimated period of the cycle for Valle del Cauca approaches the estimated four years for the national GDP, according to the study by Arango, Arias, Flores, and Jalil (2008, 9–37).

Figure 4 shows the cyclical component and **Figure 5** the estimated irregular component. The sum of these plus the fixed level (0.0031) would result in the common factor previously shown in **Figure 3**. The irregular component includes noise to be removed from the series, whereas the cycle contains the signal sought for

⁷ In the specification of the univariate structural model, neither seasonality components nor slope were included because both were eliminated in the factor estimation. The specification of the univariate structural model to estimate the unobservable components of the factor is facilitated because the series was constructed using the DFM and monthly GDP, which means that it has information about the data-generating process.

Table 2: Monthly GDP for Valle del Cauca (at constant prices and without seasonality).

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
<i>January</i>	2475.5	2518.3	2565.1	2648.1	2769.1	2852.8	2984.1	3297.5	3458.7	3483.1	3525.6	3628.5	3789.6	3926.7
<i>February</i>	2479.2	2523.0	2574.6	2663.5	2793.0	2846.8	3005.8	3325.3	3458.0	3500.1	3518.8	3649.1	3807.9	3948.9
<i>March</i>	2491.7	2521.8	2580.5	2661.2	2805.8	2861.2	3030.1	3358.0	3480.4	3483.4	3533.5	3671.5	3825.1	3965.4
<i>April</i>	2497.4	2524.0	2594.3	2666.8	2823.0	2863.6	3066.9	3371.4	3475.5	3488.9	3523.8	3672.5	3821.1	3989.7
<i>May</i>	2502.0	2524.8	2605.3	2691.0	2822.1	2860.4	3106.5	3377.1	3483.7	3493.3	3543.0	3709.7	3827.2	3999.1
<i>June</i>	2507.4	2528.4	2612.0	2677.0	2836.5	2876.5	3134.2	3406.7	3474.6	3485.0	3542.2	3712.0	3861.4	4033.0
<i>July</i>	2502.1	2529.2	2622.5	2696.6	2841.8	2873.3	3167.6	3412.1	3494.2	3510.8	3539.8	3723.8	3862.4	4055.8
<i>August</i>	2513.3	2542.1	2618.9	2713.7	2848.5	2893.4	3188.7	3434.0	3476.4	3500.5	3572.3	3751.0	3876.7	4054.4
<i>September</i>	2517.5	2549.2	2624.9	2730.3	2838.2	2907.7	3227.3	3441.6	3481.9	3504.3	3581.6	3746.7	3892.0	4089.3
<i>October</i>	2519.1	2541.9	2643.5	2738.5	2849.3	2913.8	3252.5	3452.0	3481.3	3518.0	3591.0	3766.1	3893.8	4108.6
<i>November</i>	2516.0	2555.7	2643.2	2754.4	2857.9	2928.9	3270.4	3470.6	3477.6	3524.2	3600.2	3772.4	3908.0	4119.5
<i>December</i>	2525.7	2550.6	2643.0	2760.9	2858.6	2962.5	3283.0	3463.7	3475.6	3527.4	3608.1	3792.8	3923.6	4142.6
<i>Year</i>	30047	30409	31328	32402	33944	34641	37717	40810	41718	42019	42680	44596	46289	48433

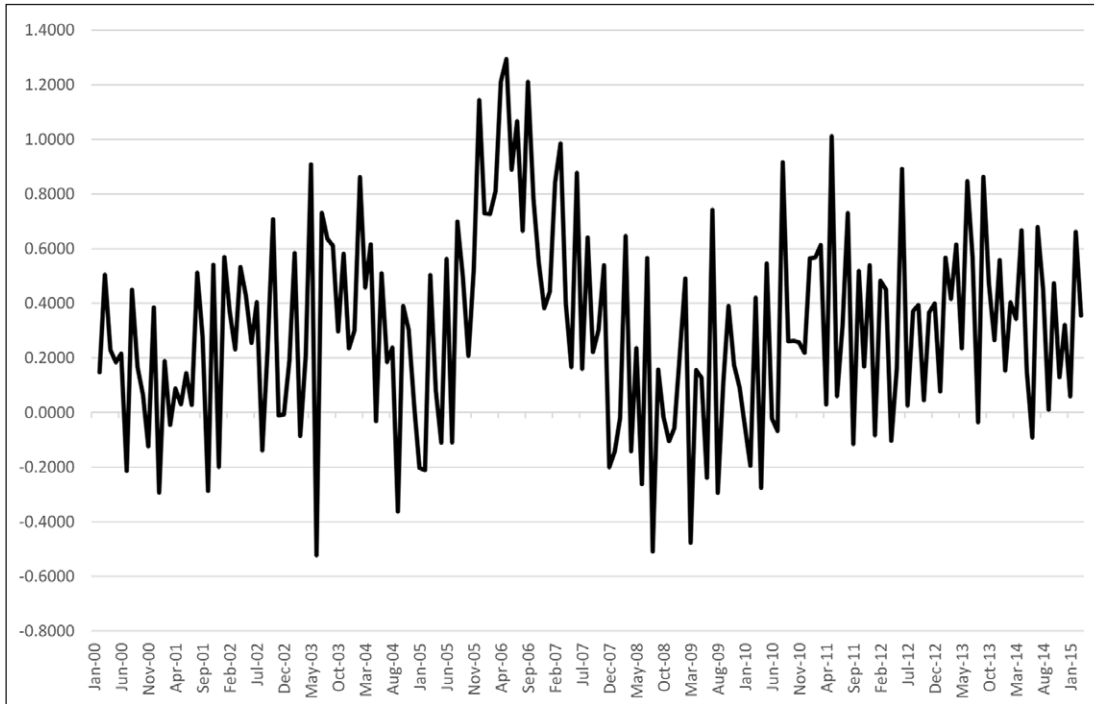


Figure 3: Common factor corrected with GDP information. Rate of monthly variation (%), January 2000–March 2015.

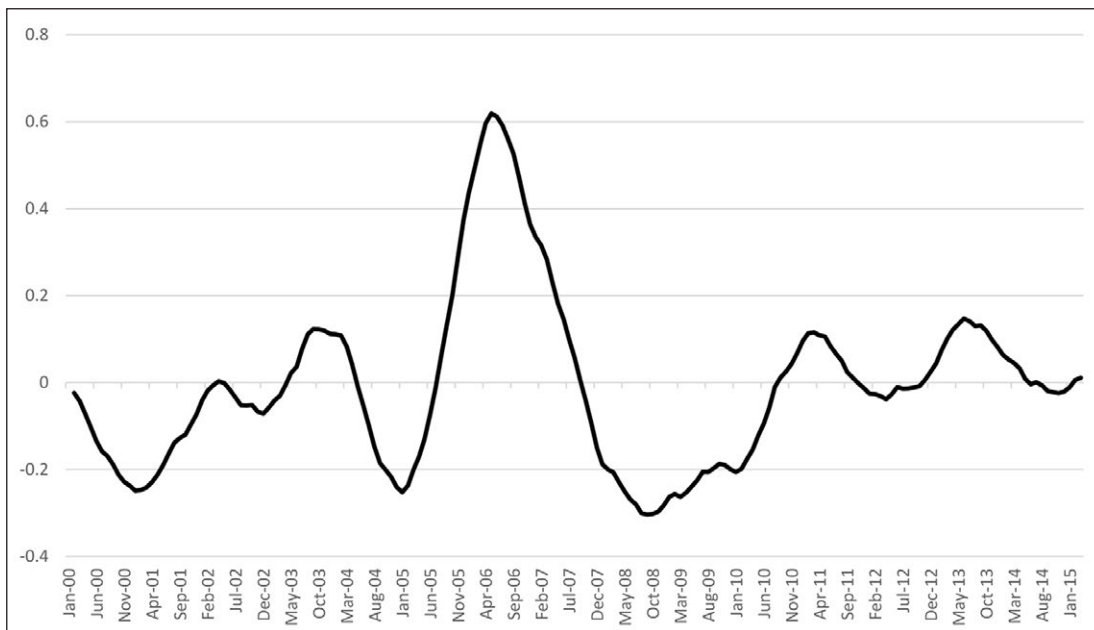


Figure 4: Cyclical components of the common factor (IMAE). Rates of monthly variation (%), January 2000–March 2015.

the indicator: the strong, continued evolution of the common factor. Indeed, the estimated cycle is the monthly indicator of economic activity (IMAE) proposed for the department of Valle del Cauca.

Figure 6 shows the IMAE and the annual growth rates of GDP for Valle del Cauca. It is observed that when the IMAE exceeds zero, the GDP grows above its historical average (3.8 percent). This occurs in 2004, 2006, 2007, 2011, and 2013. In 2012, the IMAE remained near zero, and the economy grew by 3.79 percent. In

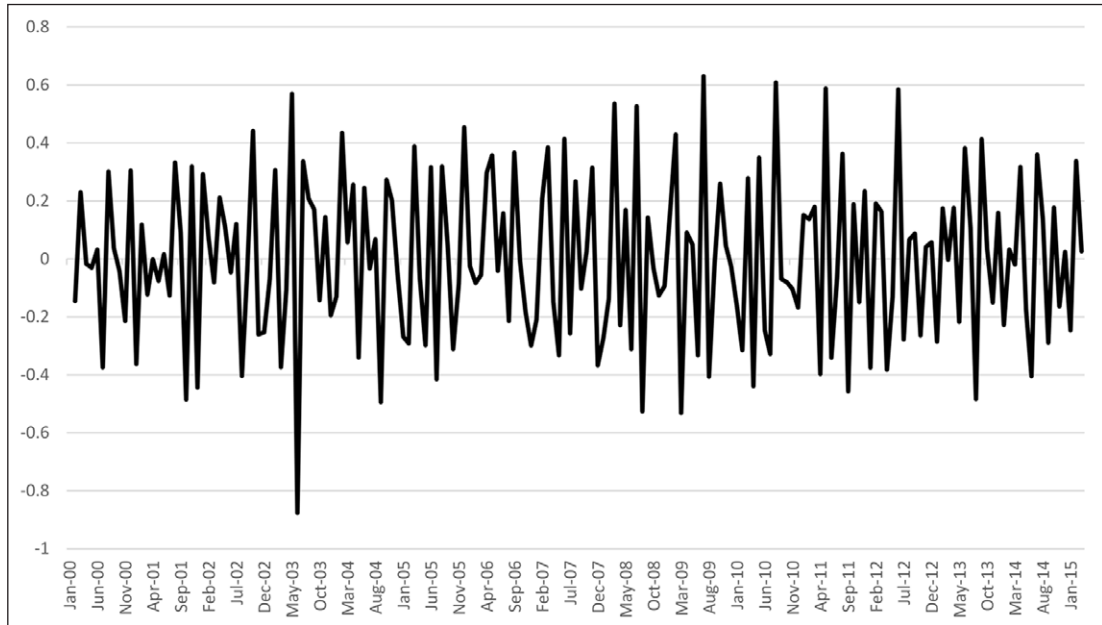


Figure 5: Irregular components of the common factor (IMAE). Rates of monthly variation (%), January 2000–March 2015.

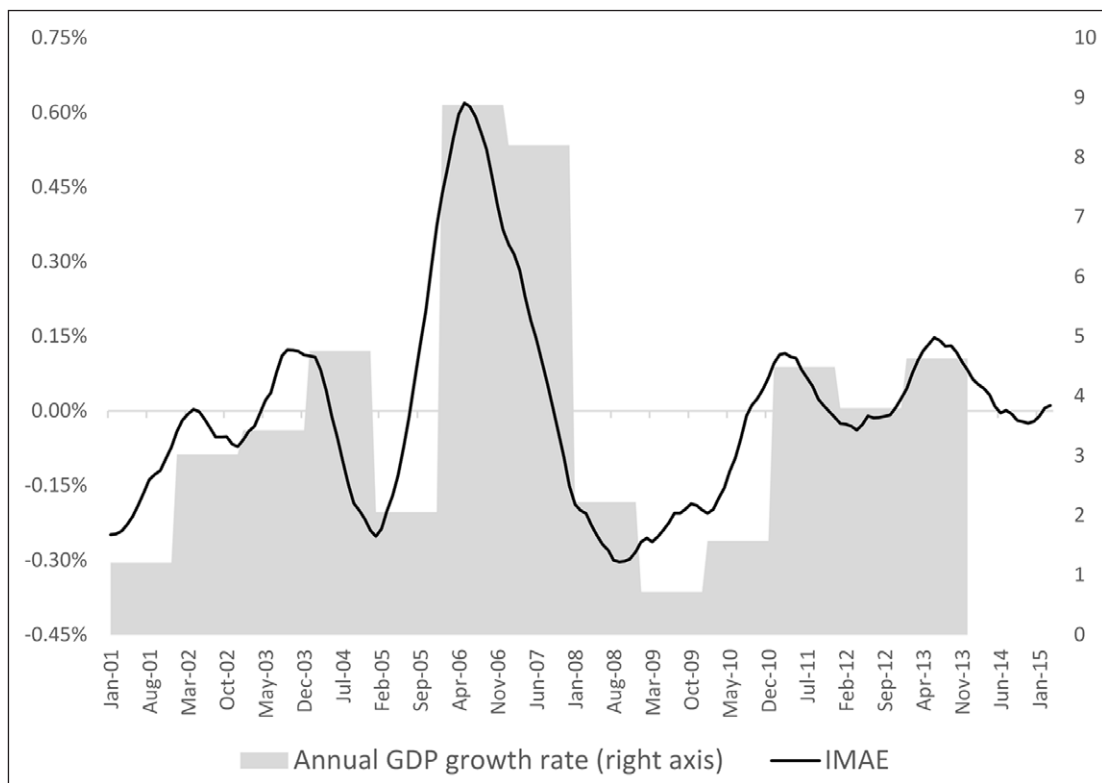


Figure 6: IMAE and annual GDP growth.

the years when the IMAE is more negative, the GDP grows more slowly, as is the case of 2001, 2005, 2009, and 2010. The correlation between the average annual value of the IMAE and the annual GDP growth rate reported by DANE is 0.84.

Table A3 in the appendix shows the maximum and minimum points of the estimated growth cycle. The last turning point of the IMAE (0.14) is reached in June 2013. In this year, the department showed its fastest growth rate in the previous five years (GDP grew by 4.6 percent). Since then, the IMAE has estimated a slowdown, which may have bottomed out in November 2014. In the first quarter of 2015, Valle del Cauca grew at values that are very close to its historical average (3.8 percent) and below the growth in the first quarter of 2014.

An analysis of the contribution of the variables that constitute the IMAE suggests that from November 2014 to March 2015, industry in Valle de Cauca seems to be leading the growth, replacing vehicle sales and foreign trade, which were the elements that contributed to the economy's dynamism in the first three quarters of 2014.

Finally, we calculate the correlations between IMAE and two monthly indicators of national economic activity, *Indicador de Seguimiento a la Economía (ISE)*, prepared by DANE, and *Índice Mensual de Actividad Económica Colombiana (IMACO)*, calculated by the central bank. The correlations are 0.45 and 0.34, respectively. They are relatively low correlations, which suggest that the economic cycle in the region has a high degree of independence in relation to the macroeconomic business cycle. This result further justifies the need for the Valle del Cauca to have its own monthly indicator of economic activity.

Conclusion

Although it is typically applied in the construction of indicators of economic activity at the macroeconomic level, in this paper, a version of the methodology of Stock and Watson (1991) was applied to a regional context.

On the basis of seven key historical series from the department of Valle del Cauca in the period from January 2000 to March 2015, a monthly coincident economic activity indicator was constructed. To that end, a methodology was employed that uses dynamic factor models estimated with the Kalman filter. Additionally, the estimated common factor for the series was adjusted to annual GDP growth data and then smoothed for a more stable signal about the evolution of economic activity. We present a three-step methodology that can be easily replicated in other regions in Latin America.

The regional industrial production index, vehicle sales, and exports are series with greater weight within the estimated indicator. These are the variables that, to a greater extent, determine the economic cycle of the department. This signal approximates a cycle for the department's economic activity, with an average periodicity of four years and ten months.

The IMAE reveals a low correlation between the economic cycle of Valle del Cauca and the macroeconomic business cycle in Colombia. This can be explained by the peculiarities of the economic structure of Valle del Cauca. Unlike the predominant structure in the country, Valle del Cauca does not depend on extractive activities such as oil and minerals, has rather a diversified industry, trade concentrated with border partners, an agricultural predominance of sugar cane, and consumption fueled largely by remittances. While in general the Colombian economy has been slowing from 4.9 percent in 2013 to 2.5 percent in the first quarter of 2016, the economy of the Valle del Cauca shown resilience and maintains a growth rate close to 4 percent.

Such a situation in the cycle and structure of Valle del Cauca is similar to several regions in Latin America. These are regions not directly benefited from the commodities boom and increasing Asian demand in the 2000s. Nevertheless, they can now offer strength and provide momentum to the growth of exports and GDP in the current scenario of low commodity prices and slowing Chinese economy.

A necessary extension of the present work would be to estimate time series models with the IMAE and the leading variables of economic activity, not only to count on nowcasting, but also to be able to project GDP growth in the region based on different scenarios. That is, to develop new regional analysis concerning the transmission mechanisms of different external and domestic shocks on the monthly economic activity of the region, such as variation of interest rates, the exchange rate, and commodity prices.

Appendix

Table A1: All monthly series available.

	Description	Source	Beginning	End	Lag
1	Overall participation rate: Cali and the metropolitan area	DANE	Jan-01	Mar-15	One month
2	Employment rate: Cali and the metropolitan area	DANE	Jan -01	Mar-15	One month
3	Unemployment rate: Cali and the metropolitan area	DANE	Jan -01	Mar-15	One month
4	Thousands of employees: Cali and the metropolitan area	DANE	Jan -01	Mar-15	One month
5	Sugar production: Southwest	Asocaña	Jan-00	Mar-15	One month
6	Ethanol production: Southwest	Asocaña	Oct-05	Mar-15	One month
7	Nonresidential electric energy demand: Valle	X.M. Filial de ISA	Jan-00	Mar-15	Two weeks
8	Industrial gas consumption: Valle	SUI – Superservicios	Jan-00	Mar-15	One month
9	Would you say your home is faring better or worse economically than a year ago?: Cali. Question 1.	Fedesarrollo. Consumer Confidence Survey	Nov-01	Mar-15	One month
10	Would you say you are currently able to save any money or not? – Cali.	Fedesarrollo. Consumer Confidence Survey	Nov-01	Mar-15	One month
11	Vehicle sales: Cali	Asonac	Jan-00	Mar-15	Two months
12	Was the annual growth in sales (quantity) this month higher, lower or equal to the annual growth from this month last year? Cali and the metropolitan area. Question 1.	Banco de la República. Monthly Survey of Economic Expectations (Encuesta Mensual de Expectativas Económicas –EMEE)	DEc-05	Mar-15	One month
13	Consumer Confidence Index (Índice de confianza del consumidor, Cali, ICC.)	Fedesarrollo. Consumer Confidence Survey	Nov-01	Mar-15	Two weeks
14	Exports US\$FOB: Valle	DANE-DIAN	Jan-00	Mar-15	One month
15	Imports US\$CIF: Valle	DANE-DIAN	Jan-00	Mar-15	Two months
16	Index of economic conditions (Índice de condiciones económicas, Cali, ICE)	Fedesarrollo. Consumer Confidence Survey	Nov-01	Mar-15	One month
17	What do you expect the growth of sales (quantity) over the next 12 months to be compared with the growth of sales (quantity) in the past 12 months?: Cali metropolitan area	Banco de la República. Monthly Survey of Economic Expectations (EMEE)	Dec-05	Mar-15	One month

(contd.)

	Description	Source	Beginning	End	Lag
18	Ground sugarcane: Southwest	Asocaña	Jan-00	Mar-15	One month
19	Tons of chicken carcasses: Valle	Fenavi	Jan-05	Mar-15	One month
20	Egg production: Valle	Fenavi	Jan-05	Mar-15	One month
21	Passengers transported by the Alfonso Bonilla Aragón Airport: Cali	Aerocivil	Jan-00	Mar-15	Two weeks
22	Cement shipments: Valle	DANE	Jan-00	Mar-15	One month
23	Industrial Production Index (Índice de producción industrial–national, IPI)	DANE	Jan-00	Mar-15	Six weeks
24	Economic monitoring index (Indicador de seguimiento a la economía–national, ISE)	DANE	Jan-11	Mar-15	Three months
25	Quarterly balance of net loans: Valle	Superfinanciera	Mar-00	Dec-14	Three months
26	Sale of new homes	Galería Inmobiliaria	Aug-06	Mar-15	One month
27	Regional Industrial Production Index (Índice de Producción Industrial Regional, IPIR)	Banco República Cali	Jan-02	Mar-15	Six weeks

Note: Shaded rows indicate that the series was used in estimating the indicator.

Table A2: Description of the treatment used for the series comprising the IMAE.

	Description	Source	Seasonality	Degree of integration	Transformation
1	Ground sugarcane	Asocaña	Yes	I(0)	S,E,D
2	New vehicle sales: Cali	Asonac	Yes	I(1)	S,E,L,D
3	Cement shipments	DANE	Yes	I(1)	S,E,L,D
4	Regional Industrial Production Index, IPIR	Banco República Cali	Yes	I(1)	S,E,L,D
5	Exports US\$FOB: Valle	DANE-DIAN	Yes	I(1)	S,E,L,De,D
6	Imports US\$CIF: Valle	DANE-DIAN	Yes	I(1)	S,E,L,De,D
7	Non-residential energy demand: Valle	X. M. Filial de ISA	Yes	I(1)	S,E,L,D

Note: S = seasonally adjusted series; E = standardized series; L = series transformed with the natural logarithm; De = series deflated by US producer price index (PPI); and D = series transformed with the first difference.

Table A3: Dates of cycles for Valle del Cauca.

Maximum	Minimum
April 2002	December 2002
September 2003	January 2005
May 2006	October 2008
March 2011	April 2012
June 2013	November 2014

Author Information

Pavel Vidal is associate professor at the Pontificia Universidad Javeriana Cali, Colombia. He previously worked at the Central Bank of Cuba. He has been a visiting researcher at Harvard University, Columbia University, Universidad Complutense de Madrid, University of Oslo, and Institute for Developing Economies (Japan External Trade Organization). Vidal has been a consultant to the World Bank, Inter-American Development Bank, Brookings Institution, and Atlantic Council. His areas of specialization include monetary policy and time-series econometric models.

Lya Paola Sierra is associate professor at Pontificia Universidad Javeriana Cali, Colombia, and holds a PhD in economics from Universidad Autónoma de Madrid. Sierra's recent research analyzes the determinants of commodity prices in the international markets and macroeconomic factors affecting the employment and the manufacturing value-added in Colombia and other Latin American countries. On these topics Sierra has published several articles in academic journals. Lately, her work has contributed to create an index to measure the economic activity in Valle del Cauca.

Johana Sanabria is an economist at Centro de Estudios Económicos Regionales (CEER), Banco de la República (central bank of Colombia). Her main areas of interest are regional economy, labor market, macroeconomic indicators, and international trade. Sanabria is currently involved in a project to construct indicators of economic activity for three regions in Colombia.

Jaime Andrés Collazos is an economist at Centro de Estudios Económicos Regionales (CEER), Banco de la República (central bank of Colombia). His research interest include regional economics, especially economic activity indexes, infrastructure, and public finances. He is the author of several articles in academic journals, most recently "Global Ranking of Cities to Attract Industrial Investment into the Latin American Pacific Basin." Collazos is coeditor of *Boletín Económico Regional del Suroccidente* in Colombia.

References

- Angelini, Elena, Marta Banbura, and Gerthard Rünstler. 2008. "Estimating and Forecasting the Euro Area Monthly National Accounts from a Dynamic Factor Model." *Journal of Business Cycle Measurement and Analysis* 1: 1–22.
- Arango, Luis Eduardo, Fernando Arias, Luz Adriana Flórez, and Munir Jalil. 2008. "Cronología de los ciclos de negocios recientes en Colombia." *Lecturas de economía* 68: 9–37.
- Aruoba, S., Francis Diebold, and Chiara Scotti. 2009. "Real-Time Measurement of Business Conditions." *Journal of Business and Economic Statistics* 27: 417–427. DOI: <https://doi.org/10.1198/jbes.2009.07205>
- Banco de la República. 2015. *Boletín Económico Regional de Suroccidente*, 1st trimester.
- Boivin, Jean, and Serena Ng. 2006. "Are More Data Always Better for Factor Analysis?" *Journal of Econometrics* 132: 169–194. DOI: <https://doi.org/10.1016/j.jeconom.2005.01.027>
- Camacho, Máximo, and Gabriel Pérez-Quirós. 2010. "Introducing the Euro-STING: Short-Term Indicator of Euro Area Growth." *Journal of Applied Econometrics* 25: 663–694. DOI: <https://doi.org/10.1002/jae.1174>
- Camacho, Máximo, Gabriel Pérez-Quirós, and Pilar Poncela. 2013. "Short-Term Forecasting for Empirical Economists. A Survey of the Recently Proposed Algorithms." Working Paper No. 1318, Banco de España, Madrid.
- Camacho, Máximo, and Rafael Doménech. 2012. "MICA-BBVA: A Factor Model of Economic and Financial Indicators for Short-Term GDP Forecasting." *Journal of the Spanish Economic Association* 3: 475–497. DOI: <https://doi.org/10.1007/s13209-011-0078-z>
- Castro, Carlos Alberto. 2003. "Yet Another Lagging, Coincident and Leading Index for the Colombian Economy." *Archivos de economía* 233: 1–26.
- Chow, Gregory, and An-loh Lin. 1971. "Best Linear Unbiased Distribution and Extrapolation of Economic Time Series by Related Series." *Review of Economics and Statistics* 53(4): 372–375. DOI: <https://doi.org/10.2307/1928739>
- Crone, Theodore. 2005. "Consistent Economic Indexes for the 50 States." *Review of Economics and Statistics* 87(4): 593–603. DOI: <https://doi.org/10.1162/003465305775098242>
- Escobar, Julio. 2015. "Indicador mensual de la producción industrial regional a partir de los Índices de la EMM del DANE." Working paper, Banco de la República, Bogotá.
- Escobar, Julio, Stefany Moreno, and Jaime Collazos. 2013. "Composición de la economía de la región suroccidente de Colombia." *Ensayos sobre economía regional* 52: 3–10.
- Fernández, Roque. 1981. "Methodological Note on the Estimation of Time Series." *Review of Economic and Statistics* 63: 471–478. DOI: <https://doi.org/10.2307/1924371>

- Forni, Mario, Marc Hallin, Marco Lippi, and Lucrezia Reichlin. 2009. "Coincident and Leading Indicators for the Euro Area." *Economic Journal* 111: 62–85. DOI: <https://doi.org/10.1111/1468-0297.00620>
- Gómez, Victor, and Augustin Maravall. 1996. "Programs TRAMO and SEATS: Instructions for the User." Working Paper No. 9628, Servicio de Estudios, Banco de España.
- Guerrero, Santiago, and Juan Carlos Martínez-Ovando. 2014. "Índices de difusión generalizados de la actividad económica estatal y sectorial para México." Working Paper No. 16, Banco de México.
- Gupta, Rangan, and Alain Kabundi. 2011. "A Large Factor Model for Forecasting Macroeconomic Variables in South Africa." *International Journal of Forecasting* 27: 1076–1088. DOI: <https://doi.org/10.1016/j.ijforecast.2010.10.001>
- Haldane, Andrew, and Stephen Hall. 1991. "Sterling's Relationship with the Dollar and the Deutschemark: 1976–1989." *Economic Journal* 101: 436–443. DOI: <https://doi.org/10.2307/2233550>
- Hamilton, James. 1994. *Time Series Analysis*. Princeton, NJ: Princeton University Press.
- Harvey, Andrew. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. New York: Cambridge University Press.
- Hurtado, Jorge Luis, and Luis Fernando Melo. 2010. "Una metodología multivariada de desagregación temporal." *Borradores de economía* 586: 1–36.
- Jorrat, Juan Mario. 2005. "Construcción de índices compuestos mensuales coincidente y líder de Argentina." In *Progresos en econometría*, edited by Mariana Marchionni, 43–100. Buenos Aires: Asociación Argentina de Economía Política.
- Kalman, Rudolf. 1960. "A New Approach to Linear Filtering and Prediction Problems." *Journal of Basic Engineering* 82: 35–45. DOI: <https://doi.org/10.1115/1.3662552>
- Kamil, Herman, José David Pulido, and José Luis Torres. 2010. "El IMACO: Un índice mensual líder de la actividad económica de Colombia." *Borradores de economía* 609: 1–40.
- Kim, Chang-Jin, and Charles Nelson. 1989. "The Time-Varying-Parameter Model for Modeling Changing Conditional Variance: The Case of the Lucas Hypothesis." *Journal of Business and Economic Statistics* 7: 433–440.
- Koopman, Siem Jan, Andrew Harvey, Jurgen Doornik, and Neil Shephard. 2009. *STAMP 8.2: Structural Time Series Analyser, Modeler, and Predictor*. London: Timberlake Consultants.
- Koopman, Siem Jan, Neil Shephard, and Jurgen Doornik. 1999. "Statistical Algorithms for Models in State Space using SsfPack 2.2 (with discussion)." *Econometrics Journal* 2: 113–166. DOI: <https://doi.org/10.1111/1368-423X.00023>
- Litterman, Robert. 1983. "A Random Walk, Markov Model for the Distribution of Time Series." *Journal of Business and Economic Statistics* 1: 169–173.
- Marcillo, Edgar. 2013. "Un indicador líder para la actividad económica de Colombia." *Archivos de economía* 404: 1–41.
- Megna, Robert, and Qiang Xu. 2003. "Forecasting the New York State Economy: The Coincident and Leading Indicators Approach." *International Journal of Forecasting* 19: 701–713. DOI: [https://doi.org/10.1016/S0169-2070\(03\)00002-5](https://doi.org/10.1016/S0169-2070(03)00002-5)
- Melo, Luis Fernando, and Fabio Nieto. 2001. "About a Coincident Index for the State of the Economy." *Borradores de economía* 194: 1–32.
- Melo, Luis Fernando, Fabio Nieto, Carlos Posada, Yanneth Betancourt, and Juan David Baron. 2001. "Un índice coincidente para la actividad económica colombiana." *Borradores de economía* 195: 1–66.
- Melo, Luis Fernando, Fabio Nieto, and Mario Ramos. 2003. "A Leading Index for the Colombian Economic Activity." *Borradores de economía* 243: 1–57.
- Poncela, Pilar, Eva Senra, and Lya Paola Sierra. 2014. "Common Dynamics of Nonenergy Commodity Prices and Their Relation to Uncertainty." *Applied Economics* 46: 3724–3735. DOI: <https://doi.org/10.1080/0036846.2014.939377>
- Quilis, Enrique. 2001. *Notas sobre desagregación temporal de series económicas*. Madrid: Instituto Nacional de Estadística.
- Rivero, Andres David. 2007. "El índice compuesto coincidente mensual de la actividad económica de Córdoba (ICA-COR) 1994–2006." *Revista de economía y estadística* 45(1): 31–73.
- Schumacher, Christian. 2007. "Forecasting German GDP Using Alternative Factor Models Based on Large Datasets." *Journal of Forecasting* 26: 271–302. DOI: <https://doi.org/10.1002/for.1026>
- Stock, James, and Mark Watson. 1991. "A Probability Model of the Coincident Economic Indicators." In *The Leading Economic Indicators: New Approaches and Forecasting Records*, edited by Kajal Lahiri,

and Geoffrey Moore, 63–90. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9781139173735.005>

Vidal, Pavel, and Annia Fundora. 2008. "Trade-Growth Relationship in Cuba: Estimation Using the Kalman Filter." *CEPAL Review* 94: 97–116.

How to cite this article: Vidal, P, Sierra, L P, Sanabria, J and Collazos, J A. A Monthly Regional Indicator of Economic Activity: An Application for Latin America. *Latin American Research Review*. 2017; 52(4), pp. 589-605. DOI: <https://doi.org/10.25222/larr.38>

Submitted: 30 November 2015

Accepted: 29 July 2016

Published: 23 October 2017

Copyright: © 2017 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



Latin American Research Review is a peer-reviewed open access journal published by the Latin American Studies Association.

OPEN ACCESS 

The Open Access icon consists of a stylized padlock with a circular arrow around it, indicating that the content is freely available.