Statistics and Applications {ISSN 2454-7395 (online)}
Volume 16 No.2, (New Series), pp. 89-102

An Application of Segmented Trend Estimation and Forecasting with Controlled Smoothness via Penalized Least Squares

Víctor M. Guerrero[1], Eliud Silva[2] and Alejandro Islas-Camargo[3]
[1, 3]Department of Statistics - ITAM, Ciudad de México, MÉXICO
[2]Universidad Anáhuac México, Estado de México, MÉXICO

Received: November 15, 2017; Revised: October 03, 2018; Accepted: October 15, 2018

Abstract

We apply a filtering methodology to estimate the segmented trend of a time series and to get forecasts with the underlying unobserved-component model of the filter employed. The trend estimation technique is based on Penalized Least Squares. The application of this technique allows us to control the amount of smoothness in the trend by segments of the data range, corresponding to different regimes. The method produces a smoothing filter that mitigates the effect of outliers and transitory blips, thus capturing an adequate smooth trend behavior for the different regimes. Obtaining forecasts becomes an easy task once the filter was applied to the series, because the unobserved-component model underlying the filter has parameters directly related to the smoothing parameter of the filter. The empirical application is useful to appreciate the appropriateness of our segmented trend approach to capture the underlying behavior of the annual rate of growth of remittances to Mexico.

Key words: Hodrick-Prescott filter, Penalized Least Squares, controlled smoothing, changing regimes, smoothing parameter

1. Introduction

Here, we propose to use a filtering methodology based on Penalized Least Squares (PLS) to estimate the smooth trend of a time series and to get forecasts with the unobserved-component model underlying the filter. There are several methods that can be used to produce smooth trends, such as smoothing splines (*e. g*. Wahba, 1990; Eilers and Marx, 1996; Paige and Trindade, 2010) and methods based on parametric or semiparametric models (Ruppert et al., 2003; Chandler and Scott, 2011). On the one hand, smoothing splines is a widely used and powerful method that can be used to smooth any type of time series, but to the best of our knowledge it does not consider the possibility of a change in regimes. On the other hand, using parametric or semiparametric trend representations are also very powerful techniques that may include changing regimes, mainly when the analyst knows the source of such a change or at least the date of the regime change. For instance, one may try using a structural model where the factors affecting the variable of interest are used as explanatory variables. Nevertheless, in a forecasting exercise with this type of models the analyst must provide forecasts of the future behavior of the exogenous variables. In the empirical application of this work we found just one attempt at forecasting remittances with a structural

Corresponding Author: Eliud Silva
Email: jose.silva@anahuac.mx

model (see Mohapatra and Ratha, 2010) and the authors of such work recognized that the data quality for this type of exercise requires a lot of improvement.

Another method that may seem appropriate at first sight is Box and Tiao's (1975) intervention analysis. This method requires knowledge of the moment when an intervention took place and the main idea of the analysis is to measure the effect of such an intervention on the time series under study. To that end, a discrete dynamic function should be postulated to capture the behavior of the intervention across time and its parameters must be estimated together with the other model parameters. In the empirical situation considered here there is no reason to believe that an explicit intervention affecting the time series took place at a given time point. In fact, there is always the possibility of many host country interventions, such as changes in migration policies and structural macroeconomic changes, *e.g*. an increase in wages. There may also be interventions derived from the home country policies, such as a change in the banking system and changes in macroeconomic conditions in general.

The method applied here is based on the work of Guerrero and Silva (2015) that is useful to control the percentage of smoothness of the trend both globally and by segments. We are convinced that estimation of trends must be done with controlled smoothness, as it has been advocated in Guerrero (2007), Guerrero and Silva (2015) and Guerrero et al. (2017), to establish valid comparisons for trends of a time series during different periods of observation and/or with different sample sizes, or even for time series of different variables.

This work originated from the need of studying the growth rate of remittances flows to Mexico, which is an important variable for developing countries in general. In some of these countries, the size of remittances is larger than that of traditional sources, such as official aids and private capital flows. See, for example, Rao and Hassan (2011) and Nyamongo et al. (2012). According to World Bank (2016), migrant remittances flows to developing countries reached $432 billion U.S. dollars in 2015. The value of remittances is not only more than three times the volume of official aid in some countries, but it amounts to more than 10 percent of gross domestic product in many of them and has almost the same size as foreign direct investment in some others (see World Bank, 2011). Thus, remittances are an important building block of economic development.

Remittances receipts have been associated with reduction in poverty, increased household resources devoted to investment, improved health and education outcomes, and higher levels of entrepreneurship, see Adams and Page (2005), Amuedo-Dorantes and Pozo (2011), Fajnzylber and Lopez (2007), Hildebrandt and McKenzie (2005) and Valero-Gil (2009). They can also improve recipient household's access to formal financial services, as mentioned in Giuliano and Ruiz-Arranz (2009) and Gupta et al. (2009). Given the size and increasing importance of migrant remittances to developing countries, there is a concern that a decline of remittances flows could affect the poorest countries and households that heavily depend on them. Therefore, policymakers need forward-looking analyses of remittances, trying to foresee the growth rate of remittances flows in the short and medium term.

Figure 1 shows the quarterly series of annual percent growth rate (i.e. year/year) of remittances flows to Mexico from the U.S. over the period 1996:I - 2016:IV, together with its estimated trend obtained with the so-called Hodrick-Prescott (HP) filter. This trend plays the role of a dynamic mean of the process, so that the smooth corresponds to the estimated trend. The smoothing parameter of this filter was chosen for this exercise as $\lambda = 266.2$ in order to produce 90% smoothness for the trend (we explain this idea in the following sections). In

Figure 1 we appreciate three different regimes, from 1996:I to 2003:IV (fast positive change of the growth rate), from 2004:I to 2009:IV (fast negative change of the growth rate) and from 2010:I onwards (medium positive change of the growth rate). Associated with these regimes we also see the following variability behavior. High variability in the first regime (10.37, as measured by the standard deviation about the estimated trend), low variability in the second (6.95) and similar or slightly lower variability (6.46) in the third regime. This fact led us to look for a segmentation of the series that will improve the forecasting ability of the technique when using the most recent segment of data, which usually contains the most useful information for forecasting purposes.

Figure 1: Quarterly series of annual percent growth rates of remittances to Mexico, 1996:I – 2016:IV and estimated trend obtained with the HP filter.

Source: *BANXICO (http://www.banxico.org.mx)*

The organization of this paper is as follows. Section 2 presents the PLS methodology that will be used to estimate the segmented trend. It is seen that the method is supported by very weak assumptions, *i.e.*, basically the existence of second order moments of the time series under study. Therefore, there is no need to verify model assumptions to validate the results. A consequence of the lack of assumptions is that no statistical inference can be made, except that of obtaining point trend estimates together with a measure of their uncertainty, expressed in terms of standard errors. In Section 3 we show how to produce forecasts from the filter previously employed for trend estimation. We recognize that these forecasts come from an unobserved-component model underlying the filter, whose parameters are related to the smoothing parameter of the filter. Section 4 is devoted to the empirical application of the methodology to the annual percent growth rates of remittances from U.S. to Mexico. Firstly, we show how to estimate the segmented trend and secondly, how to produce the required forecasts from the previously employed filter. In Section 5 we conclude with some final remarks.

The main contributions of this work are the following. (1) An application of the method proposed by Guerrero and Silva (2015) to estimate the segmented trend of remittances flows to Mexico via filtering by PLS, in order to capture the trend behavior appropriately. (2) Linking the smoothing parameter of the PLS filter with the parameters of the time series model underlying the filter, as previously shown by Kaiser and Maravall (2005) and McElroy (2008), thus avoiding the need of estimating parameters with a very short time series. (3) Obtaining forecasts of the time series under study based on the last segment of the series and its corresponding model, that is shown to provide good results.

2. Trend estimation by data segments

Let us assume that the observed time series (the annual percent growth rate of remittances in the present case) can be expressed as a signal-plus-noise model. This is not because we believe that the data were generated this way, but just to take into account the empirical regularities in the data, that is, $y_t = \tau_t + \eta_t$ where $\{\tau_t\}$ is the trend (or signal) and $\{\eta_t\}$ is the noise of $\{y_t\}$, for $t = 1, \ldots, N$. Then, we propose to use PLS for estimating the trend and we have to solve the following minimization problem

$$\min_{\{\tau_t\}} \{\sum_{t=1}^{N} \frac{1}{\sigma_1^2}(y_t - \tau_t)^2 + \sum_{t=31}^{N} \frac{1}{\sigma_0^2}(\nabla^2 \tau_t)^2\} \tag{1}$$

where σ_1^2 is the variance of the deviation from trend (y_t - τ_t) and σ_0^2 is the variance of the differenced trend ($\nabla^2 \tau_t$). Thus, the trend behavior is expressed by means of a second order difference, *i.e.,* it is expressed as $\nabla^2 \tau_t = \tau_t - 2\tau_{t-1} + \tau_{t-2}$. In this setting, the smoothing parameter $\lambda = \sigma_1^2 / \sigma_0^2$ serves to balance fidelity of the trend to the original data against smoothness of the trend, in such a way that as $\lambda \downarrow 0$ the trend gets closer to the original data and as $\lambda \uparrow \infty$ the trend behaves as a polynomial of order 1.

By solving the minimization problem (1) we obtain the HP filter that provides trend estimates for the whole range of observations ($t = 1, \ldots, N$) and can also be used to forecast future realizations of the series $\{y_t\}$, for $t = N+1, N+2, \ldots$. In the present case, we just want to describe the underlying pattern of the data with the aid of a trend and then obtain forecasts of the original series. We are not interested in using a formal statistical model or in making other type of statistical inference. Thus, rather than formally estimating the smoothing parameter λ we calibrate its value, and this becomes the most important practical decision to be made.

Sometimes the trend changes abruptly due to changes in variance, a fact reflected in the smoothness of the trend. Thus, we consider here an extension of the minimization problem posed by (1) by allowing different trend behaviors for segments of the data range, which are in turn linked to different variances, one for each regime. In that case, we pose a minimization problem that accounts for different trends in two adjacent segments and hence, two different λ values must be calibrated. This is a relevant situation when forecasting since the last segment of data is the one actually used to generate forecasts. More than two data segments could also be considered without difficulty, as shown by Guerrero and Silva (2015), but we think that considering only two segments is appropriate for forecasting purposes. So, we consider here only the two-segment problem

$$\min_{\{\tau_t\}} \left\{ \sum_{t=1}^{N_1} \frac{1}{\sigma_1^2}(y_t - \tau_t)^2 + \sum_{t=N_1+1}^{N} \frac{1}{\sigma_2^2}(y_t - \tau_t)^2 + \sum_{t=3}^{N} \frac{1}{\sigma_0^2}(\nabla^2 \tau_t)^2 \right\} \tag{2}$$

where σ_1^2 and σ_2^2 are the variances of the first and second data segments, with N_1 and $N_2 = N$

– N_1 data points, correspondingly. Here, we let $\{y_t\}$ be the observed series and $\{\tau_t\}$ be the trend.

The unobserved-component model that underlies the minimization problem posed by (2) can be written as

$$y_t = \tau_t + \eta_{1,t} \text{ with } \eta_{1,t} \sim (0,\sigma_1^2) \text{ for } t = 1, \ldots, N_1, \tag{3}$$

$$y_t = \tau_t + \eta_{2,t} \text{ with } \eta_{2,t} \sim (0,\sigma_2^2) \text{ for } t = N_1+1, \ldots, N, \tag{4}$$

$$\nabla^2 \tau_t = \varepsilon_t \text{ with } \varepsilon_t \sim (0,\sigma_0^2) \text{ for } t = 3, \ldots, N, \tag{5}$$

where we write $\nu \sim (0,\sigma_\nu^2)$ to say that the random variable ν has mean 0 and variance σ_ν^2. The sequences of random errors $\{\eta_{i,t}\}$ for i = 1, 2, are serially uncorrelated, and $\{\varepsilon_t\}$ is another sequence of serially uncorrelated random errors that is also mutually uncorrelated with the previous two sequences.

The matrix representation of model (3) - (5) becomes

$$\mathbf{y} = \boldsymbol{\tau} + \begin{pmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \end{pmatrix} \text{ and } K\boldsymbol{\tau} = \boldsymbol{\varepsilon} \tag{6}$$

where K is an $(N\text{-}2)\times N$ matrix difference operator whose i,j-th entry is the binomial coefficient $K(i, j) = (-1)^{2+i-j} 2!/[(j-i)!(2-j+i)!]$, for $i = 1, \ldots, N-2$ and $j = 1, \ldots, N$, with $K(i, j) = 0$, for $j < i$ or $j > 2+i$. The vectors $\mathbf{y}$ and $\boldsymbol{\tau}$ contain the original observations and the trend, respectively, while $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ as well as $\boldsymbol{\varepsilon}$ are random vectors. Then, an application of Generalized Least Squares (GLS) produces the Best Linear Unbiased Estimator (BLUE) of the vector of trends, that can be expressed as

$$\hat{\boldsymbol{\tau}} = \begin{pmatrix} I_{N_1} + \lambda_1(K_1'K_1 + k_1'k_1) & \lambda_1 k_1' k_2 \\ \lambda_2 k_2' k_1 & I_{N_2} + \lambda_2(K_2'K_2 + k_2'k_2) \end{pmatrix}^{-1} \mathbf{y} \tag{7}$$

with I_{N_i} the N_i-dimensional identity matrix and $\lambda_i = \sigma_i^2/\sigma_0^2$, for i = 1, 2. The matrices K_i are of size $(N_i\text{-}2)\times N_i$, for $i = 1, 2$, and have the same form as K. In fact, we have

$$K = \begin{pmatrix} K_1 & 0 \\ k_1 & k_2 \\ 0 & K_2 \end{pmatrix} \text{ with } k_1 \text{ of size } 2\times N_1 \text{ and } k_2 \text{ of size } 2\times N_2. \tag{8}$$

The BLUE given by (7) is also called predictor, since we are in fact estimating the realization of a random vector rather than a vector of constants (a justification for using GLS in this context was provided by Guerrero, 2007). Moreover, the variance-covariance matrix of the GLS estimator is given by

$$\Gamma = Var(\hat{\boldsymbol{\tau}}) = \left[\begin{pmatrix} \sigma_1^{-2} I_{N_1} & 0 \\ 0 & \sigma_2^{-2} I_{N_2} \end{pmatrix} + \sigma_0^{-2} \begin{pmatrix} K_1'K_1 + k_1'k_1 & k_1'k_2 \\ k_2'k_1 & K_2'K_2 + k_2'k_2 \end{pmatrix}\right]^{-1} \tag{9}$$

and unbiased estimators of the error variances are given by $\hat{\sigma}_0^2 = RSS/(N-2)$ and $\hat{\sigma}_i^2 = \lambda_i \hat{\sigma}_0^2$ for i = 1, 2, with RSS the Residual Sum of Squares (see Guerrero and Silva, 2015 for details). To be specific, the k_i arrays are as follows

$$k_1 = \begin{pmatrix} 0_{2\times(N_1-2)} & \begin{matrix} 1 & -2 \\ 0 & 1 \end{matrix} \end{pmatrix} \text{ and } k_2 = \begin{pmatrix} \begin{matrix} 1 & 0 \\ -2 & 1 \end{matrix} & 0_{2\times(N_2-2)} \end{pmatrix} \qquad (10)$$

where $0_{2\times(N_i-2)}$ denotes a zero matrix, for i = 1, 2.

To be able to apply (7) - (9) we must provide the values of λ_1 and λ_2, as well as the cutoff point N_1. The smoothing parameters will be chosen by applying the controlled smoothness approach (see Guerrero, 2007) that is based on measuring the relative precision attributable to the smoothness specification, that is, the second equation in model (6). To this end we should notice that the total precision achieved by estimating the trend is the inverse of the variance-covariance matrix given by (9), that is, it is provided by Γ^{-1}. Therefore, the amount of smoothness to be attained globally by the trend for the unsegmented time series is given by $S = 1 - tr(I_N + \lambda K'K)^{-1} / N$. Now, it is important to keep in mind that the amount of smoothness to be achieved globally is bounded by $1 - 2/N$ as $\lambda \to \infty$. This fact follows because the trace appearing in the expression of S can be written in terms of the d nonzero eigenvalues of $K'K$, that is, $e_1, ..., e_{N-2} > 0$, as $tr(I_N + \lambda K'K)^{-1} = (1+\lambda e_1)^{-1} + ... + (1+\lambda e_{N-2})^{-1} + 2$ so that the trace tends to 2 as $\lambda \to \infty$.

We now measure the precision shares for each segment by means of the following indices (see Guerrero and Silva, 2015 for details)

$$S_i(\lambda_1, \lambda_2; N) = tr[B_{i,\lambda}(I_N + B_{1,\lambda} + B_{2,\lambda})^{-1}]/N \quad \text{for } i = 1, 2, \qquad (11)$$

which quantify the smoothness achieved by smoothing segment i of the data, where

$$B_{1,\lambda} = \frac{N_1}{N} \begin{pmatrix} \lambda_1(\frac{N}{N_1} K_1'K_1 + k_1k_1') & 0 \\ 0 & \lambda_2 I_{N_2} \end{pmatrix}, \; B_{2,\lambda} = \frac{N_2}{N} \begin{pmatrix} \lambda_1 I_{N_1} & 0 \\ 0 & \lambda_2(\frac{N}{N_2} K_2'K_2 + k_2k_2') \end{pmatrix}. \qquad (12)$$

Thus, if we fix the values of the indices $S_i(\lambda_1, \lambda_2; N)$ for i = 1, 2, or equivalently the percentages of smoothness obtained by multiplying those indices by 100, the smoothing parameters are obtained by solving expression (11) numerically for λ_1 and λ_2.

The cutoff point can be chosen by applying the search procedure suggested in Guerrero and Silva (2015). That is, we first decide the amount of smoothness to be attained globally by the trend. Then, we choose the smoothness S_1 to be achieved in the first segment, recalling that the more variable the series the more smoothness required for the trend. Finally, the amount of smoothness for the second segment gets fixed by $S_2 = (NS - N_1S_1) / (N - N_1)$. Thus, we search for the optimal cutoff point as the value of N_1 that minimizes the estimated error variance $\hat{\sigma}_0^2$. There is no guarantee that such a cutoff point exists in practical applications and when that happens the segmentation is deemed appropriate, otherwise there is no need to split the series into segments. This is a data-based and easy-to-use method that is in line with our proposal for choosing the smoothing parameter. Nevertheless, when appropriate, subject matter considerations could lead to selecting the value of N_1 exogenously.

3. Obtaining forecasts

Forecasts of the original series can be obtained from the unobserved component model underlying the HP filter for the last segment of the time series. So, let us consider an unobserved component representation similar to (3) - (5), except that we now introduce a constant μ to account for the possible presence of a drift in the most recent segment of the trend. That is,

$$y_t = \tau_t + \eta_t \text{ for } t = N_1+1, \ldots, N \text{ and } \nabla^2 \tau_t = \mu + \varepsilon_t \text{ for } t = N_1+3, \ldots, N \quad (13)$$

where $\{\eta_t\}$ and $\{\varepsilon_t\}$ are mutually uncorrelated processes. The presence of a drift in the trend specification provides a better fit to the end of the series, since it amounts to having an underlying polynomial behavior of order 2 rather than 1, which is the case when the drift is not included (see Guerrero, 2007). Thus, representation (13) produces the following Integrated Moving Average model of order (2,2), that is, the IMA(2,2) model

$$\nabla^2 y_t = \mu + \varepsilon_t + \nabla^2 \eta_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \text{ for } t = N_1+3, \ldots, N, \quad (14)$$

where $\{a_t\}$ behaves as a zero-mean White Noise process with variance σ_a^2.

An unbiased estimator of μ is provided by the sample mean of $\{\nabla^2 y_t\}$, that is $\hat{\mu} = \sum_{t=N_1+3}^{N} \nabla^2 y_t /(N - N_1 - 3)$. Moreover, θ_1 and θ_2 are parameters fully determined by the value of the smoothing constant. Once such parameters are obtained, we can get the h-period-ahead optimal forecast, in minimum Mean Square Error (MSE) sense, with origin at time $t = N_1+3, \ldots, N$, as follows,

$$\hat{y}_t(1) = \mu + 2y_t - y_{t-1} - \theta_1 a_t - \theta_2 a_{t-1}, \quad \hat{y}_t(2) = \mu + 2\hat{y}_t(1) - y_t - \theta_2 a_t$$
$$\text{and } \hat{y}_t(h) = [(h-2)(h-1)/2]\mu + (h-1)\hat{y}_t(2) - (h-2)\hat{y}_t(1) \text{ for } h \geq 3, \quad (15)$$

so that all the forecasts for $h \geq 2$ depend on both the one-step-ahead and two-step-ahead forecasts.

3.1 Relating the IMA parameters to the smoothing parameter of the HP filter

The HP filter makes use of a smoothing parameter λ that is related to the corresponding IMA(2,2) parameters by equating variances and covariances of the processes involved. From the MA(2) structure of the differenced series $\{\nabla^2 y_t\}$ we get the equation (i) $Var(\nabla^2 y_t) = (1+\theta_1^2+\theta_2^2)\sigma_a^2 = \sigma_\varepsilon^2 + 6\sigma_\eta^2$, as well as the equations for the covariances (ii) $Cov(\nabla^2 y_t, \nabla^2 y_{t-1}) = (-\theta_1 + \theta_1\theta_2)\sigma_a^2 = -4\sigma_\eta^2$ and (iii) $Cov(\nabla^2 y_t, \nabla^2 y_{t-2}) = -\theta_2\sigma_a^2 = \sigma_\eta^2$. From (ii) and (iii) we deduce that $\theta_1 = -4\theta_2/(1-\theta_2)$, while (i) and (iii) lead to $-\theta_2^{-1}(1+\theta_1^2+\theta_2^2) = 1/\lambda + 6$, where $\lambda = \sigma_\eta^2/\sigma_\varepsilon^2$ is the signal-to-noise variance ratio. An explicit expression for θ_2 is given by (see Fernández-Macho, 2005, Kaiser and Maravall, 2005 and McElroy 2008)

$$\theta_2 = \{-4 - 1/\lambda - c + [24/\lambda + 2/\lambda^2 + (8 + 2/\lambda)c]^{1/2}\}/4 \quad (16)$$

with $c = (16/\lambda + 1/\lambda^2)^{1/2}$. Finally, from (i) we also get $\sigma_a^2 = (1+6\lambda)/(1+\theta_1^2+\theta_2^2)$. Then, it is clear that $\theta_2 \downarrow -1$, $\theta_1 \uparrow 2$ and $\sigma_a^2 \uparrow \infty$ as $\lambda \uparrow \infty$, implying that the trend gets smoother than the original series when this happens. On the other hand, if $\theta_2 \uparrow 1$, then $\lambda \downarrow 0$, in which case the trend gets closer to the original series and therefore its smoothness diminishes.

Point forecasts of the original series are obtained as indicated by (15) and interval forecasts, say of ±2 standard errors, can also be calculated around the point forecasts by means of the following expression

$$\hat{y}_t(h) \pm 2(\sum_{j=0}^{h-1}\psi_j^2)^{1/2}\sigma_a \quad \text{for } h = 1, \ldots, H, \tag{17}$$

where the ψ weights are $\psi_0 = 1$ and $\psi_j = (j-1)\theta_2 + j\theta_1 - j - 1$ for j = 1, 2, … . In Table 1 we present the values required to calculate interval forecasts for a time series of size N = 84 (the size of the series employed in the motivating empirical application to the percent growth rate of remittances to Mexico presented below) filtered with the HP filter. This table was obtained by first fixing the desired percentage of smoothness 100S%, then we searched numerically for the λ value that yields such index in expression $S = 1 - tr(I_N + \lambda K'K)^{-1}/N$ appearing in the paragraph before equation (11). The parameter θ_2 is calibrated by using formula (16) and the remaining model parameters are obtained from $\theta_1 = -4\theta_2/(1-\theta_2)$ and $\sigma_a^2 = (1+6\lambda)/(1+\theta_1^2+\theta_2^2)$.

Table 1: Values of the quantities required to produce forecasts with the HP filter for different percentages of smoothness and N = 84.

Percent smoothness (100S%)	Smoothing parameter (λ)	IMA(2,2) parameters (θ_2) and (θ_1)		Variance of IMA model (σ_a^2)	Two standard error for h = 1 ($2\sigma_a$)	Two standard error, h = 2
80%	14.012	-0.404	1.151	34.189	11.694	15.340
85%	45.828	-0.479	1.295	94.956	19.489	23.845
87.5%	99.746	-0.519	1.366	191.144	27.651	32.734
90%	266.250	-0.559	1.434	474.468	43.565	50.056
92.5%	998.493	-0.596	1.494	1670.108	81.734	91.599
93.21%	1600*	-0.606	1.509	2635.061	102.666	114.379
95%	7448.443	-0.626	1.541	11867.428	217.875	239.765

*Value usually employed with the HP filter, for quarterly series.

4. Empirical Application

The original dataset employed in this application is a quarterly series of workers' remittances in U.S. dollars recorded by BANXICO and expressed as relative growth rates by applying an annual difference to the logarithm of remittances. All numerical computations were carried out with MatLab version R2018a and WinRATS version 9.0 (www.estima.com).

4.1 Trend estimation

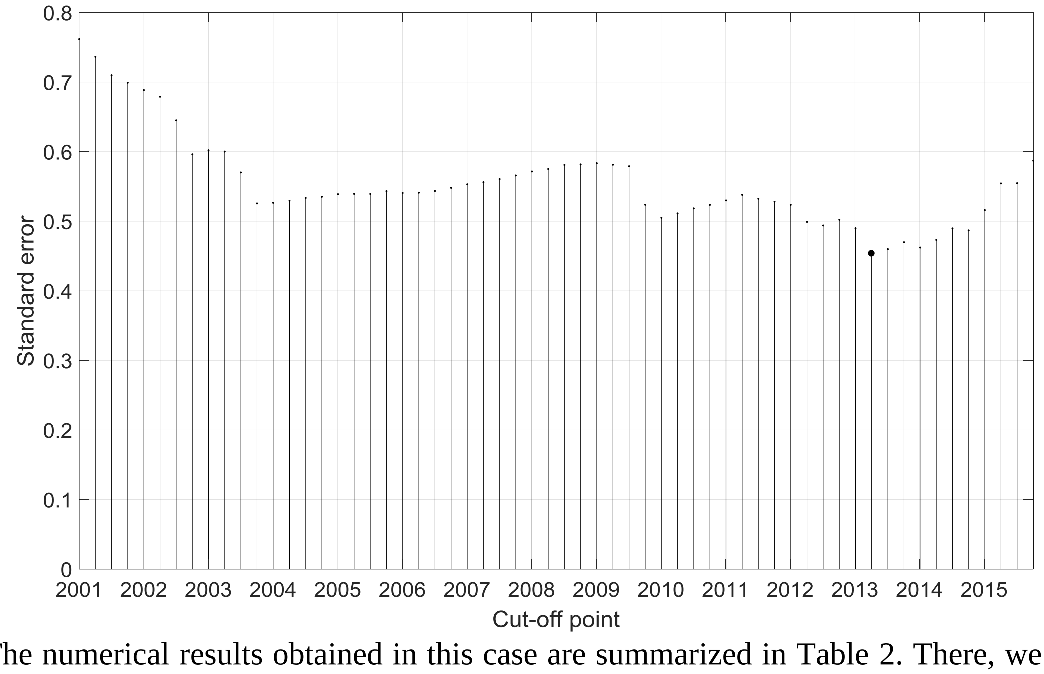
We first apply the HP filter to the percent growth rates that runs from 1996:I to 2016:IV. To that end we used the following strategy that comes out from a simulation study carried out by Guerrero, et al. (2017).

(i) When the series behaves as a straight line, a large percentage of smoothness should be chosen, starting from 90% for $N > 48$, and increasing the smoothness for larger values of N.
(ii) If the series does not show a straight line pattern, the percentage of smoothness should start from 85% for $N > 48$ and increase its value as N gets larger.

So, we decided to fix in advance the trend smoothness in 90% because the series has $N = 84$ data points and it does not resemble a straight line. Thus, we obtained the estimated trend shown in Figure 1. We notice that the usual value of the smoothing parameter employed with the HP filter, $\lambda = 1600$, would produce in this case an inappropriately large percentage of smoothness (93.206%), as indicated in Table 1. Then, we started the empirical search for a possible cutoff point in the series by assuming the existence of two regimes. For practical purposes, we considered that the numerical search converged if the imposed global percentage of smoothness is within the range $100(S \pm \varepsilon)\%$, with $\varepsilon = 0.0002$.

As shown in Figure 2, we found a cutoff point located at 2013:II ($N_1 = 70$), which is a sensible segmentation from a statistical standpoint since the standard deviation without segmentation is $\hat{\sigma}_0 = 0.5315$ and it diminishes to $\hat{\sigma}_0 = 0.4535$ by segmenting the series. This segmentation can be explained by a recovery of remittances flows to Mexico. In fact, the improvement in U.S. employment indicators at the end of 2013 and the beginning of 2014, specifically in the states where Mexican immigrants mostly reside, such as California and Texas, provides an explanation of the recent recovery of remittances to Mexico starting at the end of 2013. It should be noticed that the search of the cutoff point does not cover the end of the series, this is because the procedure employed here does not apply to the first 4 nor the last 4 data points of the series.

Figure 2: Estimated standard error for the trend estimate of the annual percent rate of growth of remittances to Mexico.



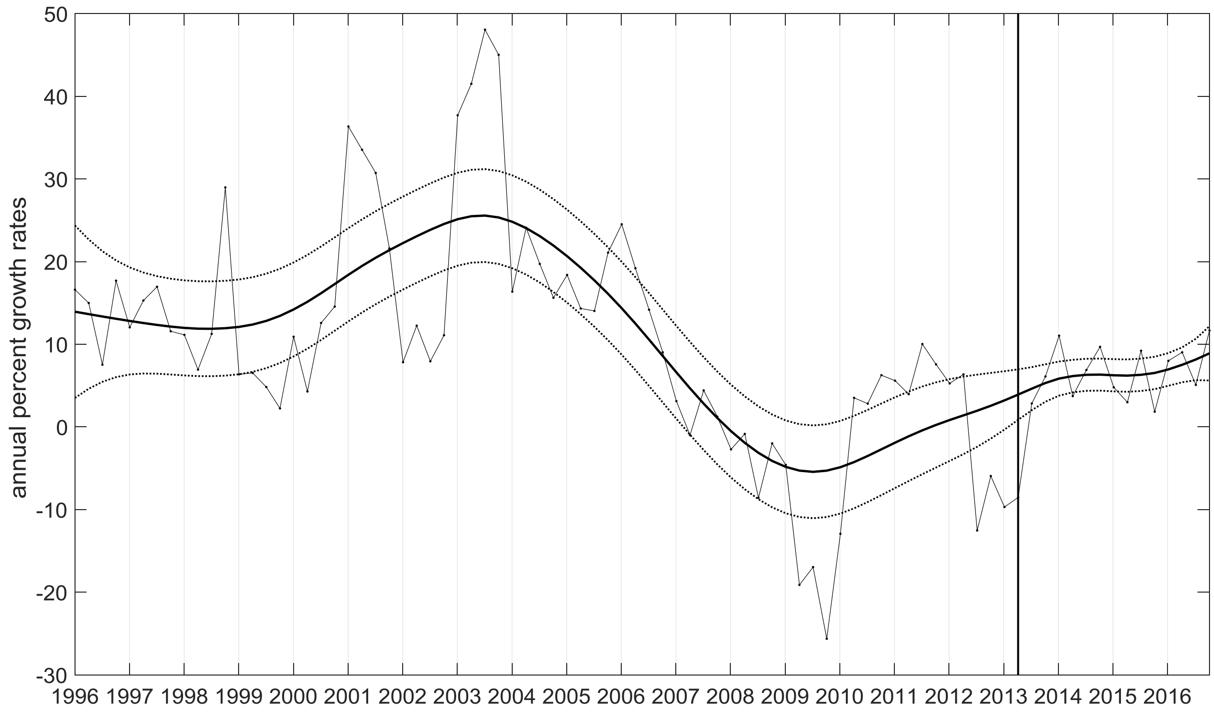
The numerical results obtained in this case are summarized in Table 2. There, we see that once we fix the percentage of smoothness for the first segment (92.5%) higher than that for the whole series (90%), much less smoothness is required (77.5%) for the second segment, implying a trend closer to the observed data at the end of the time series. The choice of such a percentage follows from noticing that there is more variability in the first segment (so that more smoothness is needed for that segment) than in the second one. Moreover, the reduction in standard deviation of the estimated trend in this example amounts to 100(1 - 0.4535/0.5315)% = 14.68% and this reduction in variability provides a justification for the extra effort of segmenting the data.

Table 2: Smoothness results for quarterly percent growth rates of remittances with the HP filter using one and two segments. Data-based segmentation at N_1 = 70 (2013:II).

	Global	First segment	Second segment
Percentage of smoothness	90%	92.5%	77.5%
Smoothing parameter	266.25	514.2	28.5
Estimated standard error	$\hat{\sigma}_0 = 0.5315$	$\hat{\sigma}_0 = 0.4535$	

Figure 3 shows the estimated trend by segments, with 90% global smoothness. There we see that the variability in the second (the most recent) segment of the data is much smaller than that in the first segment.

Figure 3: Quarterly series of annual percent growth rates of remittances to Mexico, 1996:I – 2016:IV. Trend with 90% global smoothness, 92.5% smoothness for the first segment and 77.5% for the second one.



4.2 Comparing forecasting performance

Remittances have had a significant positive effect on the Mexican economy and on household well-being for the families that receive them. The 2008 economic crisis, which severely affected the U.S. economy and hence some economic sectors that traditionally employ Mexican immigrants (such as construction and manufacturing industries) led to a fall of remittances to Mexico. Hence, policymakers are concerned about the effects of this reduction in money transfers at a time when local economic growth is also slowing down. So, they require forward-looking analyses of the sustainability of remittances, trying to foresee whether remittances flows would continue at their current growth, in the short and medium term. Therefore, the need of generating forecasts is clear in this context.

We have shown that the forecasts produced with the IMA (2,2) model, by means of equation (15), depend on the amount of smoothness for the last segment of the data and on the cutoff point employed. Since the HP filter employed mitigates the effect of outliers and transitory blips adequately, it tends to show an adequate trend behavior for the last segment of the dataset and there is no reason to doubt about the amount of smoothness or the cutoff point employed. However, from a forecasting perspective, a question of interest is whether the model underlying the filter outperforms a simpler alternative forecasting model, specifically a Random Walk (RW), in terms of Mean Squared Error (MSE),

For comparison, we generated h-period ahead forecasts of the annual percent growth rates of remittances flows to Mexico and calculated the mean square forecast error as

$$\sum_{t=N_1+3}^{N-h}(\hat{y}_t(h)-y_{t+h})^2/(N-N_1-2-h), \quad (18)$$

with N = 84 and N_1 = 70, with forecast horizons h = 1, …, 4. We employed three models related to the HP filter in this comparison; the first one is an IMA(2,2) with two segments and $\lambda = 28.5$, producing $100\,S(\lambda, N-N_1)\%$ = 77.5% smoothness for the most recent part of the series. The second one is an IMA(2,2) without segmentation and with $100\,S(\lambda, N)$ = 90.0% smoothness, so that $\lambda = 266.25$. The third one is the standard HP filter, that is, an IMA(2,2) model with $\lambda = 1600$ without segmentation, that produces $100\,S(\lambda, N)\%$ = 93.206% smoothness.

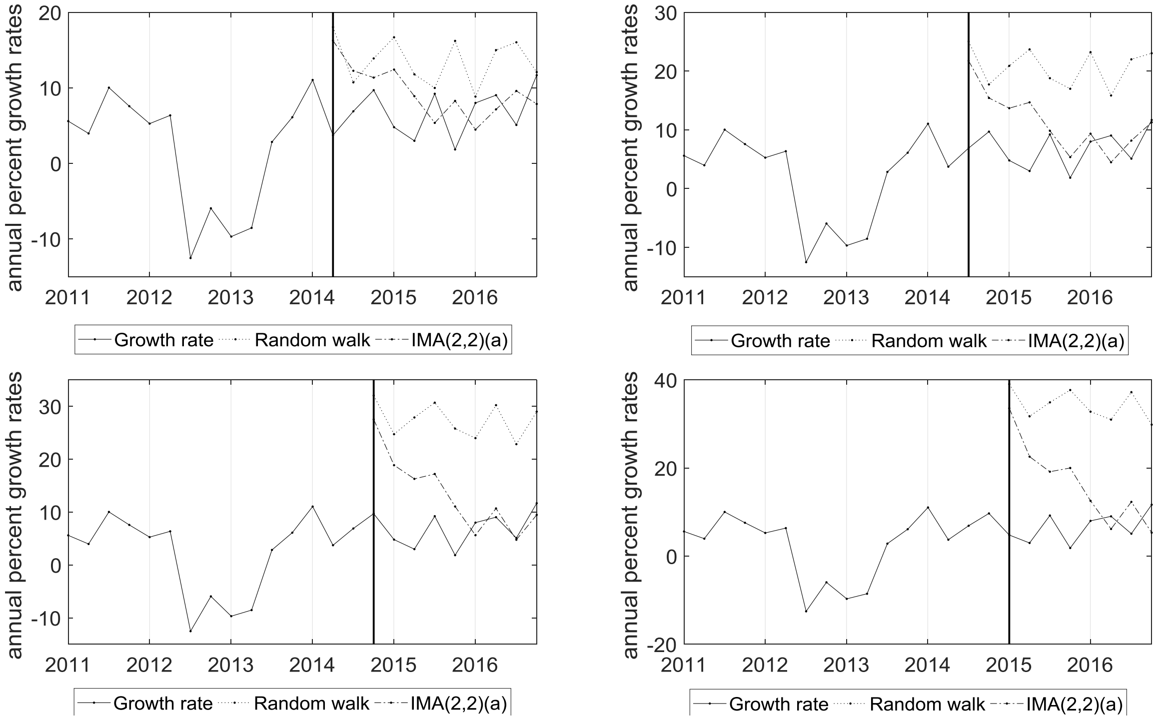
Table 3 presents the MSEs of the forecasts and compares them with those of a RW specification, whose forecasts are obtained as $\hat{y}_t(h) = y_t + h\bar{y}$, with $\bar{y} = \sum_{t=N_1+3}^{N} y_t/(N-N_1-3)$. In that table, we can see that the average improvement of forecast precision, over the four quarter-ahead horizon, is about 68.0%, 46.4% and 37.12% for the IMA(2,2)[a], IMA(2,2)[b] and IMA(2,2)[c] models, respectively. As a complement of Table 3, in Figure 4 we can visually appreciate the performance of the IMA(2,2)[a] and RW models.

Table 3: MSE of the forecasts at horizons from one to four quarters.

Mean Square Forecast Error				
	Forecast horizon (h)			
Model	1	2	3	4
RW	74.4849	209.4436	429.9089	657.8221
IMA(2,2)[a]; $\lambda = 28.5$	35.3118	51.3736	94.9000	220.1261
Improvement (%)	52.5	75.4	77.9	66.5
IMA(2,2)[b]; $\lambda = 266.25$	49.9123	86.0312	162.6456	323.0987
Improvement (%)	43.1	32.9	58.9	50.8
IMA(2,2)[c]; $\lambda = 1600$	64.2479	114.8590	212.7991	399.3122
Improvement (%)	13.7	45.1	50.5	39.2

Notes: MSEs are for forecasts with t = *2014:II+h* to *2016:IV,* where h is the forecast horizon.

Figure 4: Forecasts of the quarterly series of annual percent growth rates of remittances to Mexico with different horizons.



4.3 Forecast accuracy evaluation

As a complement of the forecasting performance analysis, we now focus on forecast accuracy. Let us recall that we did not estimate model parameters, so that each model employed here is only an approximation and cannot be regarded as a null hypothesis to be statistically tested. Furthermore, since we forecast the last segment of the publicly available data, we can use Diebold-Mariano (DM) tests to measure forecast accuracy as in a real-time comparison. Thus, in Table 4 we present some Diebold-Mariano (DM) test statistics (see Diebold and Mariano, 1995) for the null hypothesis of no-difference in accuracy of two competing forecasts, that is, each one of the three IMA(2,2) models against a RW. Since the standard DM test is known to over-reject the null hypothesis in the context of finite samples, we applied here the modified DM test proposed by Harvey et al. (1997). Each calculated statistic reported in Table 4 should be compared with a standard normal distribution to declare statistical significance. The DM test results reinforce our findings in Table 3, *i. e.* the three IMA(2,2) models are in general significantly better than the RW, in the context of forecasting. Nevertheless, the second IMA(2,2) model, without segmentation, shows no-significant difference (at the 5% level) in forecasting ability, as compared with the RW model when forecasting 4-periods ahead. The usual HP filter does not show significant difference in forecasting ability when compared with the RW in the forecast horizons 1 and 4.

Table 4: Diebold-Mariano tests for relative forecasting ability.

	Forecast horizon			
Models	1	2	3	4
IMA(2,2)[a] *vs.* RW				
MSE ratio	0.474	0.245	0.221	0.335
DM-statistic	2.465	10.462	7.909	1.965
p-value	0.007	0.000	0.000	0.025
IMA(2,2)[b] *vs.* RW				
MSE ratio	0.670	0.411	0.378	0.491
DM-statistic	1.716	6.269	5.263	1.287
p-value	0.043	0.000	0.000	0.099
IMA(2,2)[c] *vs.* RW				
MSE ratio	0.863	0.548	0.495	0.607
DM-statistic	0.724	3.754	3.888	0.988
p-value	0.235	0.000	0.000	0.162

5. Conclusion

We propose here to employ a filtering methodology to estimate the segmented trend of a time series that allows us to control the amount of smoothness in the trend by segments of the data range, corresponding to different regimes. The required smoothing parameter is implicitly chosen by fixing a desired amount of smoothness for the trend, which is decided by applying some simple data-based criteria. Besides, the cutoff point for segmenting the series is located by means of simple data-based procedure that is also in line with filtering the series by Penalized Least Squares. The filter employed makes use of the smoothing parameter previously found, which is related to the underlying IMA(2,2) model parameters. This model is then used to forecast the original series using the last segment of the data.

The empirical application to remittances flows to Mexico was carried out with three IMA(2,2) models, calibrated with different amounts of smoothness. The results obtained in this empirical example were clear in defining two different segments associated with more variability in the first segment of the time series. Therefore, more smoothness was needed for the first segment than for the second one, implying a trend closer to the observed data at the end of the time series. We believe this fact was helpful to improve the predictive ability of the entertained models.

Acknowledgements

The authors thank the Editor of this Journal and the anonymous referees for their comments and suggestions that helped to improve the presentation of this work. Eliud Silva dedicates this article to his wife. V. M. Guerrero and Alejandro Islas-Camargo also acknowledge the support provided by Asociación Mexicana de Cultura, A. C. to carry out this research.

References

Adams, R.H. and Page, J. (2005). Do international migration and remittances reduce poverty in developing countries? *World Development*, **33(10)**, 1645-1669.

Amuedo-Dorantes, C. and Pozo, S. (2011). New evidence on the role of remittances on health care expenditures by Mexican households. *Review of Economics of the Household,* **9(1)**, 69-98.

Box, G.E.P. and Tiao, G.C. (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association*, **70**, 70-79.

Ruppert, D., Wand, M.P. and Carroll, R.J. (2003). *Semiparametric Regression*. Cambridge University Press: New York.

Chandler, R. and Scott, M. (2011). *Statistical Methods for Trend Detection and Analysis in the Environmental Sciences*. John Wiley and Sons: UK.

Diebold, F.X. and Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, **13**, 253-265.

Eilers, P. and Marx, B. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science*, **11**, 89-121.

Fajnzylber, P. and Lopez, J.H. (2007). Close to home: The development impact of remittances in Latin America. *The World Bank*, Washington DC.

Fernández-Macho, J. (2005). Comments on 'Combining filter design with model-based filtering'. *International Journal of Forecasting*, **21**, 711-715.

Guerrero, V.M. (2007). Time series smoothing by penalized least squares. *Statistics and Probability Letters*, **77**, 1225-1234.

Guerrero, V.M. and Silva, E. (2015). Smoothing a time series by segments of the data range. *Communications in Statistics - Theory and Methods*, **44**, 4568-4585.

Guerrero, V.M., Islas-Camargo, A. and Ramirez-Ramirez, L. (2017). Trend estimation of multivariate time series with controlled smoothness. *Communications in Statistics - Theory and Methods*, **46**, 6704-6726.

Giuliano, P. and Ruiz-Arranz, M. (2009). Remittances, financial development and growth. *Journal of Development Economics*, **90(1)**, 144-152.

Gupta, S., Pattillo, A. and Wagh, S. (2009). Effects of remittances on poverty and financial development in Sub-Saharan. *World Development*, **37**, 104-115.

Harvey, D., Leybourne S. and Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, **13(2)**, 281-291.

Hildebrandt, N. and McKenzie, D. (2005). The effects of migration on child health in Mexico. *Economía*, **6(1)**, 257-289.

Kaiser, R. and Maravall, A. (2005). Combining filter design with model-based filtering (with an application to business-cycle estimation). *International Journal of Forecasting*, **21**, 691-710.

McElroy, T. (2008). Exact formulas for the Hodrick-Prescott filter. *Econometrics Journal*, **11**, 209-217.

Mohapatra, S. and Ratha, D. (2010). Forecasting migrant remittances during the global financial crisis. *Migration Letters*, **7(2)**, 203-213.

Nyamongo, E.M., Misati, R.N., Kipyegon, L. and Ndirangu, L. (2012). Remittances, financial development and economic growth in Africa. *Journal of Economic and Business*, **64**, 240-260.

Paige, R.L. and Trindade, A. (2010). The Hodrick-Prescott Filter: A special case of penalized spline smoothing. *Electronic Journal of Statistics*, **4**, 856-874.

Rao, B. and Hassan, G.M. (2011). Are the direct and indirect growth effects of remittances significant? *The World Economy*, **35(3)**, 351-371.

Valero-Gil, J.N. (2009). Remittances and the household's expenditures on health. *Journal of Business Strategies*, Spring issue.

Wahba, G. (1990). Spline models for observational data. *Society for Industrial and Applied Mathematics (SIAM)*, Philadelphia, PA.

World Bank (2011). *Migration and Remittances Factbook*, 2nd Edition. The World Bank Group.

World Bank (2016). *Migration and Remittances. Recent Developments and Outlook*. Migration and Development Brief 26. The World Bank Group.