Documento de Trabajo - 2019/03

Now-casting Spain

Manu García
(Universidad Complutense de Madrid and FEDEA)

Juan F. Rubio-Ramírez
(Emory University and FEDEA)

fedea

Now-casting Spain

Manu García
Universidad Complutense de Madrid
FEDEA

Juan F. Rubio-Ramírez *
Emory University
FEDEA

April 11, 2019

1 Introduction

In this paper, we follow Antolin-Diaz et al. (2017) and present a dynamic factor model (DFM) that allows for gradual changes in the mean and the variance of real output growth to track changes in the long-run growth rate of real GDP for Spain in a timely and reliable manner. In doing so we separate the long-run growth rate of real GDP from its cyclical counterpart. This model introduces two features into an otherwise standard DFM of real activity data. First, it allows the mean of real GDP growth, and possibly other series, to drift gradually over time. Second, it allows for stochastic volatility (SV) in the innovations to both factors and idiosyncratic components. Given our interest in studying the entire period since early 70's, the inclusion of SV is essential to capture the substantial changes in the volatility of output that have taken place in this sample.

Since the seminal contributions of Giannone et al. (2008) DFMs have become the standard tool for now-casting. Interestingly, Antolin-Diaz et al. (2017) showed that standard DFM forecasts revert very quickly to the unconditional mean of GDP, so taking into account the variation in long-run GDP growth substantially improves point and density GDP forecasts even at very short horizons.

The model used here basically borrows from Antolin-Diaz et al. (2017) and it is related to two strands of literature. The first one encompasses papers that allow for structural changes within the DFM framework. Del Negro and Otrok (2008) model time variation in factor loadings and

*Corresponding author: Juan Rubio-Ramírez<juan.rubio-ramirez@emory.edu>

volatilities, while Marcellino et al. (2016) show that the addition of SV improves the performance of the model for short-term forecasting of euro-area GDP. Acknowledging the importance of allowing for time variation in the means of the variables, Stock and Watson (2012) pre-filter their data set in order to remove any low-frequency trends from the resulting growth rates using a bi-weight local mean. In his comment to their paper, Sims (2012) suggests explicitly modelling, rather than filtering out, these long-run trends, and emphasizes the importance of evolving volatilities for describing and understanding macroeconomic data. The second strand of related literature takes a similar approach to decomposing long-run GDP growth into its drivers, in particular Gordon (2010, 2014) and Reifschneider et al. (2013). Relative to these studies, we emphasize the importance of using a broader information set, as well as a Bayesian approach, which allows using priors to inform the estimate of long-run growth and characterize the uncertainty around the estimate stemming from both filtering and parameter uncertainty. The remainder of this paper is organized as follows. Section 2 presents the model while Section 3 presents the data used in the model.

2 The Model

Let $\mathbf{y}_t$ be an $n \times 1$ vector of observable macroeconomic time series, and let $\mathbf{f}_t$ denote a $k \times 1$ vector of latent common factors. It is assumed that $n >> k$, i.e. the number of observables is much larger than the number of factors. Formally,

$$\mathbf{y}_t = \mathbf{c}_t + \boldsymbol{\Lambda}(L) f_t + \mathbf{u}_t, \tag{1}$$

where $\boldsymbol{\Lambda}(L)$ is a polynomial in the lag operator of order s which contains the loadings on the contemporaneous common factor and its lags. $\mathbf{u}_t$ is a vector of idiosyncratic components. Shifts in the long-run mean of $\mathbf{y}_t$ are captured by time-variation in $\mathbf{c}_t$. In principle one could allow time-varying intercepts in all or a subset of the variables in the system. Moreover, time variation in a given series could be shared by other series. $\mathbf{c}_t$ is therefore flexibly specified as

$$\mathbf{c}_t = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{c} \end{bmatrix} \begin{bmatrix} \mathbf{a}_t \\ 1 \end{bmatrix}, \tag{2}$$

where $\mathbf{a}_t$ is an $r \times 1$ vector of time-varying means, $\mathbf{B}$ is an $m \times r$ matrix which governs how the time-variation affects the corresponding observables, and $\mathbf{c}$ is an $(n - m) \times 1$ vector of constants. In our baseline specification, $\mathbf{a}_t$ will be a scalar capturing time-variation in long-run real GDP growth, which is shared by real consumption growth, so that $r = 1, m = 2$.

Throughout the paper, we focus on the case of a single dynamic factor by setting $k = 1$ (i.e. $\mathbf{f}_t = f_t$).[1] The laws of motion of the latent factor and the idiosyncratic components are

$$(1 - \phi(L))f_t = \sigma_{\varepsilon_t}\varepsilon_t, \tag{3}$$

$$(1 - \rho_i(L))u_{i,t} = \sigma_{\eta_{i,t}}\eta_{i,t}, \quad i = 1, \ldots, n \tag{4}$$

where $\phi(L)$ and $\rho_i(L)$ denote polynomials in the lag operator of order p and q, respectively. The idiosyncratic components are cross-sectionally orthogonal and are assumed to be uncorrelated with the common factor at all leads and lags, i.e. $\varepsilon_t \overset{iid}{\sim} N(0, 1)$ and $\eta_{i,t} \overset{iid}{\sim} N(0, 1)$.

Finally, the dynamics of the model's time-varying parameters are specified to follow driftless random walks:

$$a_{j,t} = a_{j,t-1} + v_{a_{j,t}}, \qquad v_{a_{j,t}} \overset{iid}{\sim} N(0, \omega^2_{a,j}) \quad j = 1, \ldots, r \tag{5}$$

$$\log \sigma_{\varepsilon_t} = \log \sigma_{\varepsilon_{t-1}} + v_{\varepsilon,t}, \qquad v_{\varepsilon,t} \overset{iid}{\sim} N(0, \omega^2_{\varepsilon}) \tag{6}$$

$$\log \sigma_{\eta_{i,t}} = \log \sigma_{\eta_{i,t-1}} + v_{\eta_{i,t}}, \qquad v_{\eta_{i,t}} \overset{iid}{\sim} N(0, \omega^2_{\eta,i}) \quad i = 1, \ldots, n \tag{7}$$

where $a_{j,t}$ are the r time-varying elements in $\mathbf{a}_t$, and σ_{ε_t} and $\sigma_{\eta_{i,t}}$ capture the SV of the innovations to factor and idiosyncratic components.

Using the data to be described in the next section, we estimate the model following the steps described in section 3.4 of Antolin-Diaz et al. (2017).

[1] Note that we order real GDP growth as the first element of $\mathbf{y}_t$, and normalize the loading for GDP to unity. This serves as an identifying restriction in our estimation algorithm. Bai and Wang (2015) discuss minimal identifying assumptions for DFMs.

3 The database

Our data set includes three key business cycle variables measured at quarterly frequency (output, employed and unemployed), as well as a set of 15 monthly indicators that are intended to provide additional information about cyclical developments in a timely manner.

The additional monthly indicators are crucial to our objective of disentangling in real time the cyclical and long-run components of GDP growth, since the quarterly variables are only available with substantial delay. In principle, a large number of candidate series are available to inform the estimate of f_t. In practice, however, macroeconomic data series are typically clustered in a small number of broad categories (such as production, employment, or income) for which desegregated series are available along various dimensions (such as economic sectors, demographic characteristics, or expenditure categories). The choice of which available series to include for estimation can therefore be broken into, first, a choice of which broad categories to include, and second, to which level and along which dimensions of desegregation.

With regard to which broad categories of data to include, previous studies agree that prices and monetary and financial indicators are uninformative for the purpose of tracking real GDP, and argue for extracting a single common factor that captures real economic activity. As for the possible inclusion of series within each category, Boivin and Ng (2006) argue that the presence of strong correlation in the idiosyncratic components of disaggregated series of the same category will be a source of misspecification that can worsen the performance of the model in terms of in-sample fit and out-of-sample forecasting of key series. Alvarez et al. (2012) investigate the trade-off between DFMs with very few indicators, where the good large-sample properties of factor models are unlikely to hold, and those with a very large amount of indicators, where the problems above are likely to arise. They conclude that using a medium-sized panel with representative indicators of each category yields the best forecasting results.

The above considerations lead us to select 15 monthly indicators that include the high-level aggregates for all of the available broad categories that capture real activity, without over-weighting any particular category. The data used in this project is described in Tables 1 and 2. While Table 1 describes the data and the frequency, Table 2 describes the sources. As the tables shows, we include

Table 1: DATA SERIES USED IN EMPIRICAL ANALYSIS

	Type	Start Date	Transform.	Lag
QUARTERLY TIME SERIES				
Real GDP	Expenditure & Inc.	Q1:1970	% QoQ Ann	26
Unemployed	Labor Market	Q1:2002	% QoQ Ann	26
Employed	Labor Market	Q1:2000	% QoQ Ann	26
MONTHLY INDICATORS				
Real Imports of Goods	Foreign Trade	Jan 85	% MoM	35
Real Exports of Goods	Foreign Trade	Jan 85	% MoM	35
Consumption of Cement	Housing & Inc.	Jan 70	% MoM	27
Electric Energy Consumption	Production & Sales	Jan 81	% MoM	15
Industrial Production	Production & Sales	Jan 92	% MoM	15
Passenger Car Registrations	Production & Sales	Jan 60	% MoM	1
Registered Unemployment	Labor Market	Jan 96	% MoM	5
Retail Trade:Gen Def Index w/out Serv Stations	Production & Sales	Jan 95	% MoM	25
Social Security Afiliation	Labor Market	Jan 82	% MoM	5
Air Traffic, Passengers	Production & Sales	Jan 69	% MoM	25
Markit Manufacturing PMI	Business Confidence	May 07	-	-7
ISM Manufacturing PMI	Business Confidence	Jan 48	-	1
Consumer Confidence Indicator	Consumer Confid.	Jun 86	Diff 12 M.	-5
Industry: Production Trend	Industry Confidence	Jan 93	-	-3
Order Book Trend	Business Confidence	Jan 93	-	-15

Notes: % QoQ Ann refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. The last column shows the average publication lag, i.e. the number of days elapsed from the end of the period that the data point refers to until its publication by the statistical agency.

representative series of expenditure and income, the labor market, production and sales, foreign trade, housing, and business and consumer confidence.

The inclusion of all the available monthly surveys is particularly important. Apart from being the most timely series available, these are unlikely to feature permanent shifts in their mean by construction and have a high signal-to-noise ratio. They thus provide a clean signal to separate the cyclical component of GDP growth from its long-run counterpart.

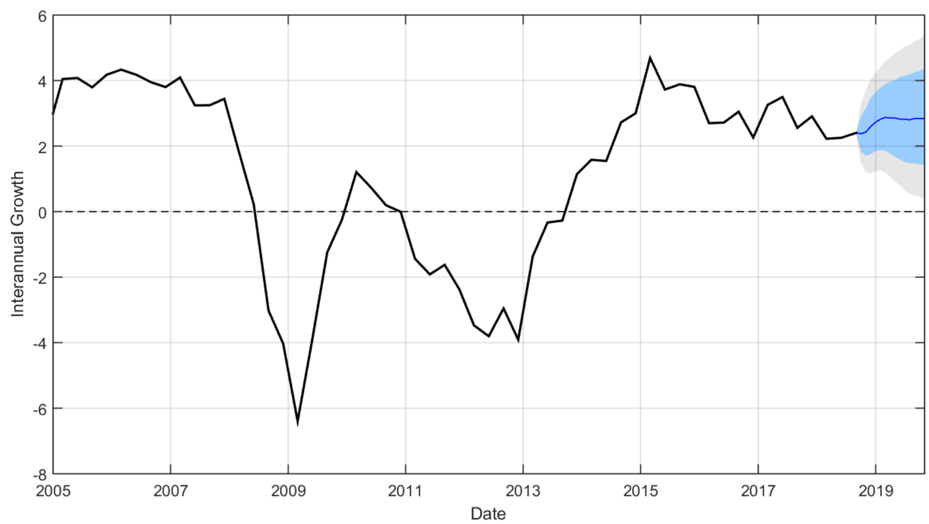
Table 2: Data series used in empirical analysis. Sources

	Source
Quarterly time series	
Real GDP	INE (Contabilidad Nacional Trimestral. Base 2010)
Unemployed	INE (Encuesta de Población Activa)
Employed	INE (Encuesta de Población Activa
Monthly indicators	
Real Imports of Goods	Ministerio de Economía y Empresa
Real Exports of Goods	Ministerio de Economía y Empresa
Consumption of Cement	OFICEMEN
Electric Energy Consumption	REE (Red Eléctrica Española)
Industrial Production	INE
Passenger Car Registrations	DGT (Dirección General de Tráfico)
Registered Unemployment	SEPE (Servicio Público de Empleo Estatal
Retail Trade: Gen Def Index w/out Serv Stations	INE
Social Security Afiliation	Ministerio de Trabajo, Migraciones y Seguridad Social
Air Traffic, Passengers	AENA (Aeropuertos Españoles y Navegación Aérea)
Markit Manufacturing	Markit Economics
Markit Manufacturing	Markit Economics
Consumer Confidence Indicator	European Comission
Industry: Production Trend	Ministerio de Industria, Turismo y Comercio
Order Book Trend	Ministerio de Industria, Turismo y Comercio

4 Results and updates

The now-casting model used in this work will be updated in a daily basis when new data updates became available. Data is updated with a certain time lag, in the last column of Table 1 can be founded the approximate lag of publication of each series. With any new data update, the model will be updated in the next 24 hours in order to show the GDP now-cast with all the available information displayed in every moment. As we can see in Figure 1 this will allow us to obtain GDP forecasts for the last, actual and next quarters.

Figure 1: GDP Now-casting for the Spanish Economy



5 Conclusion

We propose a Bayesian DFM, which incorporates low-frequency variation in the mean and variance, heterogeneous responses to common shocks, outlier observations and fat tails, endogenous modeling of seasonality. With this model, as shown in Antolin-Diaz et al. (2017), the real-time now-casting performance is substantially improved across a variety of metrics: Capturing trends and SV improves now-casting performance significantly, heterogeneous dynamics deliver substantial additional improvement, fat tails improve density forecasts of monthly variables.

References

Alvarez, R., M. Camacho, and G. Perez-Quiros (2012): "Finite sample performance of small versus large scale dynamic factor models," CEPR Discussion Papers 8867, C.E.P.R. Discussion Papers.

Antolin-Diaz, J., T. Drechsel, and I. Petrella (2017): "Tracking the Slowdown in Long-Run GDP Growth," *The Review of Economics and Statistics*, 99, 343–356.

Bai, J. and P. Wang (2015): "Identification and Bayesian Estimation of Dynamic Factor Models," *Journal of Business & Economic Statistics*, 33, 221–240.

Boivin, J. and S. Ng (2006): "Are more data always better for factor analysis?" *Journal of Econometrics*, 132, 169–194.

Del Negro, M. and C. Otrok (2008): "Dynamic factor models with time-varying parameters:measuring changes in international business cycles," NYFed Staff Rep. 326.

Giannone, D., L. Reichlin, and D. Small (2008): "Nowcasting: The real-time informational content of macroeconomic data," *Journal of Monetary Economics*, 55, 665–676.

Gordon, R. J. (2010): "Okun's Law and Productivity Innovations," *American Economic Review*, 100, 11–15.

——— (2014): "A New Method of Estimating Potential Real GDP Growth: Implications for the Labor Market and the Debt/GDP Ratio," NBER Working Paper 20423.

Marcellino, M., M. Porqueddu, and F. Venditti (2016): "Short-Term GDP Forecasting With a Mixed-Frequency Dynamic Factor Model With Stochastic Volatility," *Journal of Business & Economic Statistics*, 34, 118–127.

Reifschneider, D., W. Wascher, and D. Wilcox (2013): "Aggregate supply in the United States: recent developments and implications for the conduct of monetary policy," Finance and Economics Discussion Series 2013-77, Board of Governors of the Fed.

Sims, C. A. (2012): "Comment to Stock and Watson (2012)," *Brookings Papers on Economic Activity*, Spring, 81–156.

Stock, J. and M. Watson (2012): "Disentangling the Channels of the 2007-09 Recession," *Brookings Papers on Economic Activity*, Spring, 81–156.