Estimation of VIX futures through Gaussian factor models

Felipe do Nascimento Fernandes[†]
Fernando Antonio Lucena Aiube[‡]
Carla Gomes Costa de Souza[*]

Abstract In this paper we investigate VIX dynamics through a two-factor Gaussian model, following Avellaneda and Papanicolaou (2019). Two strategies were adopted. First, we considered constant market price of risk. Second, we included time-varying market price of risk. In both cases, we estimated the model following two different approaches. In the first, we used a sample encompassing seven VIX future contracts and also the spot VIX. In the second, we dropped the spot VIX from the sample. We found that this removal provided a better adjustment between empirical and estimated data.

Keywords: VIX futures contracts; Gaussian factor models; Kalman filter; Market price of risk.

JEL Code: G13, C32.

1. Introduction

Since their emergence in 2004, VIX future contracts have become an important instrument for investment decisions, hedging strategies and risk management. As typically occurs in commodities markets, VIX future contracts are negotiated in futures markets. Moreover, there are VIX traded options, and more recently, instruments like exchanged traded products (ETPs), such as exchange traded notes (ETNs) and exchanged traded funds (ETFs). Hence, modeling the dynamics of these instruments is of crucial importance for traders to develop successful strategies involving that index. The need for an accurate description of the index has aroused the interest of the academics. Several researchers have demonstrated their interest in the subject for better understanding the behavior of the VIX and its derivatives.

Gregory and Timcenko (2016) developed an initial study of VIX futures contracts. Hill et al. (2015) and Alexander and Korovilas (2013) described strategies and historical performance of ETPs. For details on the operational dynamics of ETPs, see Boroujerdi and Fogertey (2015).

Submitted on November 14, 2021. Revised on March 29 and August 8, 2022. Accepted on August 8, 2022. Published online in September 2022. Editor in charge: Rodrigo Targino.
[†]Instituto de Matemática Pura e Aplicada, Brazil: nasci.lipe@gmail.com
[‡]Universidade do Estado do Rio de Janeiro, Brazil: faiube@gmail.com
[*]Universidade do Estado do Rio de Janeiro, Brazil: gccarla2012@gmail.com

©2022 Sociedade Brasileira de Finanças, under a Creative Commons Attribution 3.0 license

Zhang and Zhu (2006) used the stochastic volatility to model spot VIX prices following the Heston (1993) approach. They developed an expression for the VIX future prices. The discrepancy with the observed data was between 2 to 12% when the estimation was carried out based on the last year of the sample. Using the whole sample the authors reported a difference of 12 to 44%.

Wang et al. (2017) proposed a closed-form pricing formula for VIX futures based on discrete-time Heston-Nandi GARCH model. The estimation is based on data from different sources: S&P 500 returns, VIX spot and VIX future prices. They found that the joint information based on VIX and VIX futures provided the best results.

Avellaneda and Papanicolaou (2019) developed a two-factor Gaussian model to describe VIX futures contracts dynamics. First, the authors constructed the constant maturity futures (CMFs) based on VIX futures contracts. After conducting unit root tests on VIX futures, they found evidence of a stationary process. The main hypothesis in the study was that the CMFs are stationary for each maturity contract. Hence, they modeled the spot VIX by decomposing it as the sum of two latent factors, both following a mean-reverting process. To establish the number of factors, the authors performed principal components analysis (PCA). They found that two factors accounted for the 90% of the variance of the term-structure shape. To estimate the model they used Kalman filtering and analyzed the behavior of rolling strategies involving ETNs and ETFs.

In this study, we use the original two-factor model of Avellaneda and Papanicolaou (2019) to model VIX prices. However, our analysis is based on two approaches. The first one considers the model with the market price of risk (MPR) held constant. In the second approach, we include a time-varying market price of risk (TVMPR). In both cases, we perform the estimation considering the following: (i) VIX futures contracts (future prices) and also the spot VIX (spot price); and (ii) only VIX futures contracts (future prices). The intuition behind this procedure is the same as in commodity models where the spot price is an unobservable factor and it is estimated as being an output of the model.

Unlike many studies in the literature of VIX prices that use stochastic volatility models, this paper is in line with that of Avellaneda and Papanicolaou (2019) focusing on Gaussian factor models. It contributes on VIX estimation procedures investigating the behavior of a more parsimonious model considering two different approaches when compared to the Avellaneda and Papanicolaou (2019) study: (i) excluding the VIX spot price from the sample, that is, considering only VIX future prices series; (ii) considering a constant

market price of risk in the model. We compare these approaches to their counterparts. We found evidence that removing the spot VIX results in a better adjustment to empirical data, mainly for medium and long term maturities. Moreover, the best adjustment when one removes the spot VIX from the sample, occurs with the time-varying market price of risk model. The usefulness of these results is a simpler procedure when dealing with VIX estimation.

The remainder of this paper is organized as follows. Section 2 specifies the model; Section 3 presents the data; Section 4 presents the results; Section 5 illustrates an application to an ETN; and Section 6 concludes.

2. The models

Consider $(\Omega, \mathscr{F}, \mathbb{P})$ the probability space. Furthermore, consider that the market is complete and arbitrage free, so that the uniqueness of an equivalent martingale measure (EMM) $\mathbb{Q}$ is well-established by the fundamental theorems of Finance.

Avellaneda and Papanicolaou (2019) run the PCA analysis and found that the first component accounts for 72% of the variance of the term-structure shape. Including the second component increases the explanation to 90%. The third and fourth components explain 96% and 97% of the variance of the term-structure, respectively. They considered a parsimonious model with two components to describe the behavior of VIX future contracts. We followed their study and decomposed the logarithm of the spot VIX as the sum of two latent components, $X_{1,t}$ and $X_{2,t}$. These two factors are described through an Ornstein-Uhlenbeck process, following the assumption that the time series of CMFs are stationary. This is coherent with the well-known fact, based on empirical analysis, that the variance of financial time series is a mean-reverting process.

In the $\mathbb{P}$-measure, the logarithm of the spot VIX is given by:

$$\ln \mathrm{VIX}_t = X_{1,t} + X_{2,t} \ , \tag{1a}$$

$$dX_{1,t} = \kappa_1(\mu_1 - X_{1,t})dt + \sigma_1 dB_{1,t} \ , \tag{1b}$$

$$dX_{2,t} = \kappa_2(\mu_2 - X_{2,t})dt + \sigma_2 dB_{2,t} \ , \tag{1c}$$

where the factors are driven by standard Brownian motions B_1 and B_2 respectively, which are correlated with $\rho dt = dB_{1,t}dB_{2,t}$. In addition, $\kappa_1, \kappa_2 > 0$ are the speeds of mean reversion and $\sigma_1, \sigma_2 > 0$ are the volatilities of each factor.

Define $\lambda_{1,t}$ and $\lambda_{2,t}$ as the time-varying risk premiums of both latent fac-

tors, respectively

$$\lambda_{1,t} = p_1 + q_1 X_{1,t} \ , \tag{2a}$$
$$\lambda_{2,t} = p_2 + q_2 X_{2,t} \ , \tag{2b}$$

where p_1, p_2, q_1 and $q_2 \in \mathbb{R}$.

Changing to the EMM $\mathbb{Q}$, we obtain:

$$dX_{1,t} = \kappa_1^\star(\mu_1^\star - X_{1,t})dt + \sigma_1 d\tilde{B}_{1,t} \ , \tag{3a}$$
$$dX_{2,t} = \kappa_2^\star(\mu_2^\star - X_{2,t})dt + \sigma_2 d\tilde{B}_{2,t} \ . \tag{3b}$$

The change of measure preserves the mean-reverting dynamics of both factors in the $\mathbb{Q}$-measure, where the new long-run means are $\mu_1^\star = \frac{\kappa_1\mu_1 - \sigma_1 p_1}{\kappa_1 + \sigma_1 q_1}$ and $\mu_2^\star = \frac{\kappa_2\mu_2 - \sigma_2 p_2}{\kappa_2 + \sigma_2 q_2}$. In the same manner, the new speeds of mean reversion are given by $\kappa_1^* = \kappa_1 + \sigma_1 q_1$ and $\kappa_2^* = \kappa_2 + \sigma_2 q_2$.

Let F_{t,T_j} denote the jth VIX future price observed at time t maturing at time T_j. We represent by V_{t,T_j} the CMF contract constructed for the jth contract based on VIX future prices of adjacent maturities F_{t,T_j} and $F_{t,T_{j+1}}$. Using the fact that $V_{t,T_j} = E^{\mathbb{Q}}(V_{T_j}|\mathscr{F}_t)$, one can prove that the logarithm of V_{t,T_j} is given by:

$$\ln V_{t,T_j} = X_{1,t}\, e^{-\kappa_1^\star(T_j - t)} + X_{2,t}\, e^{-\kappa_2^\star(T_j - t)} + A(\tau) \ , \tag{4}$$

where

$$A(\tau) = \mu_1^\star(1 - e^{-\kappa_1^\star\tau}) + \mu_2^\star(1 - e^{-\kappa_2^\star\tau}) + \frac{\sigma_1^2}{4\kappa_1^\star}(1 - e^{-2\kappa_1^\star\tau}) + \frac{\sigma_2^2}{4\kappa_2^\star}(1 - e^{-2\kappa_2^\star\tau}) + \frac{\rho\sigma_1\sigma_2}{\kappa_1^\star + \kappa_2^\star}(1 - e^{-(\kappa_1^\star + \kappa_2^\star)\tau}) \ ,$$

which is a function of the hyperparameters and the remaining time to maturity for each contract.

3. The data

The analyzed sample consists of seven VIX futures contracts and the spot VIX, from 21st April, 2008 to 28th January, 2019, forming a complete panel. Each series contains 510 weekly observations. We constructed the CMFs based on these VIX futures as in Avellaneda and Papanicolaou (2019). This is the same concept used for Treasuries. The details can be found in Alexander and Korovilas (2013). In the Appendix A we present the computation of the

Table 1
Main statistics of the CMF and the spot VIX series

	Spot VIX	V_{t,T_1}	V_{t,T_2}	V_{t,T_3}	V_{t,T_4}	V_{t,T_5}	V_{t,T_6}	V_{t,T_7}
Minimum (%)	9.34	9.88	11.33	12.25	13.23	13.78	14.23	14.83
Maximum (%)	80.06	67.95	56.29	52.72	48.63	46.44	44.49	44.06
Mean (%)	19.85	20.18	20.96	21.5	21.87	22.19	22.48	22.76
Std. Dev.(%)	10.03	8.96	7.96	7.28	6.78	6.43	6.15	5.96
Skewness	2.37	2.19	1.86	1.63	1.43	1.27	1.16	1.10
Kurtosis	6.97	5.61	3.97	3.04	2.06	1.33	0.90	0.74
Jarque-Bera	1427.3	1175.4	2012.7	1777.4	370.7	186.9	147.0	400.3

Note: The Jarque-Bera statistic was computed using the log-returns series.

CMF. The maturities obtained for V_{t,T_j} span from one through seven months ahead. Table 1 shows the main statistics of the CMF and the spot VIX series.

Observe that the skewness is positive in all series, meaning that the distribution is asymmetric to the right. Moreover, the standard deviation is higher for V_{t,T_j} with short maturities than longer ones. The Jarque-Bera test, conducted on the log-returns $\ln\left(V_{t,T_j}\right) - \ln\left(V_{t-1,T_j}\right)$, $j = 1 \ldots 7$, rejects the null hypothesis of normality of all contracts at 1% significance level. Additionally, for the spot VIX and the first three contracts the kurtosis indicates the presence of fat tails. It is important to note that, although the test statistics of the time series reject the Gaussian property, it is a well-established procedure in the literature to approximate future contracts through Gaussian models. Regarding this point we cite Schwartz and Smith (2000), Geman and Nguyen (2005), Cortazar and Naranjo (2006), to mention a few.

Since $\ln V_{t,T_j}$, in equation (4), is an affine function of the latent factors $X_{1,t}$ and $X_{2,t}$ and both are Gaussian, we applied the Kalman filter method as the estimation procedure. This is the common method used in the literature on affine Gaussian factor models. To proceed with the estimation, the first step is to rewrite the model in the state-space form. The Kalman filter has many applications in different areas of knowledge. In econometrics, we can mention Harvey (1989), Durbin and Koopman (2012) and Shumway and Stoffer (2006). In Appendix B we rewrite the model in state-space form, detail the Kalman filter algorithm and discuss the estimation issues.

4. Results

Following the equation (4), one can specify the difference between the logarithm of the observed (V_{t,T_j}) and estimated ($\widehat{V}_{t,T_j}$) CMFs of the jth future contract is given by $\ln V_{t,T_j} - \ln \widehat{V}_{t,T_j} = \varepsilon_{t_j}$ and $\varepsilon_{t_j} \sim N(0, s^2_{\varepsilon_j})$. The vector

$\Theta = (\kappa_1, \kappa_2, \sigma_1, \sigma_2, p_1, p_2, q_1, q_2, \rho, \mu_1, \mu_2, s^2_{\varepsilon_j})$, where $j = 1, \ldots, 7$, contains the hyperparameters of the model; and when the spot VIX time series is included in the sample, there is also $s^2_{\varepsilon_{spot}}$ in the vector Θ.

It is worth mentioning that once the measure is changed, the resulting model includes both types of parameters: the real measure parameters and the parameters involved in the specification of the market price of risk. They are jointly estimated and are all included in the Θ vector above.

As mentioned above, we performed two different analysis. First, we considered the MPR constant which means that $\lambda_1 = p_1$ and $\lambda_2 = p_2$ in equations (2a) and (2b). Second, we considered the TVMPR, which means that $q_1 \neq 0$ and $q_2 \neq 0$ in these equations.

Table 2 presents the estimation results for a constant MPR, for two different models. The Model 1 corresponds to the estimation procedure that includes the spot VIX and seven monthly CMF contracts while Model 2 consists of an application of Model 1 to the same sample as before but excluding the spot VIX series from the analysis.

For both models, the mean reversion is significant at 1% level. The same can be said about volatilities (σ_1 and σ_2) and the MPRs. For Model 2, the long-run growth rates for both latent factors are significant at 1%, whereas for Model 1 only μ_1 is significant. All CMF contract variances ($s^2_{\varepsilon_j}$), in both models, are significant at 1% level.

Tables 3 and 4 present the errors between observed and estimated prices, for constant MPR Model 1 and Model 2, respectively. As can be seen, removing the spot VIX from the sample improves the model fit to observed data, on average. This improvement occurs for almost all CMF contracts, except for V_{t,T_1}, V_{t,T_2} and marginally for V_{t,T_5}. The reason for this effect can be explained due the presence of spot VIX series in Model 1. The spot VIX contract affects the whole term-structure error absorbing mainly the existing errors of the next two contracts, V_{t,T_1} and V_{t,T_2}. This lets V_{t,T_1} and V_{t,T_2} with smaller errors compared to the same contracts in Model 2. This effect dissipates throughout longer maturities. In terms of information criteria, Model 2 presents a smaller likelihood and greater AIC in comparison with Model 1.

Table 5 presents the estimation results for Models 1 and 2 considering the TVMPR approach. As can be seen, the speed of mean reversion parameters are significant at 1% level for both latent factors, in both models. The volatilities follow the same pattern, as does the long-run growth rate, except for μ_2 in the TVMPR Model 1. Both correlation factors are significant at 1%. Additionally, all CMF contract variances are significant at the same level.

In Model 1, the TVMPR parameters are significant for only one of the model's factors, namely the second one. For this factor, both the variable

Table 2
Estimation results for constant MPR approach

	Model 1		Model 2	
Parameters	Value	t-stat	Value	t-stat
κ_1	3.0942	41.57	2.6161	31.3
κ_2	0.3124	24.38	0.2347	11.9
σ_1	0.7138	24.29	0.6232	23.3
σ_2	0.8520	23.95	1.1163	14.4
p_1	-1.4300	-4.35	-1.6803	-5.3
p_2	-1.2154	-6.43	0.3205	3.1
ρ	0.2159	8.87	0.1169	6.6
μ_1	1.4179	16.23	1.0775	12.2
μ_2	0.2082	0.61	1.1708	2.8
$s^2_{\varepsilon_{spot}}$	7.32×10^{-03}	582.10		
$s^2_{\varepsilon_1}$	2.10×10^{-03}	492.46	3.31×10^{-03}	355.3
$s^2_{\varepsilon_2}$	1.60×10^{-04}	135.09	3.03×10^{-04}	132.7
$s^2_{\varepsilon_3}$	2.29×10^{-04}	475.13	1.60×10^{-04}	237.9
$s^2_{\varepsilon_4}$	1.64×10^{-04}	454.75	1.31×10^{-04}	356.6
$s^2_{\varepsilon_5}$	1.03×10^{-04}	373.54	1.07×10^{-04}	471.7
$s^2_{\varepsilon_6}$	6.00×10^{-05}	170.72	4.72×10^{-05}	178.8
$s^2_{\varepsilon_7}$	1.68×10^{-04}	376.62	1.36×10^{-04}	396.7
Log-likelihood	12,600.04		11,527.34	
AIC	-25,349.17		-23,427.2	

Table 3
Constant MPR with spot VIX - Model 1

	Spot VIX	V_{t,T_1}	V_{t,T_2}	V_{t,T_3}	V_{t,T_4}	V_{t,T_5}	V_{t,T_6}	V_{t,T_7}	Mean
RMSE	2.34467	1.19300	0.29889	0.38356	0.30035	0.22207	0.16588	0.30777	0.65202
MAE	1.35914	0.74388	0.16491	0.23830	0.20184	0.14560	0.11245	0.21779	0.39799
MAPE	0.06436	0.03414	0.00716	0.01015	0.00870	0.00645	0.00479	0.00911	0.01811

Table 4
Constant MPR without spot VIX - Model 2

	V_{t,T_1}	V_{t,T_2}	V_{t,T_3}	V_{t,T_4}	V_{t,T_5}	V_{t,T_6}	V_{t,T_7}	Mean
RMSE	1.55895	0.47835	0.31856	0.26364	0.23072	0.15489	0.28382	0.46985
MAE	0.95654	0.26773	0.19859	0.17662	0.14993	0.10265	0.19544	0.29250
MAPE	0.04399	0.01165	0.00843	0.00757	0.00663	0.00437	0.00809	0.01296

and the constant components of the TVMPR are significant at 1% level. The same can be said of Model 2. As in the previous case with constant MPR, removing the spot VIX from the sample, decreases the information criteria. In terms of fitting, it can be seen, from Tables 6 and 7, that Model 2 better fits the observed data than Model 1. That is, removing the spot VIX significantly contributes to improve the model fit, on average, especially for medium and long term contracts, despite the marginal deterioration for contracts V_{t,T_5} and V_{t,T_6}. Again one observes the same phenomenon that occurs in the MPR case explained earlier, that is the presence of spot VIX in Model 1 reduces the error of the next two contracts and affects less longer maturities. Hence, in both cases (MPR and TVMPR) removing the spot VIX series improves the adjustment to empirical data for medium and long-term maturities.

Comparing the estimation results in Table 2 and in Table 5 one can observe that the major difference between TVMPR and constant MPR cases is that the market price of risk of the first latent factor is not significant for the TVMPR case. This fact occurred for Model 1 and Model 2. The presence of an non-significant market price of risk in factor models appears in many studies, see for example Schwartz and Smith (2000), Schwartz (1997) and Cortazar and Naranjo (2006). Again, we emphasize that the speed of mean reversion is significant in both cases for Model 1 and Model 2 justifying the hypothesis of a mean-reverting process for the VIX.

Now, comparing the constant MPR with the TVMPR, one can see that considering TVMPR models does not improve the information criteria for the case with the spot VIX in the sample. The same can be said marginally for the case where the spot VIX is omitted.

Figures 1a, 1b, 1c and 1d exhibit the estimated latent factors $\hat{X}_{1,t}$ and $\hat{X}_{2,t}$. In Model 1 (MPR and TVMPR cases in Figures 1a and 1c) both factors have similar output, that is, the presence of spot VIX in the sample makes the evolution a little apart. In Model 2 (MPR and TVMPR in Figures 1b and 1d) both factors evolve approximately in the same range, that is, the exclusion of spot VIX of the sample implies that both factors have nearby paths. Compared to the observed variable V_{t,T_j}, shown on next the figures, one can note that the $X_{1,t}$ factor captures the VIX main movements of ups and downs. We highlight the spike occurred in February 2018. Until this month, the market was stable and calm that a small drop on S&P 500 triggered the VIX and related instruments to spike abruptly.

Figures 2a, 2b, 2c and 2d exhibit the estimated and observed first future contract. Among all seven contracts the first contract presented the greatest error in all cases and models. This is expected since the nearest future contract is the most affected by shocks in the economy. Notwithstanding, one observes

Table 5
Estimation Results for TVMPR approach

	Model 1		Model 2	
Parameters	Value	t-stat	Value	t-stat
κ_1	3.1227	4.81	2.4841	2.87
κ_2	0.9184	7.43	0.7475	5.62
σ_1	0.7367	26.85	0.6460	22.90
σ_2	0.3516	8.36	0.3470	5.80
p_1	1.2815	0.66	-1.0306	-0.53
p_2	-1.8910	-4.19	1.2359	3.03
q_1	-0.2703	-0.30	0.2999	0.22
q_2	-1.7418	-3.26	-1.4357	-2.33
ρ	0.3769	3.02	0.4967	4.29
μ_1	2.0939	14.56	1.3669	13.13
μ_2	-0.3158	-2.30	1.1117	5.53
$s^2_{\varepsilon_{spot}}$	7.45×10^{-03}	600.82		
$s^2_{\varepsilon_1}$	2.15×10^{-03}	514.15	3.22×10^{-03}	359.10
$s^2_{\varepsilon_2}$	1.42×10^{-04}	106.20	2.94×10^{-04}	134.07
$s^2_{\varepsilon_3}$	2.31×10^{-04}	479.30	1.63×10^{-04}	246.98
$s^2_{\varepsilon_4}$	1.68×10^{-04}	458.38	1.33×10^{-04}	360.57
$s^2_{\varepsilon_5}$	1.06×10^{-04}	385.32	1.07×10^{-04}	467.21
$s^2_{\varepsilon_6}$	4.94×10^{-05}	129.26	4.83×10^{-05}	186.55
$s^2_{\varepsilon_7}$	1.62×10^{-04}	413.19	1.35×10^{-04}	396.12
Log-likelihood	12,558.46		11,661.89	
AIC	-25,078.97		-23,417.81	

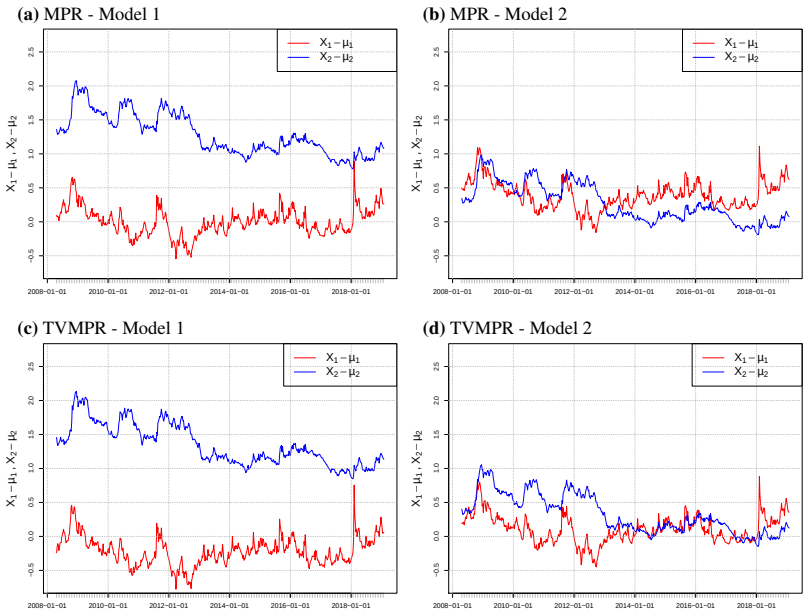
Table 6
TVMPR with spot VIX - Model 1

	Spot VIX	V_{t,T_1}	V_{t,T_2}	V_{t,T_3}	V_{t,T_4}	V_{t,T_5}	V_{t,T_6}	V_{t,T_7}	Mean
RMSE	2.41372	1.17870	0.22932	0.38712	0.30650	0.21998	0.12886	0.28620	0.64380
MAE	1.36695	0.74392	0.12841	0.23790	0.20583	0.14583	0.08779	0.20527	0.39024
MAPE	0.06430	0.03393	0.00557	0.01012	0.00887	0.00647	0.00374	0.00868	0.01771

Table 7
TVMPR without spot VIX - Model 2

	V_{t,T_1}	V_{t,T_2}	V_{t,T_3}	V_{t,T_4}	V_{t,T_5}	V_{t,T_6}	V_{t,T_7}	Mean
RMSE	1.49294	0.43875	0.31243	0.26535	0.22707	0.14536	0.27387	0.45082
MAE	0.92603	0.24827	0.19221	0.17776	0.14885	0.09663	0.19290	0.28323
MAPE	0.04251	0.01079	0.00815	0.00763	0.00658	0.00413	0.00803	0.01255

Figure 1
Output for the latent factors $\hat{X}_{1,t}$ and $\hat{X}_{2,t}$.



that the estimated $\hat{V}_{1,t}$ evolves closely to observed data. The highest level of VIX is related to subprime crises in 2008.

Among all contracts, the lowest error occurred in the sixth future contract. Figures 3a, 3b, 3c and 3d exhibit the estimated and observed variables. One can note that they are almost superimposed.

Table 8 presents the log-likelihood ratio test (LR). One can see that Model 1 (MPR constant) is better than Model 1 (TVMPR). That is, when considering the presence of the spot VIX series in the sample, it is better to use the constant MPR model. The same can not be said about Model 2 (TVMPR) in comparison with Model 2 (MPR). That is, excluding the spot VIX series from the sample makes it more suitable to use the TVMPR model.

5. Applications to ETPs

There are many applications that can be performed based on the results of the estimation above. VIX products such as ETFs and ETNs can be simulated based on the hyperparameters from Tables 1 and 2 . In this section

Figure 2
Estimated and observed first future contract

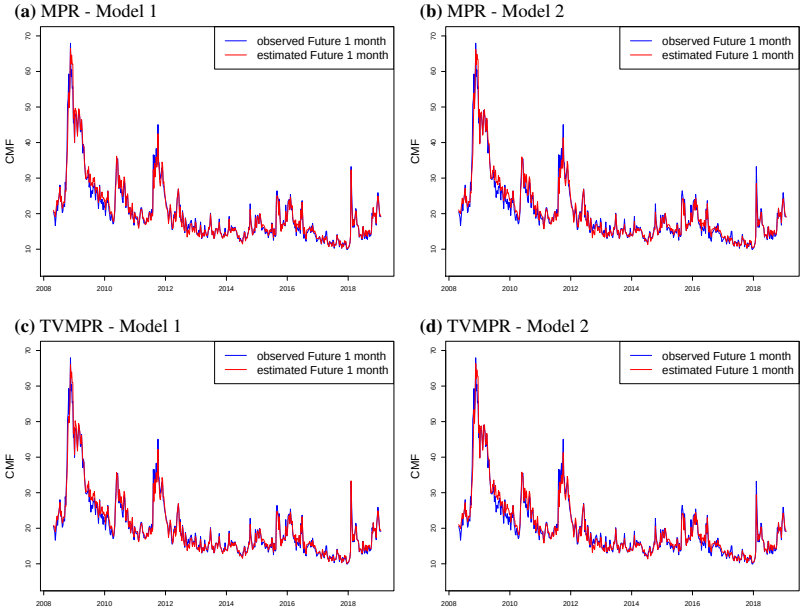
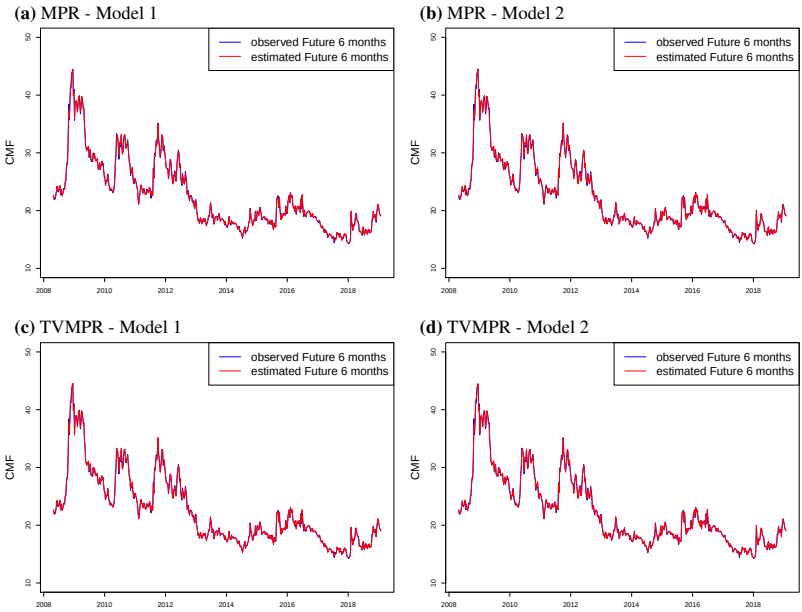


Table 8
Log-likelihood ratio test (LR)

Models	df	$\mathscr{L}$	LR	$\chi^2_{\alpha,df}$	Result
Model 1 (MPR) Model 1 (TVMPR)	2	$\mathscr{L}_R = 12{,}600.04$ $\mathscr{L}_{UR} = 12{,}558.46$	-83.16	5.99	H_0 is not rejected
Model 2 (MPR) Model 2 (TVMPR)	2	$\mathscr{L}_R = 11{,}527.34$ $\mathscr{L}_{UR} = 11{,}661.89$	269.10	5.99	H_0 is rejected

Note: $\mathscr{L}_R$ is the restricted (simple) model's log-likelihood and $\mathscr{L}_{UR}$ is the unrestricted model's log-likelihood. The test statistic is $LR = -2\log(\mathscr{L}_R/\mathscr{L}_{UR})$. If $LR > \chi^2_{\alpha,df}$, where α denotes the significance level and df denotes the degrees of freedom, H_0 is rejected, that is, the unrestricted model is better than its restricted counterpart. Otherwise, H_1 is rejected. Here, $\alpha = 0.05$ was considered.

Figure 3
Estimated and observed sixth future contract



we illustrate the use of simulation on VIX products. ETFs and ETNs are strategies that use rolling positions on VIX future contracts. We refrain from including the derivation of the equations of these instruments to make this section shorter. The details can be found in Bergomi (2016), Avellaneda and Papanicolaou (2019) and Fernandes (2019). Let I_t be the total return strategy of the positions on VIX future contracts. In the present case the CMF V_{t,T_j} represents the VIX future price F_{t,T_j}. Hence, following the references above we have

$$\frac{dI_t}{I_t} = rdt + \sum_j \omega_t^j \frac{dV_{t,T_j}}{V_{t,T_j}} , \tag{5}$$

where ω_t^j denotes the weight allocated to the jth CMF contract. The VXX is an ETN with a long position in the first and second futures. The total return for the VXX strategy is given by:

$$\frac{dI_t}{I_t} = rdt + \omega_t^1 \frac{dV_{t,T_1}}{V_{t,T_1}} + \omega_t^2 \frac{dV_{t,T_2}}{V_{t,T_2}} , \tag{6}$$

where

$$\omega_t^1 = \frac{(T_1 - t)V_{t,T_1}}{(T_1 - t)V_{t,T_1} + (\Delta - (T_1 - t))V_{t,T_1}} ,$$
$$\omega_t^2 = 1 - \omega_t^1 ,$$

and Δ is the time between the expiration of both maturities, approximately 30 days. We proceeded the simulation for the VXX generating many paths. Figure 4 and Figure 5 exhibit the simulated VXX and observed VXX data. The simulated curve presented is the average of all paths. In Figure 4 the simulation was performed for Model 1 using constant market price of risk. Figure 5 presents the simulation using Model 2, again considering a constant market price of risk. In both cases, it can be seen that the mean of simulations is approximately in the middle of the real data which means that the model is appropriate to reproduce the VXX instrument.

Figure 4
Comparison between the simulations exhibiting mean, median, quantiles and one realized path for ETN VXX using Model 1 (MPR)

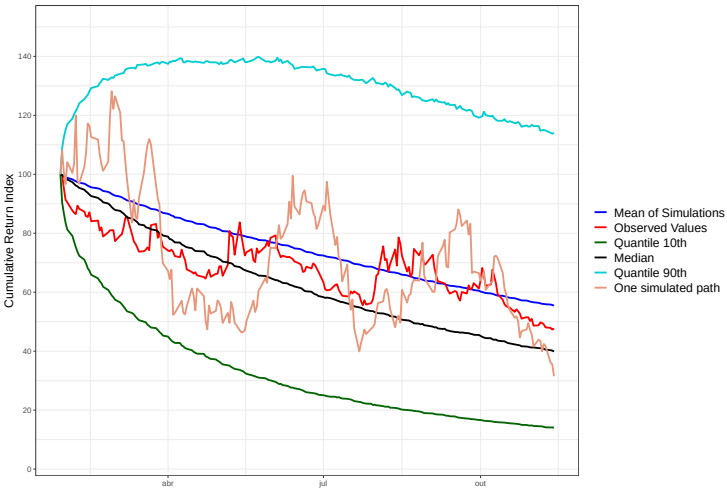
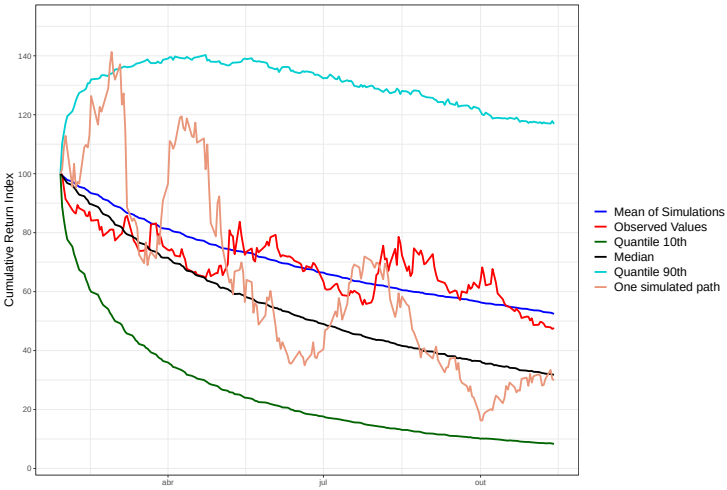


Figure 5
Comparison between the simulations exhibiting mean, median, quantiles and one realized path for ETN VXX using Model 2 (MPR)



6. Conclusions

In this paper we analyzed the estimation of VIX futures contracts following the two-factor model of Avellaneda and Papanicolaou (2019). Inspired by the literature on commodities models, we removed the spot VIX of the sample data. Additionally, we considered the model with constant and time-varying market price of risk. The exclusion of the spot VIX from the sample, resulted in a better fit based on average error between the observed and modeled prices. This fact prevailed in both models with constant MPR and TVMPR. Moreover, the best adjustment occurred mainly for medium and long future maturities. This is a relevant fact for investors positioned on longer maturities.

We also applied the LR test to evaluate the effectiveness of including a time-varying market price of risk in both estimations. We then compared the restricted model (with constant market price of risk) with the unrestricted one (with TVMPR), and the results were mixed. When including the spot VIX in the sample, the constant market price of risk model was preferred. However, when removing the spot VIX from the sample, the TVMPR prevailed.

Finally, we presented an application of the model to VXX strategy using the MPR (with and without the spot VIX). We compare the simulation with the out of sample data and the results showed that the model is appropriate to reproduce the VXX instrument.

References

Alexander, C. and Korovilas, D. (2013). Volatility exchange-traded notes: curse or cure?, *The Journal of Alternative Investments* **16**(2): 52–70.

Avellaneda, M. and Papanicolaou, A. (2019). Statistics of VIX futures and applications to trading volatility exchange-traded products, *International Journal of Theoretical and Applied Finance* **22**(01): 1–30.

Bergomi, L. (2016). *Stochastic volatility modeling*, CRC Press, Taylor & Francis Group.

Boroujerdi, R. and Fogertey, K. (2015). ETFs: The rise of the machines, *Goldman Sachs Equity Research April* .

Cortazar, G. and Naranjo, L. (2006). An N-factor gaussian model of oil futures prices, *Journal of Futures Markets: Futures, Options, and Other Derivative Products* **26**(3): 243–268.

Durbin, J. and Koopman, S. (2012). *Time series analysis by state space methods*, Oxford university press.

Fernandes, F. (2019). *Modelo de 2 fatores para a curva de contratos futuros do VIX: Derivação, calibragem e aplicações*, Master's thesis, Instituto de Matemática Pura e Aplicada (in Portuguese).

Geman, H. and Nguyen, V. (2005). Soybean inventory and forward curve dynamics, *Management Science* **51**(7): 1076–1091.

Gregory, K. and Timcenko, A. (2016). When to get long the VIX: Assessing the risk-reward of long VIX trades, *Goldman Sachs Equity Research April* .

Harvey, A. (1989). *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* pp. 327–343.

Hill, J., Nadig, D. and Hougan, M. (2015). *A comprehensive guide to exchange-traded funds (ETFs)*, CFA Institute Research Foundation.

Schwartz, E. (1997). The stochastic behavior of commodity prices: implications for valuation and hedging, *Journal of Finance* **52**(3): 923–973.

Schwartz, E. and Smith, J. (2000). Short-term variations and long-term dynamics in commodity prices, *Management Science* **46**(7): 893–911.

Shumway, R. and Stoffer, D. (2006). *Time series analysis and its applications, 2nd edition*, Springer Verlag.

Wang, T., Shen, Y., Jiang, Y. and Huang, Z. (2017). Pricing the CBOE VIX futures with the Heston–Nandi GARCH model, *Journal of Futures Markets* **37**(7): 641–659.

Zhang, J. and Zhu, Y. (2006). VIX futures, *Journal of Futures Markets: Futures, Options, and Other Derivative Products* **26**(6): 521–531.

A. Computation of the CMF

In this Appendix we present the computation of the CMF. We define $V_{t,T}$ as the price at time t of a future contract with maturity T. $V_{t,T}$ is obtained as a linear interpolation between two successive monthly contracts which expire immediately before and after T. That is

$$V_{t,T} = \frac{T_{j+1} - T}{T_{j+1} - T_j} F_{t,T_j} + \frac{T - T_j}{T_{j+1} - T_j} F_{t,T_{j+1}} \ ,$$

where T assumes the values (in days) of the vector $[0, 30, 60, 90, 120, 150, 180, 210]$ for each one of the seven maturities considered and the spot VIX.

B. State-Space formulation, Kalman filter and parameter estimation

B.1 State-Space formulation

Let $\boldsymbol{y}_t$ be a multivariate time series of size N. These are the observable variables. They are related to the state variables $\boldsymbol{x}_t$ through the measurement equation as

$$\boldsymbol{y}_t = \boldsymbol{H}\boldsymbol{x}_t + \boldsymbol{d}_t + \boldsymbol{\varepsilon}_t, \quad t = 1, \ldots, T \ ,$$

where $\boldsymbol{H}$ is an $N \times M$ matrix, $\boldsymbol{d}_t$ is an $N \times 1$ vector, $\boldsymbol{\varepsilon}_t$ is a multivariate normal distribution with covariance matrix $\mathbb{Q}$, that is, $\boldsymbol{\varepsilon}_t \sim N(\boldsymbol{0}, \mathbb{Q})$ and $\boldsymbol{x}_t$ is a $M \times 1$ vector of latent (or unobservable) variables.

The latent variables follow a Markovian process given by

$$\boldsymbol{x}_t = \boldsymbol{A}\boldsymbol{x}_{t-1} + \boldsymbol{c}_t + \boldsymbol{\nu}_t \quad t = 1, \ldots, T \ ,$$

where $\boldsymbol{A}$ is an $M \times M$ matrix, $\boldsymbol{c}_t$ is an $M \times 1$ vector and $\boldsymbol{\nu}_t \sim N(\boldsymbol{0}, \mathbb{T})$. Furthermore, the initial state vector $\boldsymbol{x}_0$ has mean given by $\hat{\boldsymbol{x}}_0 = E(\boldsymbol{x}_0)$ and covariance matrix $\boldsymbol{P}_0$. The errors $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\nu}_t$ are uncorrelated and also uncorrelated with the initial state.

B.2 The Kalman filter algorithm

The Kalman filter algorithm is composed of two groups of equations: time update equations and measurement update equations. The time update equations are responsible for projecting forward (in time) the current state $\hat{\boldsymbol{x}}_{t-1}$ to obtain the *a priori* estimates $\hat{\boldsymbol{x}}_t^-$, and the error covariance estimates $\boldsymbol{P}_{t-1}$ to obtain $\boldsymbol{P}_t^-$. In our notation the superscript "minus" sign means the *a priori* stage. The time update step are described by the following state update and covariance update equations, respectively:

$$\hat{\boldsymbol{x}}_t^- = \boldsymbol{A}\hat{\boldsymbol{x}}_{t-1} + \boldsymbol{c}_t \ ,$$

$$\boldsymbol{P}_t^- = \boldsymbol{A}\boldsymbol{P}_{t-1}\boldsymbol{A}^\top + \mathbb{T}\,.$$

Once the *a priori* state and the covariance error estimates are obtained, one follows to the measurement update equations. This step is responsible for incorporating a new measurement into the *a priori* estimate to obtain an improved *a posteriori* estimate. In this set of equations, first the Kalman gain, $\boldsymbol{K}_t$, is computed. Next the *a posteriori* state $\hat{\boldsymbol{x}}_t$ is computed using the *a priori* state $\hat{\boldsymbol{x}}_t^-$, the Kalman gain $\boldsymbol{K}_t$ and the new measurement $\boldsymbol{y}_t$. Finally the covariance error $\boldsymbol{P}_t$ is computed from the *a priori* counterpart $\boldsymbol{P}_t^-$. These are the measurement update equations:

$$\boldsymbol{K}_t = \boldsymbol{P}_t^-\boldsymbol{H}^\top\left(\boldsymbol{H}\boldsymbol{P}_t^-\boldsymbol{H}^\top + \mathbb{Q}\right)^{-1}\,,$$

$$\hat{\boldsymbol{x}}_t = \hat{\boldsymbol{x}}_t^- + \boldsymbol{K}_t\left(\boldsymbol{y}_t - \boldsymbol{H}\hat{\boldsymbol{x}}_t^- - \boldsymbol{d}_t\right)\,,$$

$$\boldsymbol{P}_t = \left(\boldsymbol{I} - \boldsymbol{K}_t\boldsymbol{H}\right)\boldsymbol{P}_t^-\,.$$

After the measurement update is completed, the process is repeated recursively with the previous *a posteriori* estimates ($\hat{\boldsymbol{x}}_t$ and $\boldsymbol{P}_t$) used to predict the new *a priori* estimates ($\hat{\boldsymbol{x}}_{t+1}^-$ and $\boldsymbol{P}_{t+1}^-$).

B.3 The parameter estimation

The initialization of the above algorithm uses the inicial state $\boldsymbol{x}_0$, the initial covariance error $\boldsymbol{P}_0$ and the vector $\boldsymbol{\Theta}$ that contains the initial hyperparameters. Once the recursive procedure from $t = 0$ to $t = T$ is completed a new vector $\boldsymbol{\Theta}$ can be obtained from the likelihood of the prediction errors as following.

Let $f_{\boldsymbol{Y}}\left(\boldsymbol{y}_t|\boldsymbol{Y}_{t-1}\right)$ be the conditional density function given the information up to $t-1$, that is $\boldsymbol{Y}_{t-1} = (\boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}, \ldots, \boldsymbol{y}_1)$. The likelihood function is

$$L(\boldsymbol{y};\boldsymbol{\Theta}) = \prod_{t=1}^{T} f_{\boldsymbol{Y}}\left(\boldsymbol{y}_t|\boldsymbol{Y}_{t-1}\right).$$

Since the errors $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\nu}_t$ are Gaussians, the conditional distribution of $\boldsymbol{y}_t$ given $\boldsymbol{Y}_{t-1}$ is also Gaussian. From previous equations one can obtain the conditional mean $E\left(\boldsymbol{y}_t|\boldsymbol{Y}_{t-1}\right) = \boldsymbol{H}\hat{\boldsymbol{x}}_t^- + \boldsymbol{d}_t$ and the conditional variance $Var\left(\boldsymbol{y}_t|\boldsymbol{Y}_{t-1}\right) = \mathbb{F}_t = \boldsymbol{H}\boldsymbol{P}_t^-\boldsymbol{H}^\top + \mathbb{Q}$.

One can observe that the conditional mean of $\boldsymbol{y}_t$, given $\boldsymbol{x}_t$, depends on matrix $\boldsymbol{H}_t$, on vector $\boldsymbol{d}_t$ and on $\hat{\boldsymbol{x}}_t^-$ which depends on $\boldsymbol{A}_t$ and $\boldsymbol{c}_t$. Hence the conditional mean contains all the hyperparameters. A finite sample of observations $\boldsymbol{y}_t, t = 1, \ldots, T$ will yield some estimates of $\hat{\boldsymbol{H}}, \hat{\boldsymbol{A}}, \hat{\boldsymbol{d}}$ and $\hat{\boldsymbol{c}}$ which will

deviate from the true $\boldsymbol{H}$, $\boldsymbol{A}$, $\boldsymbol{d}$ and $\boldsymbol{c}$ due to the fluctuations of random sampling. The estimates of $\boldsymbol{H}$ and $\boldsymbol{d}$ may be determined with any desired degree of precision and the model is identifiable.

The same point of identifiability can be seen in context of the likelihood. Define the prediction errors as $\boldsymbol{\eta}_t = \boldsymbol{y}_t - \hat{\boldsymbol{y}}_t^-$, $t = 1, \ldots, T$. Hence the likelihood function of the conditional Gaussian prediction errors $\boldsymbol{\eta}_t$ is given by

$$\ln\left[L\left(\boldsymbol{y};\boldsymbol{\Theta}\right)\right] = -\frac{NT}{2}\ln\left(2\pi\right) - \frac{1}{2}\sum_{t=1}^{T}\ln\left|\mathbb{F}_t\right| - \frac{1}{2}\sum_{t=1}^{T}\boldsymbol{\eta}_t^\top\mathbb{F}_t^{-1}\boldsymbol{\eta}_t.$$

The optimal set of parameters $\boldsymbol{\Theta}^\star$ is given by

$$\boldsymbol{\Theta}^\star = \arg\max_{\boldsymbol{\Theta}\in\bar{\boldsymbol{\Theta}}}\left\{\ln\left[L\left(\boldsymbol{y};\boldsymbol{\Theta}\right)\right]\right\},$$

where $\bar{\boldsymbol{\Theta}}$ represents the hyperparameter space. The likelihood function is thus completely specified by the hyperparameters in $\boldsymbol{H}$, $\boldsymbol{A}$, $\boldsymbol{d}$ and $\boldsymbol{c}$ and the model is identifiable.