A Multi-Factor Transformed Diffusion Model with Applications to VIX and VIX Futures

August 2018

Abstract

Transformed diffusions (TDs) have become increasingly popular in financial modelling for their model flexibility and tractability. While existing TD models are predominately one-factor models, empirical evidence often prefers models with multiple factors. We propose a novel distribution-driven nonlinear multi-factor TD model with latent components. Our model is a transformation of a underlying multivariate Ornstein Uhlenbeck (MVOU) process, where the transformation function is endogenously specified by a flexible parametric stationary distribution of the observed variable. Computationally efficient exact likelihood inference can be implemented for our model using a modified Kalman filter algorithm and the transformed affine structure also allows us to price derivatives in semi-closed form. We compare the proposed multi-factor model with existing TD models for modelling VIX. Our results show that the proposed model outperforms all existing TD models both in the sample and out of the sample consistently across all categories and scenarios of our comparison.

JEL Classification: CC13, C32, G13, G15
Keywords: Transformation Model; Nonlinear Diffusion; Latent Factor; Kalman Filter; Volatility Index

1. Introduction

A parametric univariate or one-factor continuous-time diffusion process, say $\{Y_t, t \geq 0\}$, is usually described by the following Stochastic Differential Equation (SDE):

$$dY_t = \mu_Y (y; \psi) \, dt + \sigma_Y (Y_t; \psi) \, dW_t \tag{1}$$

where $\mu_Y (y; \psi)$ and $\sigma_Y^2 (y; \psi)$ are, respectively, the drift and diffusion functions with parameter ψ and $\{W_t, t \geq 0\}$ is a standard Brownian motion. Maximum Likelihood (ML) is usually the preferred method of estimation. However, except for a few special cases such as the Geometric Brownian Motion (GBM) (c.f. Black and Scholes 1973), the Ornstein Uhlenbeck (OU) process (c.f. Vasicek 1977) and the square-root or CIR process (c.f. Cox et al. 1985), most continuous-time diffusion models do not possess closed-form transition densities. Nevertheless, nonlinearities beyond the assumptions of these models are often documented in the literature (c.f. Aït-Sahalia 1996b, Stanton 1997, Bu et al. 2011, Eraker and Wang 2015, and Bu, Cheng and Hadri 2017). One strand of literature aims to find a balance between model flexibility and tractability with no recourse to density approximations. These studies advocate the use of the so-called transformed diffusion (TD) models. TDs are usually nonlinear transformations of tractable, typically affine, underlying diffusions (UDs). Hence, TDs are potentially flexible diffusion models capable of capturing nonlinear features in the data while at the same time possess some desirable analytical and statistical tractability inherited from the more tractable UDs. Primary examples of TDs, among others, include Ahn and Gao (1999), Bu et al. (2011), Goard and Mazur (2013), Forman and Sørensen (2014), Eraker and Wang (2015), and most recently Bu, Jawadi and Li (2017), Bu et al. (2018).

We propose a novel distribution-driven nonlinear multi-factor TD model with latent components. While our approach is applicable to general multi-factor UDs, we propose more specifically a model that is the transformation of a multivariate Ornstein Uhlenbeck (MVOU) process with latent components, where the transformation function is endogenously determined by a flexible parametric specification of the stationary distribution of the observed variable. We show that exact ML inference for our model can be made efficiently by a modified Kalman filter algorithm. We examine the empirical performance of a two-factor specification of the proposed model in comparison with existing models for modelling the dynamics of VIX and for pricing VXF contracts. We base our comparison on both the in-sample model fitness criteria and the out-of-sample Root Mean Square Forecasting Error (RMSFE) for modelling VIX. Our results strongly favor our distribution-driven two-factor model, which outperforms all alternative TD models strongly and consistently across all the categories and scenarios of our comparison.

2. A Multi-Factor Transformed Diffusion Model

2.1. The Framework

The TD approach assumes that the observed diffusion process Y is a strictly monotone and sufficiently smooth function of some UD X. More specifically, it assumes that

$$Y_t = V(X_t; \vartheta) \tag{2}$$

$$dX_t = \mu_X (X_t; \omega) \, dt + \sigma_X (X_t; \omega) \, dW_t \tag{3}$$

where $\mu_X(X_t;\omega)$ and $\sigma_X^2(X_t;\omega)$ are the drift and diffusion functions of X with parameter ω. $V(x;\vartheta)$ or equivalently its unique inverse $U(y;\vartheta) = V^{-1}(y;\vartheta)$ is known as the transformation function with parameter ϑ, satisfying $\partial V(x;\vartheta)/\partial x \neq 0$ for all x on its domain D_X. More specifically, let $p_X(x|x_0,\Delta;\omega)$ and $p_Y(y|y_0,\Delta;\psi)$ be the transition density function of X and Y, respectively, where Δ is the time interval. It follows immediately that

$$p_Y(y|y_0,\Delta;\psi) = |U'(y;\vartheta)|\, p_X(U(y;\vartheta)\,|U(y_0;\vartheta), \Delta;\omega)$$

Suppose that we wish to model a diffusion process Y, assuming that $Y_t = V(X_t;\vartheta)$. Crucially, we now assume that the SDE of X can be written as

$$dX_t = \mu_X(X_t;\theta_t,\omega)\,dt + \sigma_X(X_t;\theta_t,\omega)\,dW_{X,t} \quad (4)$$
$$d\theta_t = \mu_\theta(\theta_t;X_t,\omega)\,dt + \sigma_\theta(\theta_t;X_t,\omega)\,dW_{\theta,t} \quad (5)$$

where θ_t is a latent process. The two-dimensional vector $Z_t = (X_t,\theta_t)^T$ follows a bivariate diffusion system with parameter ω. It is important to assume that the bivariate diffusion Z satisfy the regularity conditions set out in Aït-Sahalia (2008, Assumptions 1-4) which ensure that Z admits a unique weak solution in terms of the bivariate transition density $p_Z(z|z_0,\Delta;\omega)$. It then follows that continuous-time dynamics and the transition density of the transformed system $\tilde{Z}_t = (Y_t,\theta_t)^T = (V(X_t),\theta_t)^T$ can be obtained by the multivariate version of the Ito's Lemma and the usual Jacobian method, respectively.

3. Empirical Comparison

3.1. The Data

We compare the empirical performance of the newly proposed distribution-driven multi-factor TD with latent component model with existing TD models for modelling the dynamics of VIX and pricing VXFs. Our data consist of daily VIX indices from January 2, 1990 to March 20, 2015 (6352 observations) and VXF closing prices from March 26, 2004 to February 17, 2015 (19215 observations).

We plot the time series of daily VIX and the term structure of constant maturity VXF prices in Figure 1 and 2, respectively, and some summary statistics are reported in Table 1. The evolution of VIX indicates that the mean reversion is weak when the level of VIX is low but much stronger when it is high. This suggests that a suitable diffusion model for VIX should have a drift function that is close to zero when VIX is low and strongly negative when VIX is high. Meanwhile, the local volatility of VIX is also low when VIX is low and substantially higher otherwise. This suggests that a suitable diffusion model should also have a diffusion function that increases rapidly in VIX. The mean of VIX is 20.61 and the standard deviation is 10.19. The large skewness 2.21 and kurtosis 9.25 suggest strong deviation from normality. Augmented Dickey-Fuller tests on these time series all rejected the unit root hypothesis with 4 lags at 5% significance level. Therefore, the use of stationary diffusion models is justified. More importantly, Mencía and Sentana (2013) show that the daily VIX series exhibits the ARMA(2,1) autocorrelation structure. Thus, the use of our proposed two-factor TD model is justified, since it can be easily verified that it implies the ARMA(2,1) structure. Meanwhile, the term structure of VXFs is relatively flat and the evolutionary paths of the seven series are highly correlated. The first two eigenvalues of the correlation matrix dominate the others, explaining approximately 99.9% of the cross sectional variation in these series. This is further justification for the use of a two-factor models for pricing VXFs.

[Figure 1 and 2]
[Table 1]

3.2. Analysis of Time Series of VIX

We first examine the performance of competing models for modelling the VIX time series. We investigate both the in-sample goodness-of-fit measure and out-of-sample forecasting accuracy as well as consider three specification tests. One of the main advantages of TDs is the availability of closed-form transition densities. Thus, ML is our preferred choice of estimation method. The ML estimates of the parameters of competing models are reported in the top panel Table 2. Our initial unconstrained estimation of the CIREW model resulted in a negative estimate of δ. Since δ is the lower bound of the support implied by the CIREW model, a negative δ is inconsistent with the nature of VIX. We then re-estimated the model by imposing $\delta = 0$. Meanwhile, η determines the upper bound of support. Therefore, no standard errors are reported for these two parameters. Moreover, when estimating the three distribution-driven models, we profiled out the parameters m_2 and s_2 of the M2LN distribution by matching the model-implied stationary mean and variance with the sample mean and variance of the VIX data. Furthermore, for the BVOUM2LN model, normalization requires the parameters θ and σ_θ to be constrained, and thus no standard errors are reported for the estimates of m_2, s_2, θ and σ_θ. For the same reason, no standard errors are reported for σ of the CIRM2LN model and θ and σ of the OUM2LN model either.

[Table 2]

3.2.1. In-Sample Performance

We first examine the general goodness-of-fit of each model to the VIX data in terms of LL, AIC and BIC measures reported in the middle panel of Table 2. The relative ranking of each model is the same in terms of any of the three measures. The worst performing model is the CIR model, followed closely by the OUDO model. This is expected, because the CIR model is a simple linear model and also the exponential transformation of the OUDO model offers no effective degrees of freedom to the simple linear underlying OU model. The remaining one-factor models all have parameter-dependent transformations. Consequently, the goodness-of-fit of these models are significantly better than the two benchmark models.

It is interesting to examine the relative performance of the three transformed CIR models, since they represent the drift-driven, the diffusion-driven and the distribution-driven models, respectively. The distribution-driven CIRM2LN model provided slightly better fit than the diffusion-driven CIRCEV, which slightly outperformed the drift-driven CIREW model. Compared to the CIREW model, the CIRCEV model has a well defined support on $D_Y = (0, \infty)$, making it a naturally coherent model for variables such as nominal interest rates and VIX. In addition, the CIRCEV model has a closed-form conditional mean and hence a closed-form pricing formula for the VXFs. These features make the CIRCEV model a very attractive alternative to the CIREW model in practice.

We plot in Figure 3 the estimated drift and diffusion functions of the one-factor models. We can see that when VIX is low, the estimated functions are relatively close among different models, but their differences increase quite dramatically as VIX increases. As we have seen from Figure 1, strong mean reversion and high volatility at high levels of VIX is a prominent feature of the VIX data. However, both functions of the CIR model are linear and very flat, unable to generate strong enough mean reversion or large enough volatility at high levels of VIX. All the other one-factor models have nonlinear drift and diffusion functions, but the distribution-driven CIRM2LN model has the strongest mean reversion and the largest volatility at high levels of VIX, unsurprisingly making it the best fitting one-factor model. Intuitively, the flexible M2LN distribution captures the information particularly in the right tail of the distribution much better than other models. This information is then suitably incorporated into the shapes of the drift and the diffusion functions to produce a better fit to the data. The estimated functions for the remaining one-factor models are relatively close.

[Figure 3]

We now turn our attention to the two-factor models. It is very interesting to note that despite the presence of a latent central tendency factor, the BVOUMS model only performed better than the benchmark CIR and OUDO models and was even outperformed by all other one-factor models. This is potentially an extremely important observation, as this suggests that at least for our data, the flexibility provided by the parameter-dependent nonlinear transformations play a more important role than the additional latent factor, if either but not both is included. Meanwhile, the BVOUM2LN model outperformed all other models by quite clear margins. This is expected, because the BVOUM2LN model contains not only a flexible parameter-dependent distribution-driven transformation function, designed to capture potentially crucial information in the stationary long-run behavior of VIX, but also a latent factor, which tracks the stochastic short-run central tendency of movement of VIX.

We plot in Figure 4 the estimated drift functions for the two two-factor models, conditional on the latent factors, taking several values between the 1st and the 99th quantile of their estimated stationary distributions. Note that the conditional drift functions for both models are nonlinear and relatively close when VIX is low, but their differences start to emerge as VIX goes up. Specifically, as VIX increases, the spread of the conditional drift functions across different values of the latent factor becomes wider for the BVOUM2LN model than for the BVOUMS model. Another striking difference is that the conditional drift functions of the BVOUMS model are globally concave in the level of VIX, but those of the BVOUM2LN model are not. We can see quite clearly that when VIX is in the middle range, the drift functions of the BVOUM2LN model have some degrees of convexities conditional on medium to low values of the latent factor. The wider spread and the higher degrees of nonlinearities of the conditional drift functions of the BVOUM2LN model are but potentially vital differences in explaining the dynamics of VIX. We can only attribute these to the flexible distribution-driven transformation. That is, the M2LN distribution can more flexibly capture the spread and variation in the density curve of the stationary distribution of the VIX data, and crucially such distributional features are then constructively translated into the variations in the estimated drift functions.

[Figure 4]

To further demonstrate the differences of the two models, we plot in the left panel of Figure 5 the estimated stationary densities of the two two-factor models together with that of the benchmark CIR model and the nonparametric kernel density. As we can see, the implied stationary density by the BVOUM2LN model matches the kernel density very closely, incorporating most, if not all, key distributional features of the data. In contrast, a very large proportion in the middle of the stationary density implied by the BVOUMS model departed significantly from the kernel density, leading to significant differences in the estimated functions and goodness of fit to the data. Furthermore, we plot in the right panel of Figure 5 the estimated diffusion functions for the three models together with the nonparametric kernel diffusion estimate. Compared to the flat linear diffusion function of the CIR model, that of the BVOUMS model is nonlinear and increases in VIX, but it is only to a limited extent. The BVOUM2LN model, however, shows much stronger nonlinearity and produces almost twice as much volatility for high levels of VIX, which is more consistent with our observation from the time series plot of the VIX. Most importantly, the estimated diffusion function of the BVOUM2LN model matches quite closely with the nonparametric estimate in terms of both the level and the slope. In clear contrast, however, that of the BVOUMS model deviates quite substantially from the nonparametric estimate, with no overlapping whatsoever except for very low levels of VIX. Above all, the superior suitability of the BVOUM2LN model over other models is quite clear.

[Figure 5]

3.2.2. Out-of-Sample Performance

We now compare models in terms of their out-of-sample forecasting accuracy. For each model, we produce six series of rolling sample conditional mean forecasts for the VIX corresponding to forecasting horizons of 1 day (1D), 1 week (1W), and 1, 3, 5, 7 months (1M, 3M, 5M, 7M). For each horizon, we compute the RMSFE based on the observed out-of-sample series and its rolling sample forecasts produced by each model, and report the results in the bottom panel of Table 2.

For forecasting at 1D and 1W horizons, the CIR model is the best performing one-factor model (1.207 and 2.601) followed immediately by the OUDO model (1.208 and 2.618). At 1M horizon, however, the best performing one-factor model is the OUDO model (3.616) followed by the CIR model (3.626). The fact that they outperform the remaining one-factor models at these horizons suggests that at relatively short horizons, nonlinear parameter-dependent transformations may not significantly improve the ability of one-factor TD models to track the conditional mean. However, as the forecasting horizon increases to medium range (3M) and long range (5M and 7M), the performance of the CIR and OUDO models deteriorate significantly and are then exceeded by other one-factor TD models. This is not surprising, because at short forecasting horizons, the conditional distributions of all diffusion models are close to the normal distribution. Thus, the ability of more flexible models is minimized, but more parsimonious models usually have the advantage. At longer horizons, however, the conditional mean tends to depend more on the information in the stationary distribution (long-run behavior) implied by the forecasting model, for which more sophisticated models, particularly our distribution-driven models have the advantages. This explains why, among one-factor models, the two simplest models performed the best in short horizons and the worst in medium to long horizons forecasts.

The main advantage of the two two-factor models is that the additional central tendency variable can model the evolution of the conditional mean with more flexibility. Thus, we expect the two two-factor models to perform well at varied horizons. We also expect the newly proposed BVOUM2LN model to perform better than the BVOUMS model, since the distribution-driven transformation is expected to capture the nonlinear dynamics and particularly the information in the stationary distribution more effectively. Both of our expectations are confirmed by the forecasting results. Firstly, we find that both two-factor models outperformed all one-factor models at all forecasting horizons. In particular, the margins are more substantial for longer forecasting horizons than for shorter horizons. More importantly, both two-factor models outperformed one-factor models even at the shortest horizon. Comparing between the two two-factor models, the newly proposed BVOUM2LN model, which has additional degrees of freedom in the transformation function, outperformed the BVOUMS model by significant margins at all forecasting horizons. Most importantly, the advantage increases monotonically as forecasting horizon increases, confirming that the superiority of the BVOUM2LN model can indeed be attributed to its distribution-driven transformation design to incorporate information in the stationary distribution (long-run behavior) of the dynamics of VIX. In summary, the new BVOUM2LN model outperformed all competing models both in-sample and out-of-sample in every category and scenario of comparison that we considered.

References

Ahn, D. H., and Gao, B. (1999). A Parametric Non-linear Model of Term Structure Dynamics. Review of Financial Studies, 12, 721-762.

Aït-Sahalia, Y. (1996a). Nonparametric Pricing of Interest Rate Derivatives. Econometrica, 64, 527-560.

Aït-Sahalia, Y. (1996b). Testing Continuous-Time Models of the Spot Interest Rate. Review of Financial Studies, 9, 385-426.

Aït-Sahalia, Y. (2002). Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-form Approximation Approach. Econometrica, 70, 223-262.

Aït-Sahalia, Y. (2008). Closed-Form Likelihood Expansions for Multivariate Diffusions. Annals of Statistics, 36, 906-937.

Aït-Sahalia, Y., R. Kimmel., (2007). Maximum Likelihood Estimation of Stochastic Volatility Models. Journal of Financial Economics, 83, 413-452.

Aït-Sahalia, Y., and Park, J. (2016). Bandwidth Selection and Asymptotic Properties of Local Nonparametric Estimators in Possibly Nonstationary Continuous-Time Models, Journal of Econometrics, 192(1), 119-138.

Amengual, D., and Xiu, D. (2018). Resolution of Policy Uncertainty and Sudden Declines in Volatility, Journal of Econometrics, 203(2), 297-315.

Bahra, B. (1997). Implied Risk-Neutral Probability Density Functions from Option Prices: Theory and Application. Working Paper No 66, Bank of England.

Balduzzi, P., Das, S.R., Foresi, S., and Sundaram, R.K. (1996). A Simple Approach to Three-Factor Affine Term Structure Models. Journal of Fixed Income, 6(3), 43-53.

Black, F., and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities, Journal of Political Economy, 81, 637-654.

Brandt, M., and Santa-Clara, P. (2002). Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets. Journal of Financial Econometrics, 63, 161–210.

Brigo, D., and Mercurio, F. (2002). Lognormal-Mixture Dynamics and Calibration to Market Volatility Smiles. International Journal of Theoretical and Applied Finance, 5(4), 427-446.

Bu, R., Cheng, J., and Hadri, K. (2016). Reducible Diffusions with Time-Varying Transformations with Application to Short-Term Interest Rates. Economic Modelling, 52, 266-277.

Bu, R., Cheng, J., and Hadri, K. (2017). Specification Analysis in Regime-Switching Continuous-Time Diffusion Models for Market Volatility. Studies in Nonlinear Dynamics and Econometrics, 21(1), 65-80.

Bu, R., Giet, L., Hadri, K., and Lubrano, M. (2011). Modelling Multivariate Interest Rates Using Time-Varying Copulas and Reducible Non-linear Stochastic Differential Equations. Journal of Financial Econometrics, 9(1), 198-236.

Bu, R., and Hadri, K. (2007). Estimating Option Implied Risk-Neutral Densities using Spline and Hypergeometric Functions. Econometrics Journal, 10, 216-244.

Bu, R., Hadri, K., and Kristensen, D. (2018). Diffusion Copulas: Identification and Estimation, Working Paper.

Bu, R., Jawadi, F., and Li, Y. (2017). An Empirical Comparison of Transformed Diffusion Models for VIX and VIX futures. Journal of International Financial Markets, Institutions & Money, 46, 116-127.

Chan, K. C., Karolyi, A., Longstaff, F., and Sanders, A. (1992). An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. Journal of Finance, 47, 1209-1227.

Choi, S. (2009). Regime-Switching Univariate Diffusion Models of the Short-Term Interest Rate. Studies in Nonlinear Dynamics and Econometrics, 13(1): Article 4.

Conley, T. L., Hansen, L., Luttmer, E., and Scheinkman, J. (1997). Short Term Interest Rates as Subordinated Diffusion. Review of Financial Studies, 10, 525-578.

Cox, J., Ingersoll, J., and Ross, S. (1985). In Intertemporal General Equilibrium Model of Asset Prices. Econometrica, 53, 363-384.

Dai, Q., and Singleton, K.J. (2000). Specification Analysis of Affine Term Structure Models, Journal of Finance, 55, 1943-1978.

Detemple, J., and Osakwe, C. (2000). The Valuation of Volatility Option, European Finance Review, 4, 21-50.

Duffie, D., and Kan, R. (1996). A Yield Factor Model of Interest Rates. Mathematical Finance, 6, 379-406.

Elerian, O., Chib, S., and Shephard, N. (2001). Likelihood Inference for Discretely Observed Non-Linear Diffusions. Econometrica, 69, 959-993.

Eraker, B. (2001). MCMC Analysis of Diffusion Models with Application to Finance. Journal of Business and Economic Statistics, 19(2), 177-191.

Eraker, B., and Wang, J. (2015). A Non-linear Dynamic Model of the Variance Risk Premium. Journal of Econometrics, 187, 547-556.

Forman, J. L., and Sørensen, M., (2014). A Transformation Approach to Modelling Multi-Modal Diffusions. Journal of Statistical Planning and Inference, 146, 56-69.

Goard, J., and Mazur, M. (2013). Stochastic Volatility Models and the Pricing of VIX Options. Mathematical Finance, 23, 439-458.

Grünbichler, A., and Longstaff, F. (1996). Valuing Futures and Options on Volatility, Journal of Banking Finance, 20, 985-1001.

Hansen, B. (1994). Autoregressive Conditional Density Estimation. International Economic Review, 35, 705-730.

Hansen, L.P. and Sargent, T.J. (1983). The Dimensionality of the Aliasing Problem in Models with Rational Spectral Densities. Econometrica, 51, 377-387.

Jiang, G., and J. Knight., (1997). A Nonparametric Approach to the Estimation of Diffusion Processes with an Application to a Short-Term Interest Rate Model. Econometric Theory, 13, 615–645.

Karlin, S., and Taylor, H.M. (1981). A Second Course in Stochastic Processes, Academic Press, New York.

Kessler, M. (1997). Estimation of an Ergodic Diffusion from Discrete Observations. Scandinavian Journal of Statistics, 24, 211-229.

Kessler, M., and Rahbek, A. (2004). Identification and Inference for Multivariate Cointegrated and Ergodic Gaussian Diffusions. Statistical Inference for Stochastic Process, 7, 137-151.

Karatzas, I., and Shreve, S.E. (1991). Brownian Motion and Stochastic Calculus, Second Edition, Springer-Verlag, New York.

Litterman, R., and Scheinkman, J. (1991). Common Factors Affecting Bond Returns. Journal of Fixed Income, 1(1), 54-61.

Lo, A.W. (1988). Maximum Likelihood Estimation of Generalized Ito Processes with Discretely Sampled Data. Econometric Theory, 4, 231-247.

Melick, W.R., and Thomas, C.P., (1997). Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis. Journal of Financial and Quantitative Analysis, 32, 91-115.

Mencía, J., and Sentana, E. (2013). Valuation of VIX Derivatives. Journal of Financial Economics, 108, 367-391.

Merton, R. C. (1973). Theory of Rational Option Pricing. Bell Journal of Economics and Management Science, 4, 141-183.

Pedersen, A.R. (1995). A New Approach to Maximum-likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations. Scandinavian Journal of Statistics, 22, 55-71.

Pritsker, M. (1998). Nonparametric Density Estimation and Tests of Continuous Time Interest Rate Models. Review of Financial Studies, 11(3), 449-487.

Shoji, I., and Ozaki, T. (1998). Estimation for Non-linear Stochastic Differential Equations by a Local Linearization Method. Stochastic Analysis and Applications, 16, 733-752.

Söderlind, P., and Svensson, L. (1997). New Techniques to Extract Market Expectations from Financial Instruments, Journal of Monetary Economics, 40, 383-429.

Song, Z., and Xiu, D. (2016). A Tale of Two Option Markets: Pricing Kernels and Volatility Risk. Journal of Econometrics, 190, 176-196.

Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure, Journal of Financial Economics, 5, 177-188.

Figure 1: Time Series of Daily VIX

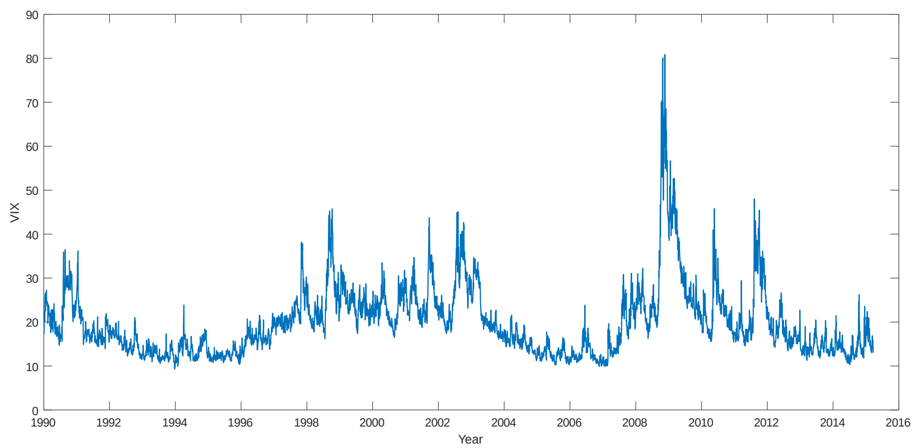

Figure 2: Term Structure of Constant Maturity VXFs

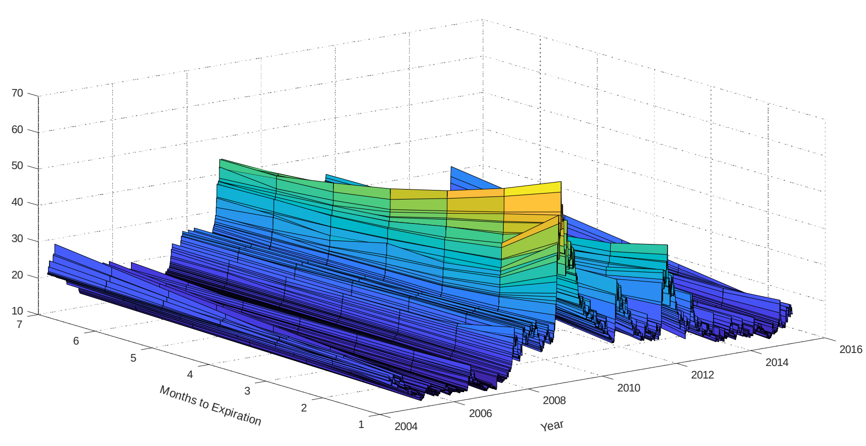

Figure 3: Estimated Drift and Diffusion Functions of One-Factor TD Models

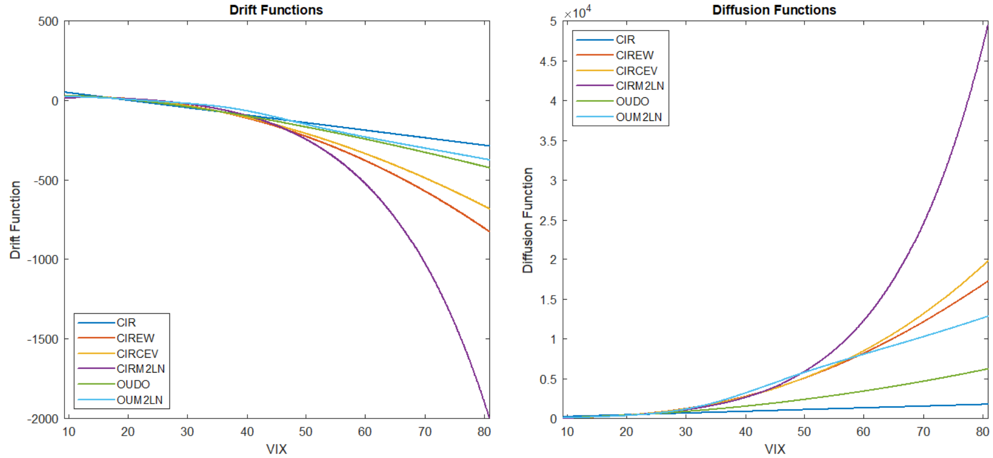

Figure 4: Estimated Conditional Drift Functions of Two-Factor TD Models

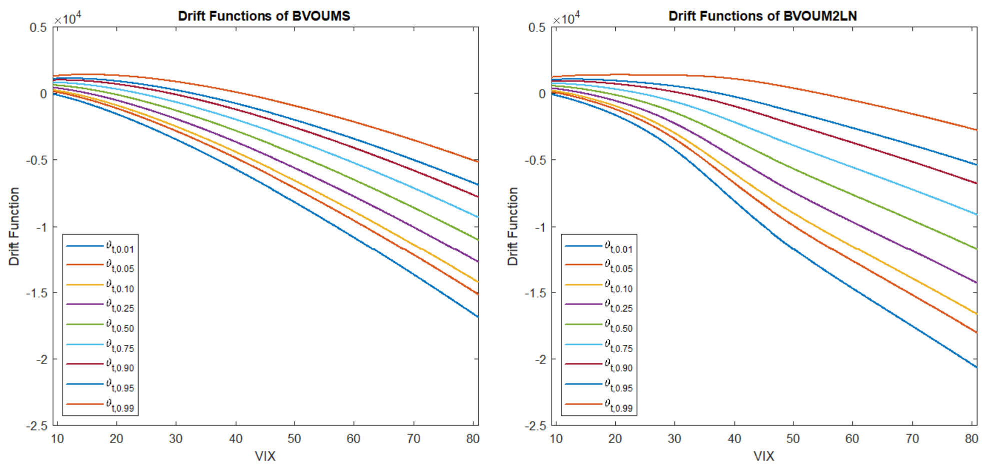

Figure 5: Estimated Stationary Densities and Diffusion Functions

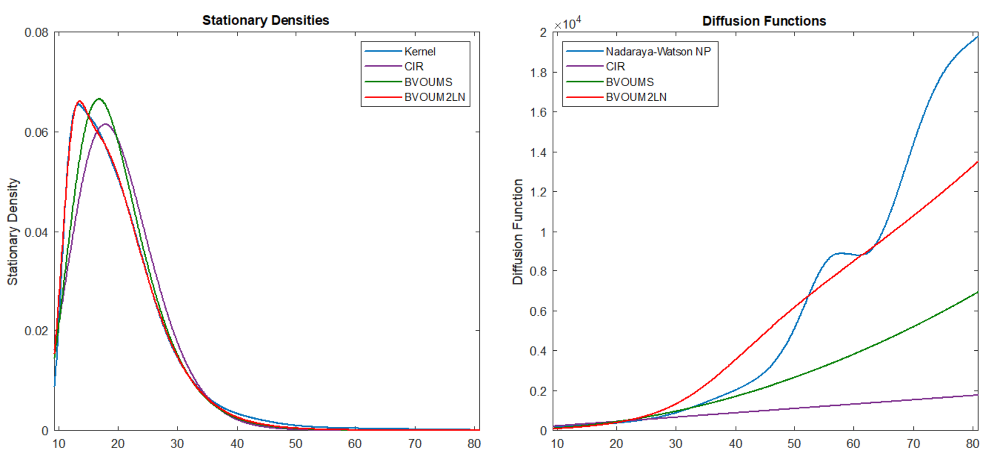

Table 1: Summary of VIX and Constant Maturity VXFs

	VXF							VIX
	1M	2M	3M	4M	5M	6M	7M	
Correlation	1.000	0.989	0.974	0.958	0.941	0.925	0.912	
		1.000	0.996	0.986	0.975	0.963	0.953	
			1.000	0.997	0.990	0.982	0.973	
				1.000	0.998	0.992	0.987	
					1.000	0.998	0.994	
						1.000	0.999	
							1.000	
Eigenvalue	6.852	0.136	0.008	0.002	0.001	0.000	0.000	
Maximum	65.462	59.004	53.782	49.727	46.707	44.635	44.085	80.860
Minimum	11.287	12.230	12.540	12.870	13.200	13.530	13.687	9.310
Mean	20.502	21.128	21.521	21.824	22.096	22.343	22.539	19.921
Median	18.180	19.312	19.944	20.506	20.972	21.272	21.534	18.130
Std. Dev.	8.351	7.672	7.153	6.806	6.566	6.384	6.231	7.982
Skewness	1.845	1.535	1.279	1.082	0.951	0.854	0.794	2.072
Kurtosis	7.070	5.815	4.717	3.936	3.508	3.232	3.085	10.466

Table 2: **Estimation and Forecasting Results for VIX**

	One-Factor Models						Two-Factor Models	
	Transformed CIR				Transformed OU		Transformed BVOU	
	CIR	CIREW	CIRCEV	CIRM2LN	OUDO	OUM2LN	BVOUMS	BVOUM2LN
κ	4.717	3.630	3.749	2.901	3.923	3.240	93.801	89.578
	(0.634)	(0.558)	(0.565)	(0.253)	(0.578)	(0.116)	(13.008)	(12.888)
θ	20.158	0.053	0.133	0.462	2.938	0	2.934	0
	(0.906)	(0.003)	(0.038)	(0.044)	(0.051)	-	(0.076)	-
σ	4.666	0.231	0.282	1	0.974	1	1.030	2.669
	(0.043)	(0.005)	(0.022)	-	(0.009)	-	(0.012)	(0.055)
κ_θ							1.588	1.157
							(0.372)	(0.140)
σ_θ							0.592	1.500
							(0.032)	-
γ			1.414					
			(0.027)					
δ		0						
		-						
η		0.005						
		-						
w				0.916		0.865		0.852
				(0.030)		(0.039)		(0.038)
θ_1				2.983		2.879		2.873
				(0.017)		(0.010)		(0.011)
s_1				0.365		0.331		0.330
				(0.005)		(0.006)		(0.006)
θ_2				2.483		3.336		3.333
				-		-		-
s_2				0.155		0.449		0.444
				-		-		-
LL ($\times 10^3$)	-10.006	-9.307	-9.305	-9.283	-9.436	-9.324	-9.364	**-9.251**
AIC ($\times 10^4$)	2.002	1.862	1.862	1.858	1.888	1.866	1.874	**1.852**
BIC ($\times 10^4$)	2.004	1.865	1.865	1.861	1.890	1.868	1.877	**1.856**
RMSFE_1D	1.207	1.211	1.210	1.211	1.208	1.210	1.206	**1.204**
RMSFE_1W	2.601	2.640	2.635	2.646	2.618	2.633	2.495	**2.484**
RMSFE_1M	3.626	3.697	3.674	3.727	3.616	3.639	3.402	**3.371**
RMSFE_3M	4.718	4.358	4.342	4.331	4.362	4.193	3.410	**3.244**
RMSFE_5M	5.474	5.185	5.164	5.160	5.162	5.093	4.066	**3.870**
RMSFE_7M	5.136	4.922	4.898	4.904	4.884	4.891	3.811	**3.630**