Implied Volatility Surface Dynamics: from Parameter-Driven to Observation-Driven Models and Beyond

Xia Zou,[a] *Yicong Lin,*[a] *André Lucas*[a]

[a] Vrije Universiteit Amsterdam and Tinbergen Institute

August 27, 2024

Table of Contents

1 Introduction

2 The Model

3 Empirical study

Motivation

- Modeling the dynamics in the Implied Volatilities (IV) *surface*: common factors.
- A typical approach for the dynamics in factors: state-space framework (parameter driven).
- Challenge: state-space models beyond the linear Gaussian is computationally expensive.
- An alternative: score-driven approach (observation-driven) with an explicit expression for the likelihood function.

Motivation

- Modeling the dynamics in the Implied Volatilities (IV) *surface*: common factors.
- A typical approach for the dynamics in factors: state-space framework (parameter driven).
- Challenge: state-space models beyond the linear Gaussian is computationally expensive.
- An alternative: score-driven approach (observation-driven) with an explicit expression for the likelihood function.

Motivation

- Modeling the dynamics in the Implied Volatilities (IV) *surface*: common factors.
- A typical approach for the dynamics in factors: state-space framework (parameter driven).
- Challenge: state-space models beyond the linear Gaussian is computationally expensive.
- An alternative: score-driven approach (observation-driven) with an explicit expression for the likelihood function.

Motivation

- Modeling the dynamics in the Implied Volatilities (IV) *surface*: common factors.
- A typical approach for the dynamics in factors: state-space framework (parameter driven).
- Challenge: state-space models beyond the linear Gaussian is computationally expensive.
- An alternative: score-driven approach (observation-driven) with an explicit expression for the likelihood function.

Motivation

- Koopman et al. (2016) in a univariate framework: point forecasts and density forecasts based on simple score-driven models perform similarly to their state-space counterpart.
- Koopman et al. (2017) in a multivariate framework:
  - Point forecasts: comparable performance.
  - Density forecasts: much worse performance of score-driven models.
- We explore the origins of this difference in a multivariate framework between the two model classes in more detail.
- Koopman et al. (2017): by assuming the error terms in the score-driven have an equicorrelation structure, the performances become more comparable.

Motivation

- Koopman et al. (2016) in a univariate framework: point forecasts and density forecasts based on simple score-driven models perform similarly to their state-space counterpart.
- Koopman et al. (2017) in a multivariate framework:
  - Point forecasts: comparable performance.
  - Density forecasts: much worse performance of score-driven models.
- We explore the origins of this difference in a multivariate framework between the two model classes in more detail.
- Koopman et al. (2017): by assuming the error terms in the score-driven have an equicorrelation structure, the performances become more comparable.

Motivation

- Koopman et al. (2016) in a univariate framework: point forecasts and density forecasts based on simple score-driven models perform similarly to their state-space counterpart.
- Koopman et al. (2017) in a multivariate framework:
    - Point forecasts: comparable performance.
    - Density forecasts: much worse performance of score-driven models.
- We explore the origins of this difference in a multivariate framework between the two model classes in more detail.
- Koopman et al. (2017): by assuming the error terms in the score-driven have an equicorrelation structure, the performances become more comparable.

Motivation

- Koopman et al. (2016) in a univariate framework: point forecasts and density forecasts based on simple score-driven models perform similarly to their state-space counterpart.
- Koopman et al. (2017) in a multivariate framework:
  - Point forecasts: comparable performance.
  - Density forecasts: much worse performance of score-driven models.
- We explore the origins of this difference in a multivariate framework between the two model classes in more detail.
- Koopman et al. (2017): by assuming the error terms in the score-driven have an equicorrelation structure, the performances become more comparable.

Motivation

- Koopman et al. (2016) in a univariate framework: point forecasts and density forecasts based on simple score-driven models perform similarly to their state-space counterpart.
- Koopman et al. (2017) in a multivariate framework:
  - Point forecasts: comparable performance.
  - Density forecasts: much worse performance of score-driven models.
- We explore the origins of this difference in a multivariate framework between the two model classes in more detail.
- Koopman et al. (2017): by assuming the error terms in the score-driven have an equicorrelation structure, the performances become more comparable.

Contributions

- We pinpoint the origin of performance difference in the state-space and the score-driven framework: a (too) restrictive assumption on the covariance structure of the measurement noise.
- We introduce a simple adaptation of the measurement equation in the score-driven model $\rightarrow$ comparable density forecast performance of score-driven models with their state-space counterparts.
- After closing the performance gap, the score-driven approach can easily be adapted to accommodate non-Gaussian features, without any complication to the ML estimation and inference procedures.

Contributions

- We pinpoint the origin of performance difference in the state-space and the score-driven framework: a (too) restrictive assumption on the covariance structure of the measurement noise.
- We introduce a simple adaptation of the measurement equation in the score-driven model $\rightarrow$ comparable density forecast performance of score-driven models with their state-space counterparts.
- After closing the performance gap, the score-driven approach can easily be adapted to accommodate non-Gaussian features, without any complication to the ML estimation and inference procedures.

Contributions

- We pinpoint the origin of performance difference in the state-space and the score-driven framework: a (too) restrictive assumption on the covariance structure of the measurement noise.
- We introduce a simple adaptation of the measurement equation in the score-driven model $\rightarrow$ comparable density forecast performance of score-driven models with their state-space counterparts.
- After closing the performance gap, the score-driven approach can easily be adapted to accommodate non-Gaussian features, without any complication to the ML estimation and inference procedures.

Empirical Study

- We apply our findings to model the dynamics of IV surfaces of S&P500 index options using data from January 2010 to December 2022.
- We find that a linear Gaussian state-space model outperforms a plain-vanilla score-driven model by a large margin, both in terms of density fit, and Value-at-Risk (VaR) violation rates.
- After adapting the score-driven model with the adjusted covariance structure for the measurement errors, the score-driven model behaves roughly at par with the state-space model.
- Adding Student t error terms to the score-driven model even increases the density fit beyond that of its state-space counterpart.

Empirical Study

- We apply our findings to model the dynamics of IV surfaces of S&P500 index options using data from January 2010 to December 2022.
- We find that a linear Gaussian state-space model outperforms a plain-vanilla score-driven model by a large margin, both in terms of density fit, and Value-at-Risk (VaR) violation rates.
- After adapting the score-driven model with the adjusted covariance structure for the measurement errors, the score-driven model behaves roughly at par with the state-space model.
- Adding Student t error terms to the score-driven model even increases the density fit beyond that of its state-space counterpart.

Empirical Study

- We apply our findings to model the dynamics of IV surfaces of S&P500 index options using data from January 2010 to December 2022.
- We find that a linear Gaussian state-space model outperforms a plain-vanilla score-driven model by a large margin, both in terms of density fit, and Value-at-Risk (VaR) violation rates.
- After adapting the score-driven model with the adjusted covariance structure for the measurement errors, the score-driven model behaves roughly at par with the state-space model.
- Adding Student t error terms to the score-driven model even increases the density fit beyond that of its state-space counterpart.

Empirical Study

- We apply our findings to model the dynamics of IV surfaces of S&P500 index options using data from January 2010 to December 2022.
- We find that a linear Gaussian state-space model outperforms a plain-vanilla score-driven model by a large margin, both in terms of density fit, and Value-at-Risk (VaR) violation rates.
- After adapting the score-driven model with the adjusted covariance structure for the measurement errors, the score-driven model behaves roughly at par with the state-space model.
- Adding Student t error terms to the score-driven model even increases the density fit beyond that of its state-space counterpart.

Table of Contents

1. Introduction
2. The Model
3. Empirical study

We model log implied volatilities $\boldsymbol{IV}_t \in \mathbb{R}^{q_t}$ for $t = 1, \ldots, T$, over a grid of moneyness values $\boldsymbol{m}_t \in \mathbb{R}^{K_t}$ and times-to-maturity $\boldsymbol{\tau}_t \in \mathbb{R}^{\mathcal{T}_T}$.

We assume that $\boldsymbol{IV}_t$ evolves as follows:

$$
\begin{aligned}
\boldsymbol{IV}_t &= \boldsymbol{M}_t \beta_t + \varepsilon_t, \qquad \varepsilon_t \sim h(\varepsilon_t \mid \boldsymbol{H}_t; \vartheta), \\
\beta_{t+1} &= (\mathbf{I} - \boldsymbol{B})\,\theta + \boldsymbol{B}\,\beta_t + \xi_t \qquad ,
\end{aligned}
$$

- We gather all static parameters of the model into a vector ψ that needs to be estimated.
- This set-up accommodates both a state-space and a score-driven framework, depending on our choice of ξ_t.

Standard state-space and score-driven models

We model log implied volatilities $\boldsymbol{IV}_t \in \mathbb{R}^{q_t}$ for $t = 1, \ldots, T$, over a grid of moneyness values $\boldsymbol{m}_t \in \mathbb{R}^{K_t}$ and times-to-maturity $\boldsymbol{\tau}_t \in \mathbb{R}^{\mathcal{T}_T}$.

We assume that $\boldsymbol{IV}_t$ evolves as follows:

$$\begin{aligned} \boldsymbol{IV}_t &= \boldsymbol{M}_t \boldsymbol{\beta}_t + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim h\left(\boldsymbol{\varepsilon}_t \mid \boldsymbol{H}_t; \boldsymbol{\vartheta}\right), \\ \boldsymbol{\beta}_{t+1} &= (\mathbf{I} - \boldsymbol{B})\, \boldsymbol{\theta} + \boldsymbol{B}\, \boldsymbol{\beta}_t + \boldsymbol{\xi}_t \qquad\qquad , \end{aligned}$$

- We gather all static parameters of the model into a vector ψ that needs to be estimated.
- This set-up accommodates both a state-space and a score-driven framework, depending on our choice of ξ_t.

Standard state-space and score-driven models

We model log implied volatilities $\boldsymbol{IV}_t \in \mathbb{R}^{q_t}$ for $t = 1, \ldots, T$, over a grid of moneyness values $\boldsymbol{m}_t \in \mathbb{R}^{K_t}$ and times-to-maturity $\boldsymbol{\tau}_t \in \mathbb{R}^{\mathcal{T}_T}$.

We assume that $\boldsymbol{IV}_t$ evolves as follows:

$$\boldsymbol{IV}_t = \boldsymbol{M}_t \boldsymbol{\beta}_t + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim h\left(\boldsymbol{\varepsilon}_t \mid \boldsymbol{H}_t; \boldsymbol{\vartheta}\right),$$
$$\boldsymbol{\beta}_{t+1} = (\mathbf{I} - \boldsymbol{B})\, \boldsymbol{\theta} + \boldsymbol{B}\, \boldsymbol{\beta}_t + \boldsymbol{\xi}_t \qquad ,$$

- We gather all static parameters of the model into a vector ψ that needs to be estimated.
- This set-up accommodates both a state-space and a score-driven framework, depending on our choice of ξ_t.

Standard state-space and score-driven models

We model log implied volatilities $\boldsymbol{IV}_t \in \mathbb{R}^{q_t}$ for $t = 1, \ldots, T$, over a grid of moneyness values $\boldsymbol{m}_t \in \mathbb{R}^{K_t}$ and times-to-maturity $\boldsymbol{\tau}_t \in \mathbb{R}^{\mathcal{T}_T}$.

We assume that $\boldsymbol{IV}_t$ evolves as follows:

$$\begin{aligned} \boldsymbol{IV}_t &= \boldsymbol{M}_t \beta_t + \varepsilon_t, \qquad \varepsilon_t \sim h(\varepsilon_t \mid \boldsymbol{H}_t; \vartheta), \\ \beta_{t+1} &= (\mathbf{I} - \boldsymbol{B})\, \boldsymbol{\theta} + \boldsymbol{B}\, \beta_t + \boldsymbol{\xi}_t \qquad , \end{aligned}$$

- We gather all static parameters of the model into a vector ψ that needs to be estimated.
- This set-up accommodates both a state-space and a score-driven framework, depending on our choice of $\boldsymbol{\xi}_t$.

Standard state-space and score-driven models

- Linear Gaussian state-space : $(\varepsilon_t^\top, \xi_t^\top)^\top$ is normally distributed.

Estimation

$$\ell(\psi) = -\frac{1}{2}\sum_{t=1}^{T}\left(\ln|2\pi \boldsymbol{F}_t| + \left(\boldsymbol{IV}_t - \boldsymbol{IV}_{t|t-1}\right)^\top \boldsymbol{F}_t^{-1}\left(\boldsymbol{IV}_t - \boldsymbol{IV}_{t|t-1}\right)\right).$$

- Score-driven framework: ξ_t is chosen as the derivative (with respect to β_t) of the log predictive density of $\boldsymbol{IV}_t$.

Estimation

$$\xi_t = \boldsymbol{A}\left(\boldsymbol{M}_t^\top \boldsymbol{H}_t^{-1}\boldsymbol{M}_t\right)^{-1}\boldsymbol{M}_t^\top \boldsymbol{H}_t^{-1}\varepsilon_t,$$

$$\ell(\psi) = -\frac{1}{2}\sum_{t=1}^{T}\left(\ln|2\pi \boldsymbol{H}_t| + \varepsilon_t^\top \boldsymbol{H}_t^{-1}\varepsilon_t\right),$$

Standard state-space and score-driven models

- Linear Gaussian state-space : $(\varepsilon_t^\top, \xi_t^\top)^\top$ is normally distributed.

Estimation

$$\ell(\psi) = -\frac{1}{2}\sum_{t=1}^{T}\left(\ln|2\pi \boldsymbol{F}_t| + \left(\boldsymbol{IV}_t - \boldsymbol{IV}_{t|t-1}\right)^\top \boldsymbol{F}_t^{-1}\left(\boldsymbol{IV}_t - \boldsymbol{IV}_{t|t-1}\right)\right).$$

- Score-driven framework: ξ_t is chosen as the derivative (with respect to β_t) of the log predictive density of $\boldsymbol{IV}_t$.

Estimation

$$\xi_t = \boldsymbol{A}\left(\boldsymbol{M}_t^\top \boldsymbol{H}_t^{-1}\boldsymbol{M}_t\right)^{-1}\boldsymbol{M}_t^\top \boldsymbol{H}_t^{-1}\varepsilon_t,$$
$$\ell(\psi) = -\frac{1}{2}\sum_{t=1}^{T}\left(\ln|2\pi \boldsymbol{H}_t| + \varepsilon_t^\top \boldsymbol{H}_t^{-1}\varepsilon_t\right),$$

Standard state-space and score-driven models

- Linear Gaussian state-space : $(\varepsilon_t^\top, \xi_t^\top)^\top$ is normally distributed.

Estimation

$$\ell(\psi) = -\tfrac{1}{2} \sum_{t=1}^{T} \left(\ln |2\pi \boldsymbol{F}_t| + \left(\boldsymbol{IV}_t - \boldsymbol{IV}_{t|t-1}\right)^\top \boldsymbol{F}_t^{-1} \left(\boldsymbol{IV}_t - \boldsymbol{IV}_{t|t-1}\right)\right).$$

- Score-driven framework: ξ_t is chosen as the derivative (with respect to β_t) of the log predictive density of $\boldsymbol{IV}_t$.

Estimation

$$\xi_t = \boldsymbol{A} \left(\boldsymbol{M}_t^\top \boldsymbol{H}_t^{-1} \boldsymbol{M}_t\right)^{-1} \boldsymbol{M}_t^\top \boldsymbol{H}_t^{-1} \varepsilon_t,$$

$$\ell(\psi) = -\tfrac{1}{2} \sum_{t=1}^{T} \left(\ln |2\pi \boldsymbol{H}_t| + \varepsilon_t^\top \boldsymbol{H}_t^{-1} \varepsilon_t\right),$$

Standard state-space and score-driven models

- Linear Gaussian state-space : $(\varepsilon_t^\top, \xi_t^\top)^\top$ is normally distributed.

Estimation

$$\ell(\psi) = -\frac{1}{2}\sum_{t=1}^{T}\left(\ln|2\pi \boldsymbol{F}_t| + \left(\boldsymbol{IV}_t - \boldsymbol{IV}_{t|t-1}\right)^\top \boldsymbol{F}_t^{-1}\left(\boldsymbol{IV}_t - \boldsymbol{IV}_{t|t-1}\right)\right).$$

- Score-driven framework: ξ_t is chosen as the derivative (with respect to β_t) of the log predictive density of $\boldsymbol{IV}_t$.

Estimation

$$\xi_t = \boldsymbol{A}\left(\boldsymbol{M}_t^\top \boldsymbol{H}_t^{-1}\boldsymbol{M}_t\right)^{-1}\boldsymbol{M}_t^\top \boldsymbol{H}_t^{-1}\varepsilon_t,$$
$$\ell(\psi) = -\frac{1}{2}\sum_{t=1}^{T}\left(\ln|2\pi \boldsymbol{H}_t| + \varepsilon_t^\top \boldsymbol{H}_t^{-1}\varepsilon_t\right),$$

Adjusted measurement equation of the score-driven model

- Consider a typical specification with diagonal error covariance matrix $\boldsymbol{H}_t$.
- Assume the DGP to be a state-space model and fit a score-driven model to it:

The true error:

$$\boldsymbol{IV}_t^{\text{ssf}} - \widehat{\boldsymbol{IV}}_t^{\text{sd}} = \varepsilon_t + \boldsymbol{M}_t \xi_t^{\text{ssf}} + \boldsymbol{M}_t \left((\mathbf{I} - \boldsymbol{B}) \left(\theta - \hat{\theta} \right) + \boldsymbol{B} \left(\beta_{t-1}^{\text{ssf}} - \widehat{\beta^{\text{sd}}}_{t-1} \right) \right),$$

where β_t^{sd} is the β_t from the score-driven model, and β_t^{ssf} is that from the state-space.

- The first component has an uncorrelated covariance structure.
- The final term is typically small given the good forecast performance of score-driven models even for a state-space DGP. (see Koopman et al., 2016).
- It is the second term that results in a cross-correlated prediction errors, which contrasts with the assumed uncorrelated structure.

Adjusted measurement equation of the score-driven model

- Consider a typical specification with diagonal error covariance matrix $\boldsymbol{H}_t$.
- Assume the DGP to be a state-space model and fit a score-driven model to it:

The true error:

$$\boldsymbol{IV}_t^{\text{ssf}} - \widehat{\boldsymbol{IV}}_t^{\text{sd}} = \varepsilon_t + \boldsymbol{M}_t \, \xi_t^{\text{ssf}} + \boldsymbol{M}_t \left((\mathbf{I} - \boldsymbol{B}) \left(\theta - \hat{\theta} \right) + \boldsymbol{B} \left(\beta_{t-1}^{\text{ssf}} - \widehat{\beta^{\text{sd}}}_{t-1} \right) \right),$$

where β_t^{sd} is the β_t from the score-driven model, and β_t^{ssf} is that from the state-space.

- The first component has an uncorrelated covariance structure.
- The final term is typically small given the good forecast performance of score-driven models even for a state-space DGP. (see Koopman et al., 2016).
- It is the second term that results in a cross-correlated prediction errors, which contrasts with the assumed uncorrelated structure.

Adjusted measurement equation of the score-driven model

- Consider a typical specification with diagonal error covariance matrix $\boldsymbol{H}_t$.
- Assume the DGP to be a state-space model and fit a score-driven model to it:

The true error:

$$\boldsymbol{IV}_t^{\text{ssf}} - \widehat{\boldsymbol{IV}}_t^{\text{sd}} = \varepsilon_t + \boldsymbol{M}_t \xi_t^{\text{ssf}} + \boldsymbol{M}_t \left((\mathbf{I} - \boldsymbol{B}) \left(\theta - \hat{\theta} \right) + \boldsymbol{B} \left(\beta_{t-1}^{\text{ssf}} - \widehat{\beta^{\text{sd}}}_{t-1} \right) \right),$$

where β_t^{sd} is the β_t from the score-driven model, and β_t^{ssf} is that from the state-space.

- The first component has an uncorrelated covariance structure.
- The final term is typically small given the good forecast performance of score-driven models even for a state-space DGP. (see Koopman et al., 2016).
- It is the second term that results in a cross-correlated prediction errors, which contrasts with the assumed uncorrelated structure.

Adjusted measurement equation of the score-driven model

- Consider a typical specification with diagonal error covariance matrix $\boldsymbol{H}_t$.
- Assume the DGP to be a state-space model and fit a score-driven model to it:

The true error:

$$\boldsymbol{IV}_t^{\text{ssf}} - \widehat{\boldsymbol{IV}}_t^{\text{sd}} = \varepsilon_t + \boldsymbol{M}_t \, \xi_t^{\text{ssf}} + \boldsymbol{M}_t \left((\mathbf{I} - \boldsymbol{B}) \left(\theta - \hat{\theta} \right) + \boldsymbol{B} \left(\beta_{t-1}^{\text{ssf}} - \widehat{\beta^{\text{sd}}}_{t-1} \right) \right),$$

where β_t^{sd} is the β_t from the score-driven model, and β_t^{ssf} is that from the state-space.

- The first component has an uncorrelated covariance structure.
- The final term is typically small given the good forecast performance of score-driven models even for a state-space DGP. (see Koopman et al., 2016).
- It is the second term that results in a cross-correlated prediction errors, which contrasts with the assumed uncorrelated structure.

Adjusted measurement equation of the score-driven model

The **solution**: include one iteration of the state-space dynamics into the score-driven measurement equation.

Adjusted measurement equation

$$\boldsymbol{IV}_t = \boldsymbol{M}_t\beta_t + \varepsilon_t, \qquad \varepsilon_t \sim h\left(\varepsilon_t \,\middle|\, \boldsymbol{H}_t + \boldsymbol{M}_t\boldsymbol{C}\boldsymbol{M}_t^\top, \vartheta\right),$$

- In our single level factor example above, the adjusted measurement equation induces an equicorrelation structure.
- This explains the huge improvements in density fit that Koopman et al. (2017) find when imposing equicorrelation structures.
- Our approach above can, however, also be used for richer factor structures.

Adjusted measurement equation of the score-driven model

The **solution**: include one iteration of the state-space dynamics into the score-driven measurement equation.

Adjusted measurement equation

$$\boldsymbol{IV}_t = \boldsymbol{M}_t \beta_t + \varepsilon_t, \qquad \varepsilon_t \sim h\left(\varepsilon_t \;\middle|\; \boldsymbol{H}_t + \boldsymbol{M}_t \boldsymbol{C} \boldsymbol{M}_t^\top, \vartheta\right),$$

- In our single level factor example above, the adjusted measurement equation induces an equicorrelation structure.
- This explains the huge improvements in density fit that Koopman et al. (2017) find when imposing equicorrelation structures.
- Our approach above can, however, also be used for richer factor structures.

Adjusted measurement equation of the score-driven model

The **solution**: include one iteration of the state-space dynamics into the score-driven measurement equation.

Adjusted measurement equation

$$\boldsymbol{IV}_t = \boldsymbol{M}_t \beta_t + \varepsilon_t, \qquad \varepsilon_t \sim h\left(\varepsilon_t \,\middle|\, \boldsymbol{H}_t + \boldsymbol{M}_t \boldsymbol{C} \boldsymbol{M}_t^\top, \vartheta\right),$$

- In our single level factor example above, the adjusted measurement equation induces an equicorrelation structure.
- This explains the huge improvements in density fit that Koopman et al. (2017) find when imposing equicorrelation structures.
- Our approach above can, however, also be used for richer factor structures.

Adjusted measurement equation of the score-driven model

The **solution**: include one iteration of the state-space dynamics into the score-driven measurement equation.

Adjusted measurement equation

$$\boldsymbol{IV}_t = \boldsymbol{M}_t \beta_t + \varepsilon_t, \qquad \varepsilon_t \sim h\left(\varepsilon_t \;\middle|\; \boldsymbol{H}_t + \boldsymbol{M}_t \boldsymbol{C} \boldsymbol{M}_t^\top, \vartheta\right),$$

- In our single level factor example above, the adjusted measurement equation induces an equicorrelation structure.
- This explains the huge improvements in density fit that Koopman et al. (2017) find when imposing equicorrelation structures.
- Our approach above can, however, also be used for richer factor structures.

Adjusted measurement equation of the score-driven model

The **solution**: include one iteration of the state-space dynamics into the score-driven measurement equation.

Adjusted measurement equation

$$\boldsymbol{IV}_t = \boldsymbol{M}_t \beta_t + \varepsilon_t, \qquad \varepsilon_t \sim h\left(\varepsilon_t \;\middle|\; \boldsymbol{H}_t + \boldsymbol{M}_t \boldsymbol{C} \boldsymbol{M}_t^\top, \vartheta\right),$$

- In our single level factor example above, the adjusted measurement equation induces an equicorrelation structure.
- This explains the huge improvements in density fit that Koopman et al. (2017) find when imposing equicorrelation structures.
- Our approach above can, however, also be used for richer factor structures.

Table of Contents

1 Introduction

2 The Model

3 Empirical study

Empirical modeling choices

- Practitioners have long tried to fit linear parametric models to the cross-section of implied volatility at a point in time, linking IV to time-to-maturity and moneyness, e.g. Andreou et al. (2010), Dumas et al. (1998), Goncalves and Guidolin (2006), and Pena et al. (1999).
- Goncalves and Guidolin (2006): simple model linear in coefficients and nonlinear in moneyness and time-to-maturity achieves a good fit to the IV surface of S&P 500 index options.
- To illustrate our main findings, we also employ the simple 5 factor models of Goncalves and Guidolin (2006) to the IV surface of S&P 500 index options.
- For robustness checks, we also consider a more flexible factor presentation that includes a nonparametric factor.

Empirical modeling choices

- Practitioners have long tried to fit linear parametric models to the cross-section of implied volatility at a point in time, linking IV to time-to-maturity and moneyness, e.g. Andreou et al. (2010), Dumas et al. (1998), Goncalves and Guidolin (2006), and Pena et al. (1999).
- Goncalves and Guidolin (2006): simple model linear in coefficients and nonlinear in moneyness and time-to-maturity achieves a good fit to the IV surface of S&P 500 index options.
- To illustrate our main findings, we also employ the simple 5 factor models of Goncalves and Guidolin (2006) to the IV surface of S&P 500 index options.
- For robustness checks, we also consider a more flexible factor presentation that includes a nonparametric factor.

Empirical modeling choices

- Practitioners have long tried to fit linear parametric models to the cross-section of implied volatility at a point in time, linking IV to time-to-maturity and moneyness, e.g. Andreou et al. (2010), Dumas et al. (1998), Goncalves and Guidolin (2006), and Pena et al. (1999).
- Goncalves and Guidolin (2006): simple model linear in coefficients and nonlinear in moneyness and time-to-maturity achieves a good fit to the IV surface of S&P 500 index options.
- To illustrate our main findings, we also employ the simple 5 factor models of Goncalves and Guidolin (2006) to the IV surface of S&P 500 index options.
- For robustness checks, we also consider a more flexible factor presentation that includes a nonparametric factor.

Empirical modeling choices

- Practitioners have long tried to fit linear parametric models to the cross-section of implied volatility at a point in time, linking IV to time-to-maturity and moneyness, e.g. Andreou et al. (2010), Dumas et al. (1998), Goncalves and Guidolin (2006), and Pena et al. (1999).
- Goncalves and Guidolin (2006): simple model linear in coefficients and nonlinear in moneyness and time-to-maturity achieves a good fit to the IV surface of S&P 500 index options.
- To illustrate our main findings, we also employ the simple 5 factor models of Goncalves and Guidolin (2006) to the IV surface of S&P 500 index options.
- For robustness checks, we also consider a more flexible factor presentation that includes a nonparametric factor.

Restricted factor specification

Parametric factor specification

$$\boldsymbol{IV}_t(\boldsymbol{m},\boldsymbol{\tau}) = \boldsymbol{M}_t\beta_t + \varepsilon_t,$$

- $\boldsymbol{IV}_t(\boldsymbol{m},\boldsymbol{\tau})$ is a q_t-dimensional vector
- $\boldsymbol{M}_t = (\boldsymbol{m}_1, \boldsymbol{m}_2, \ldots, \boldsymbol{m}_{q_t})$ is a matrix $(q_t \times 5)$ with $\boldsymbol{m}_j = (1, m_j, m_j^2, \tau_j, m_j\tau_j)'$
- β_t is vector of latent factors (5×1),
- ε_t is the disturbance $(q_t \times 1)$,
- q_t is the number of option contracts at time t.

Data

- We use a dataset comprising European call options on the S&P 500 index, encompassing all call and put options traded on the Chicago Board Option Exchange (CBOE).
- We follow van der Wel et al. (2016) to filter option contracts.
- We restrict our analysis to out-of-the-money options, defined by a Δ less than 0.5 in absolute value.
- We exclude observations characterized by time-to-maturity periods exceeding 360 days or shorter than 7 days.
- Options with implied volatilities greater than 0.7 and option prices below 0.05 are omitted from the dataset to mitigate the effect of potential data errors.

Data

- We use a dataset comprising European call options on the S&P 500 index, encompassing all call and put options traded on the Chicago Board Option Exchange (CBOE).
- We follow van der Wel et al. (2016) to filter option contracts.
- We restrict our analysis to out-of-the-money options, defined by a Δ less than 0.5 in absolute value.
- We exclude observations characterized by time-to-maturity periods exceeding 360 days or shorter than 7 days.
- Options with implied volatilities greater than 0.7 and option prices below 0.05 are omitted from the dataset to mitigate the effect of potential data errors.

Data

- We use a dataset comprising European call options on the S&P 500 index, encompassing all call and put options traded on the Chicago Board Option Exchange (CBOE).
- We follow van der Wel et al. (2016) to filter option contracts.
- We restrict our analysis to out-of-the-money options, defined by a Δ less than 0.5 in absolute value.
- We exclude observations characterized by time-to-maturity periods exceeding 360 days or shorter than 7 days.
- Options with implied volatilities greater than 0.7 and option prices below 0.05 are omitted from the dataset to mitigate the effect of potential data errors.

Data

- We use a dataset comprising European call options on the S&P 500 index, encompassing all call and put options traded on the Chicago Board Option Exchange (CBOE).
- We follow van der Wel et al. (2016) to filter option contracts.
- We restrict our analysis to out-of-the-money options, defined by a Δ less than 0.5 in absolute value.
- We exclude observations characterized by time-to-maturity periods exceeding 360 days or shorter than 7 days.
- Options with implied volatilities greater than 0.7 and option prices below 0.05 are omitted from the dataset to mitigate the effect of potential data errors.

Data

- We use a dataset comprising European call options on the S&P 500 index, encompassing all call and put options traded on the Chicago Board Option Exchange (CBOE).
- We follow van der Wel et al. (2016) to filter option contracts.
- We restrict our analysis to out-of-the-money options, defined by a Δ less than 0.5 in absolute value.
- We exclude observations characterized by time-to-maturity periods exceeding 360 days or shorter than 7 days.
- Options with implied volatilities greater than 0.7 and option prices below 0.05 are omitted from the dataset to mitigate the effect of potential data errors.

Empirical results

- We compare the performance of the linear Gaussian state-space and socre-driven models in forecasting out-of-sample IV, both statistically and economically.
- For the score-driven models, we consider four variants: normal and student-t distributions without and with adjusted measurements, of which we denote as GAS($\mathcal{N}$),GAS($\Sigma, \mathcal{N}$), GAS(t), and GAS(Σ, t) respectively.

Empirical results

- We compare the performance of the linear Gaussian state-space and socre-driven models in forecasting out-of-sample IV, both statistically and economically.
- For the score-driven models, we consider four variants: normal and student-t distributions without and with adjusted measurements, of which we denote as GAS($\mathcal{N}$),GAS($\boldsymbol{\Sigma}, \mathcal{N}$), GAS($t$), and GAS($\boldsymbol{\Sigma}, t$) respectively.

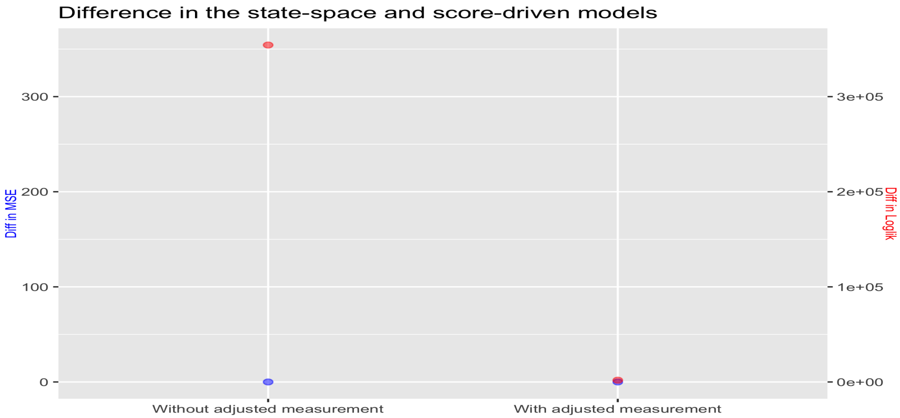
Empirical results

Table 1: STATISTICAL ACCURACY MEASURES FOR VARIOUS ESTIMATORS OF THE IMPLIED VOLATILITY SURFACE: OUT-OF-SAMPLE

Model	MSE ($\times 10^3$)	MAE ($\times 10^3$)	loglik ($\times 10^{-3}$)	AIC ($\times 10^{-3}$)	# Pars
Whole sample: 2012-2022					
State Space	2.83	33.21	1557.45	-3114.86	16
GAS($\mathcal{N}$)	2.81	32.90***	1203.19	-2406.35	16
GAS($\Sigma, \mathcal{N}$)	2.88**	33.99***	1555.56	-3111.08	21
GAS(t)	2.86	33.36	1551.76	-3103.49	17
GAS(Σ, t)	2.97***	34.59	1926.21	-3852.38	22
Static Model	6.00***	53.37***			

Note: Diebold and Mariano test is conducted with the state space model as a benchmark on MSE and MAE, and the null hypothesis is that both models are equally accurate.

Empirical results



Empirical results: VaR applications

We compare the VaR forecasts for the equally-weighted average implied volatility of S&P 500 options.

Table 2: Out-of-sample evaluation for 99% Value-at-Risk Estimation

	Violation ratio ($\times 10^2$)	Tickloss ($\times 10^3$)
The whole period: 2012-2022		
State Space	0.04	3.68
GAS($\mathcal{N}$)	41.07	9.99***
GAS(Σ,$\mathcal{N}$)	0.04	5.16***
GAS(t)	50.88	14.60***
GAS(Σ, t)	0.00	6.14***

Empirical results: VaR applications

We compare the VaR forecasts for the equally-weighted average implied volatility of S&P 500 options.

Table 2: OUT-OF-SAMPLE EVALUATION FOR 99% VALUE-AT-RISK ESTIMATION

	Violation ratio ($\times 10^2$)	Tickloss ($\times 10^3$)
The whole period: 2012-2022		
State Space	0.04	3.68
GAS($\mathcal{N}$)	41.07	9.99***
GAS($\Sigma, \mathcal{N}$)	0.04	5.16***
GAS(t)	50.88	14.60***
GAS(Σ, t)	0.00	6.14***

Conclusion

- We compare state-space and score-driven models for option implied volatility surface dynamics.
- We find that point forecasts of both models behave similarly, but density forecasts of the plain-vanilla score-driven model are substantially worse.
- We show how a simple adjustment of the measurement density of the score-driven model can put both models back on an equal footing.
- After this correction, the score-driven model can easily be extended with non-Gaussian features without complicating parameter estimation, unlike its state-space counterpart.

Conclusion

- We compare state-space and score-driven models for option implied volatility surface dynamics.
- We find that point forecasts of both models behave similarly, but density forecasts of the plain-vanilla score-driven model are substantially worse.
- We show how a simple adjustment of the measurement density of the score-driven model can put both models back on an equal footing.
- After this correction, the score-driven model can easily be extended with non-Gaussian features without complicating parameter estimation, unlike its state-space counterpart.

Conclusion

- We compare state-space and score-driven models for option implied volatility surface dynamics.
- We find that point forecasts of both models behave similarly, but density forecasts of the plain-vanilla score-driven model are substantially worse.
- We show how a simple adjustment of the measurement density of the score-driven model can put both models back on an equal footing.
- After this correction, the score-driven model can easily be extended with non-Gaussian features without complicating parameter estimation, unlike its state-space counterpart.

Andreou, Panayiotis C, Chris Charalambous, and Spiros H Martzoukos (2010). "Generalized parameter functions for option pricing". In: *Journal of Banking & Finance* 34.3, pp. 633–646.

Dumas, Bernard, Jeff Fleming, and Robert E Whaley (1998). "Implied volatility functions: Empirical tests". In: *Journal of Finance* 53.6, pp. 2059–2106.

Goncalves, Silvia and Massimo Guidolin (2006). "Predictable dynamics in the S&P 500 index options implied volatility surface". In: *Journal of Business* 79.3, pp. 1591–1635.

Koopman, Siem Jan, André Lucas, and Marcel Scharth (2016). "Predicting time-varying parameters with parameter-driven and observation-driven models". In: *Review of Economics and Statistics* 98.1, pp. 97–110.

Koopman, Siem Jan, André Lucas, and Marcin Zamojski (2017). "Dynamic term structure models with score-driven time-varying parameters: estimation and forecasting". In: *Narowdy Bank Polski, NBP Working Paper* No. 258.

Pena, Ignacio, Gonzalo Rubio, and Gregorio Serna (1999). "Why do we smile? On the determinants of the implied volatility function". In: *Journal of Banking & Finance* 23.8, pp. 1151–1179.

van der Wel, Michel, Sait R Ozturk, and Dick van Dijk (2016). "Dynamic Factor Models for the Volatility Surface". In: *Dynamic Factor Models*. Vol. 35. Emerald Group Publishing Limited, pp. 127–174.