On the Pricing of Capped Volatility Swaps using Machine Learning Techniques

23 August 2022

Eva Verschueren

EAJ 2022, Tartu

Joint work with Dr. Stephan Höcht (Assenagon) & Prof. Wim Schoutens (KU Leuven)

Outline of the presentation

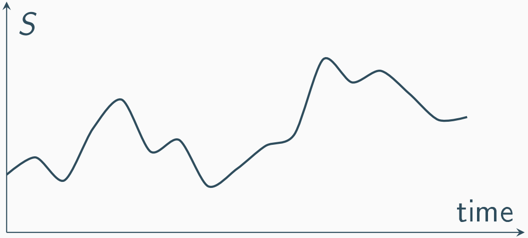
1. Introduction
2. Data
3. Modeling
4. Preliminary Results
5. Conclusion and Next Steps

Introduction

Definition

What is a capped volatility swap?

In general, a capped volatility swap is a financial **derivative contract** of which the payout depends on the performance of an underlying asset S.

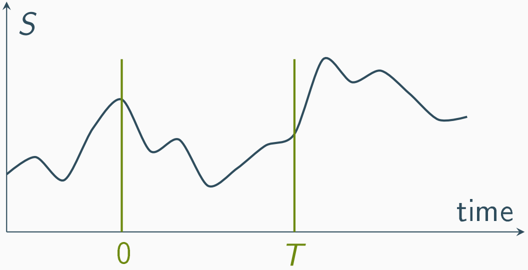


Definition

What is a capped volatility swap?

In general, a capped volatility swap is a financial **derivative contract** of which the payout depends on the performance of an underlying asset S.

More specifically, the payoff is based upon the annualized, **realized volatility** σ_R of S, over the lifetime of the contract, a period of length T.



$$\sigma_R = 100 \times \sqrt{\frac{252}{T} \sum_{t=1}^{T} \left(\ln \left[\frac{S_t}{S_{t-1}} \right] \right)^2}$$

Definition

What is a capped volatility swap?

A capped volatility swap is a two-party contract wherein both parties agree to the **exchange of cash flows** at the end of the given period of length T. It is also called a **forward** contract.

 Issuer $\longrightarrow$ $\longleftarrow$  Buyer

pays σ_R, to the buyer, at time T.

pays K, to the issuer, at time T.

Definition

What is a capped volatility swap?

A capped volatility swap is a two-party contract wherein both parties agree to the **exchange of cash flows** at the end of the given period of length T. It is also called a **forward** contract.

pays σ_R, to the buyer, at time T.

pays K, to the issuer, at time T.
K is the fair price of the contract, determined at time 0.

Definition

What is a capped volatility swap?

A capped volatility swap is a two-party contract wherein both parties agree to the **exchange of cash flows** at the end of the given period of length T. It is also called a **forward** contract.

pays σ_R, to the buyer, at time T.
To limit the risk exposure, σ_R is often capped at $2.5 \times K$.

pays K, to the issuer, at time T.
K is the fair price of the contract, determined at time 0.

Definition

> A **capped volatility swap** is a forward contract on an asset's annualized, realized volatility, over a fixed period of length T,[1]

with payoff structure

$$\text{Notional} \times [\min(\text{Cap Level}, \sigma_R) - K].$$

[1] See e.g., Demeterfi et al. (1999).

Definition

> A **capped volatility swap** is a forward contract on an asset's annualized, realized volatility, over a fixed period of length T,[1]

with payoff structure

$$\text{Notional} \times [\min(\text{Cap Level}, \sigma_R) - K].$$

⟶ Direct and **pure exposure to** the **volatility** of the underlying asset.

⟶ Used for both hedging and speculative purposes.

⟶ **Increasingly popular** tool in the financial industry.

[1] See e.g., Demeterfi et al. (1999).

Problem statement

> Volatility swaps are **traded over-the-counter**, meaning that no price is readily available on exchange.

? What is the **current price** of a specific contract?

! External entities, participating in the vol swap market, provide prices based on internal models.

⚠ Occasionally, **prices** from different pricing sources **differ** substantially.

How to tackle this pricing problem?

In general,

$$\text{Price} = \text{DF} \times \mathbb{E}(\text{Payoff}),$$

with discount factor DF and expectations taken under a pricing measure.

More specifically,

$$\boxed{\text{Price} = \text{DF} \times \mathbb{E}\big(\text{Notional} \times [\min(\text{Cap Level},\ \sigma_R) - K]\big).}$$

Not an easy problem:

- nonlinear due to cap level and square root operator
- $\mathbb{E}\left[\sqrt{X}\right] \leq \sqrt{\mathbb{E}[X]}$

How to tackle this pricing problem?

Approach 1: Current literature is focused on a **model-based** pricing approach.

- Under Black-Scholes, see e.g., Rujivan and Rakwongwan (2021)
- Under Heston, see e.g., Issaka (2020)
- Under SABR, see e.g., Kim and Kim (2020)

Disadvantages

- Many assumptions are needed on the underlying asset price process.
- Due to nonlinearity of the payoff, no exact formula can be derived.
- Literature is highly focused on the price at time 0.

How to tackle this pricing problem?

Approach 2: We focus on a **data-driven** approach.

Recently, **machine learning** provides new tools to solve challenges in finance. See e.g., De Spiegeleer et al. (2018), Gan et al. (2020) and Davis et al. (2020) for applications to the pricing of financial products.

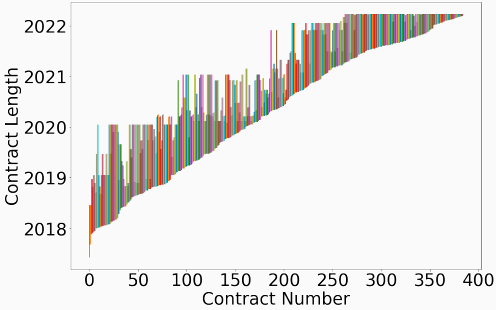
We deploy a model-free, data-driven approach to price capped volatility swaps, based on machine learning techniques.

Data

General overview

The data consists of

- 387 volatility swap contracts,
- with as underlying the **S&P 500 index**,
- spanning a time period from June 7, **2017** till March 23, **2022**.



The **price** of an individual swap is observed on **every business** day, during the lifetime of the contract, until settlement or maturity, resulting in 75066 data points.

Response variable - Price

The final payoff of a capped volatility swap is determined by the **realized volatility** of the underlying asset.

We thus rewrite the **price** of a contract, at time t, using the fact that it reflects both the already **realized** $(0\text{-}t)$ and **unrealized** $(t\text{-}T)$ part of the contract:

$$\boxed{\text{Price}_t = \text{DF}_t \times \left(\sqrt{\text{IVOL}_t^2 \times (1 - \text{Weight}_t) + \text{Accrued Vol}_t^2 \times \text{Weight}_t} - \text{Strike} \right)}$$

- Weight = proportion of the current lifetime of the swap
- Accrued Vol = already realized historical volatility

Response variable - IVOL

We model the summarizing parameter **IVOL**.

(1) reflects the expected **volatility** over the period between t and T,

(2) reflects how likely the **cap** is reached in the period between t and T.

Response variable - IVOL

We model the summarizing parameter **IVOL**.

(1) reflects the expected **volatility** over the period between t and T,

(2) reflects how likely the **cap** is reached in the period between t and T.

$\longrightarrow$ completely determined by the movements of the underlying asset S

$\longrightarrow$ **distributional features**[2] of S might be **predictive** for IVOL

- Implied Volatility and Implied Skewness (1) and Implied Kurtosis (2)
- Implied Moments, in general

[2] See Madan and Schoutens (2016) for a model-free calculation method.

Predictive variables

Contract Specific Features	Model 1	Model 2	Model 3
Accrued Volatility	✓	✓	✓
Weight	✓	✓	✓
ITM*	✓	✓	✓
Strike	✓	✓	✓
Distributional Features			
Implied Vol, Skew, Kurt		30d, RTM**	30d, RTM
Implied 1st-4th Moment			30d, RTM

*ITM=Initial Time till Maturity

**RTM=Remaining Time till Maturity

Modeling

Model candidates

Tree-based machine learning

Bagging[3]

- collection of T deep regression trees, built on **bootstrapped** samples

Random Forest[4]

- improvement over bagging
- random selection of m split candidates from the full set of candidates to **decorrelate** the trees

Gradient Boosting Machine[5]

- collection of T simple trees, all fit **sequentially**
- residuals used as response to improve the current fit

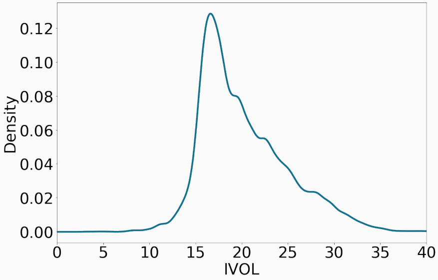
See [3]Breiman (1996), [4]Breiman (2001) and [5]Friedman (2001).

Loss function

The results shown are based on the **mean squared error** of prediction as a loss function, when training the model,

$$\mathscr{L}(\boldsymbol{y}, \widehat{f}(\boldsymbol{x})) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{f}(\boldsymbol{x}_i)\right)^2.$$

- The **normal assumption** for the response variable is however clearly **violated**.
- In one of the next steps, proper loss functions will be investigated.



Tuning via Purged, Walk-forward Validation

To prevent overfitting, k-**fold cross-validation** (CV) is often used to determine the generalization error of a machine learning algorithm.
$\longrightarrow$ Standard k-fold CV **fails** in Finance.
$\longrightarrow$ Observations cannot be assumed to be drawn from an IID process.

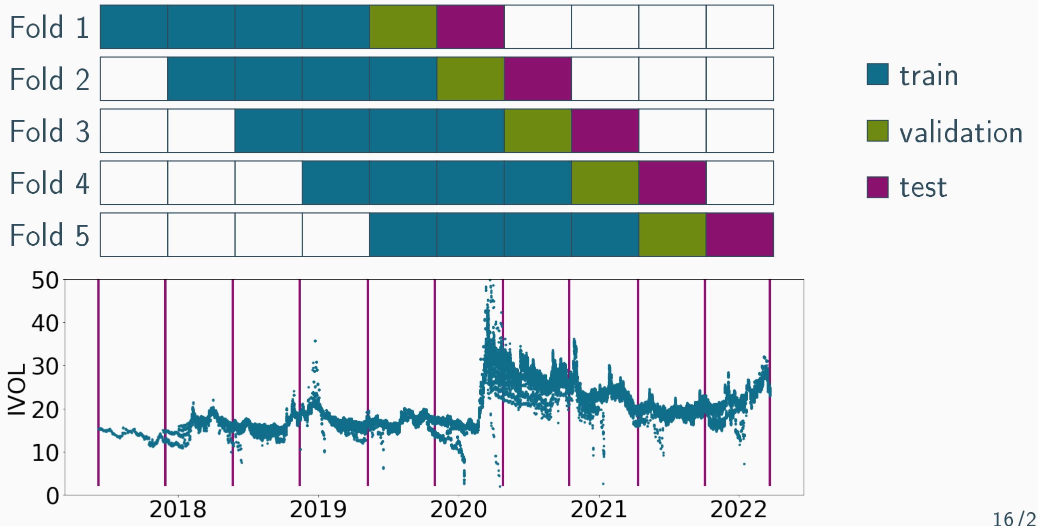
We use **purged walk-forward validation**.[6]

- Purging = all observations overlapping in time are included in the same set.
- Walk-forward = train-validation-test window rolls forward in time.

$\longrightarrow$ Information leakage between training and validation/test set is limited.

[6] See de Prado (2018).

Tuning via Purged, Walk-forward Validation



Preliminary Results

Model evaluation

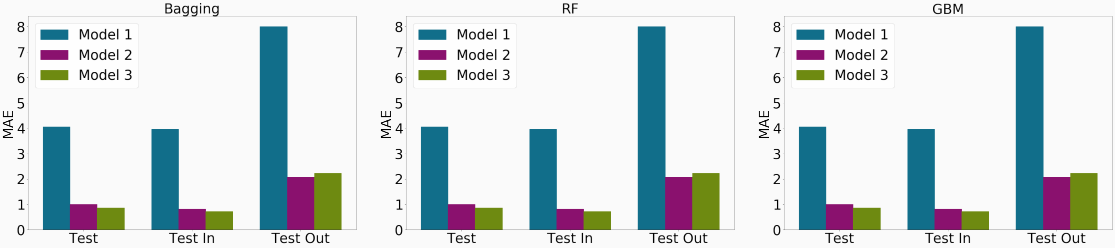
For interpretation, the models are evaluated using the **Mean Absolute Error** (MAE) of prediction, given by

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \widehat{f}(\boldsymbol{x}_i)\right|.$$

Since most data-driven machine learning approaches are **not capable** of performing reliable **extrapolations**, we make a distinction between

- Test: MAE on the total test set,
- Test In: MAE on part of the test set with features within the training boundaries,
- Test Out: MAE on part of the test set with features outside the training boundaries.

First results



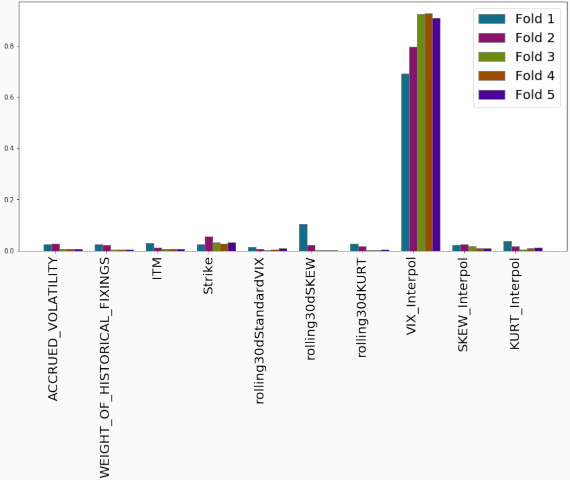
- **GBM** is the best performing technique, with minor differences between the tree-based methods.
- Including **distributional features** significantly increases the predictive accuracy.
- The increase in accuracy when including implied moments is minor.

Feature importance - Model 2

⚠ Highly correlated features!

- 30d Skew vs. 30d Kurtosis >0.9
- Interpolated Skew vs. Interpolated Kurtosis >0.9

- Multicollinearity does not affect prediction accuracy.
- Multicollinearity does affect feature importance estimates.



MAE when only using implied vol $= 2.81$.

Conclusion and Next Steps

Conclusion

The pricing of **capped volatility swaps**, at any time, during the lifetime of the product, is an unsolved problem.

(1) While current literature is mainly focused on model-based approximations of the price,

(2) we develop a **model-free, data-driven pricing** approach.

- This presentation deploys the use of **tree-based machine learning** techniques.
- First results show that the inclusion of **distributional features** results in a significant increase of predictive accuracy.

Next steps

Multiple questions are still to be answered:

- What proper **loss function** can be used to replace the mean squared error when training the model?
- What is the performance of **other ML techniques**, e.g. based on neural networks?
- How can we draw conclusions about feature importance when dealing with correlated features? What features to in/exclude?
- What is the performance of the model when tested on contracts with a **less liquid** underlying **asset**?

Thank you!

eva.verschueren@kuleuven.be

References

Breiman, L. (1996). Bagging predictors. Machine Learning, 24:123-140.

Breiman, L. (2001). Random forests. Machine learning, 45(1):5–32.

Davis, J., Devos, L., Reyners, S. and Schoutens, W. (2020). Gradient Boosting for Quantitative Finance. Journal of Computational Finance, 24(4).

Demeterfi, K., Derman, E., Kamal, M. and Zou, J. (1999). More Than You Ever Wanted To Know About Volatility Swaps. Quantitative Strategies Research Notes, Goldman Sachs, March, 1–52.

De Spiegeleer, J., Madan, D., Reyners, S. and Schoutens, W. (2018). Machine learning for quantitative finance: fast derivative pricing, hedging and fitting. Quantitative Finance, 18(10):1635-1643.

Friedman, J.H. (2001). Greedy function approximation: a gradient boosting machine. Annals of statistics, 29(5):1189–1232.

References

Gan, L., Wang, H. and Yang, Z. (2020). Machine learning solutions to challenges in finance: An application to the pricing of financial products. Technological Forecasting & Social change, 153.

Madan, D. and Schoutens, W. (2016). Applied Conic Finance. Cambridge University Press: Cambridge, UK.

Rujivan, S. and Rakwongwan, U. (2021). Analytically pricing volatility swaps and volatility options with discrete sampling: Nonlinear payoffvolatility derivatives. Communications in Nonlinear Science and Numerical Simulation, September, 100.

Kim, S-W. and Kim, J-H (2020). Volatility and variance swaps and options in the fractional SABR model. The European Journal of Finance, 26(17):1725-1745.

Issaka, A. (2020). Variance swaps, volatility swaps, hedging and bounds under multi-factor Heston stochastic volatility model. Stochastic Analysis and Applications, 38(5):856-874.

Additional information

Correlation Matrix Model 2