VAR - DSFM Modeling for Implied Volatility String Dynamics

Ralf Brüggemann
Wolfgang Härdle
Julius Mungo
Carsten Trenkler

CASE-Center for Applied Statistics and Economics
Institut für Statistik and Ökonometrie
Humboldt-Universität zu Berlin

Aims

Dynamic Semiparametric Factor Models (DSFM) yield time dependent factor loadings

In the context of Implied Volatility (IV) Dynamics, risk factors

- explaining the nature of volatility risk
- allowing to hedge positions of 'volatility derivatives'
- characterizing and quantifying risk in relation to economic indicators e.g. interest rates, oil prices etc.

Modeling the dynamics of these factors is important for accurate assessment of market risk.

Challenges

- Large number of observations (> 6 million contracts, $> 5\,000$ observations per day).
- Data appear in ‘strings’.
- Strings are not locally fixed, but ‘move’ through the observation space (expiry effect).
- In the moneyness dimension observations may be missing in certain sub-regions for some dates
- Standard smoothing techniques are necessarily **biased**.

Challenges

- How to model dynamics of this high-dimensional string object?
- Volatility characteristics, a strong day to day variation and a number of volatility clusters
- Structural breaks in series with sudden downward movements e.g. for z_{t1} in September 2001
- Influence of possible outliers e.g. for z_{t2} in November 2001

Overview

1. Motivation ✓
2. Literature review on factor times series modeling
3. Factor loadings series from DSFM
4. Integration analysis and unit root tests
5. VAR modeling and dynamic interaction between factors
6. Results
7. Outlook

Literature review

Recent research towards analyzing the behavior of the IVS:

- [Skiadopoulos et al. (1999)] analyzed the IVS of S&P 500 and reported that at least two and at most six factors are necessary to capture the dynamics
- [Cont and Fonseca (2002)], on dynamics of the S&P 500 implied volatility reported that the first three principal components account for 95% of the daily variance.

Literature review cont.

- [Fengler et al. (2003)] indicated three factors are sufficient to capture 95% variation in DAX implied volatilities.
- [R. Hafner (2004)] with a parametric approach, uses a four-factor model for DAX implied volatilities.
- [Borak, Härdle and Fengler (2005)] identified three loading series $z_t = (z_{t1}, z_{t2}, z_{t3})^\top$, after fitting a DSFM.

An Implied Volatility Surface

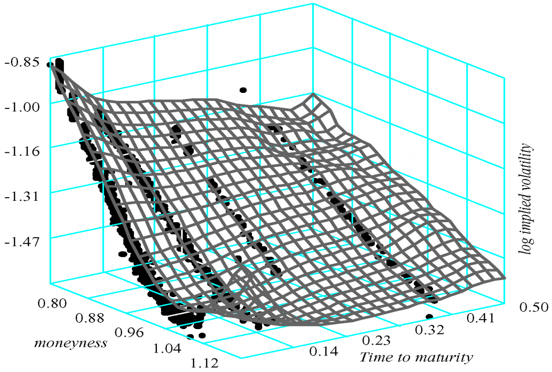

Figure 1: Implied volatility surface from DSFM fit for the DAX-Option on 20000502 (2 May 2000)

The semiparametric factor model

$$Y_{t,j} = \sum_{k=0}^{K} z_{tk} m_k(X_{t,j}) + \varepsilon_{t,j} \qquad (1)$$

where $z_{t0} = 1, \quad j = 1, \ldots, J_t \ (t = 1, \ldots, T)$ represents the number of IV observations on day t and K is the number of basis functions.

$X_{t,j}$ are the exogenous variables like strike and maturity.

z_{tk} are time dependent factors or weights of the smooth basis function m_k, for $(k = 0, \ldots, K)$.

[Borak, Härdle and Fengler (2005)]

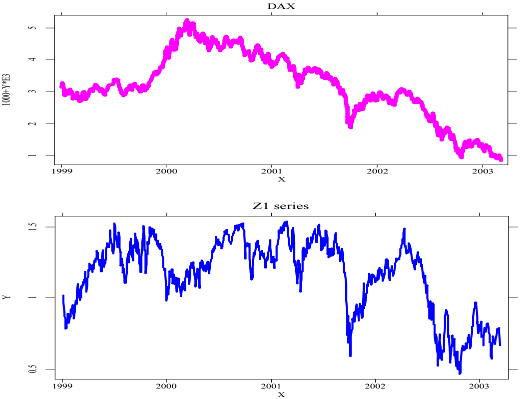

Figure 2: Upper panel: Time series of the Underlying DAX. Lower panel: Time series of first factor, Z_{t1} from 04.01.1999 – 31.07.2001

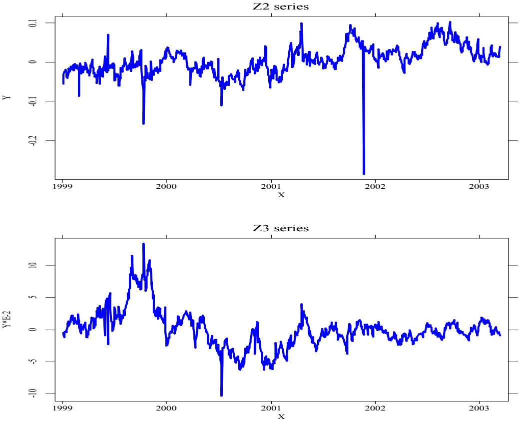

Figure 3: Upper panel: Time series of second factor, z_{t2}. Lower panel: Time series of third factor, z_{t3} from 04.01.1999 – 31.07.2001

Factor Loadings series

Factor loadings determine the movements of the Implied Volatility Surface (IVS)

- z_{t1} may be interpreted as representing the overall shift (up and down movement factor) of the IVS.
- z_{t2} represent tilt or slope effect
- z_{t3} represent changes in curvature (or convexity) of the IVS.

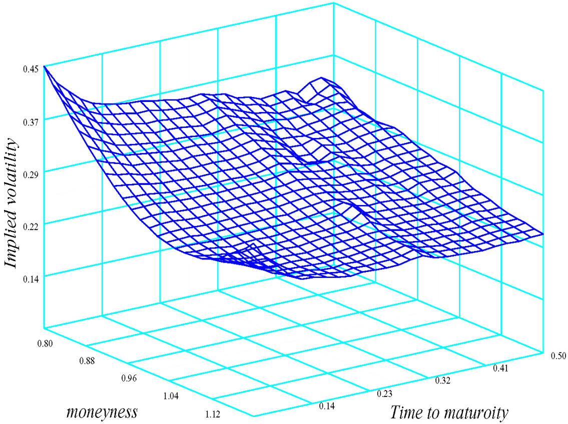

Figure 4: Effect of z_{t1} on IVS

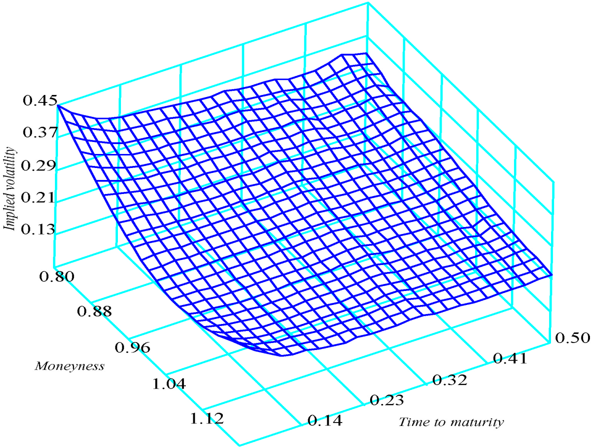

Figure 5: Effect of z_{t2} on IVS

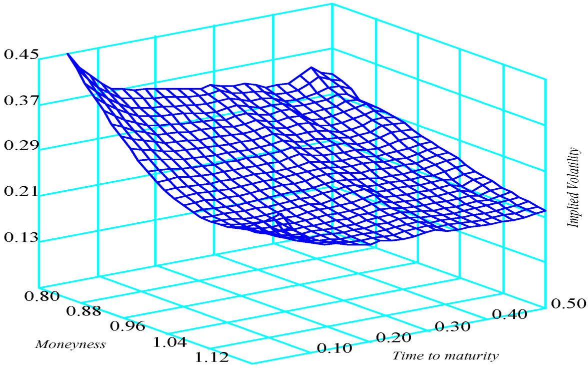

Figure 6: Effect of z_{t3} on IVS

The Data and Unit Root Tests

- time series data on factor loading $z_t = (z_{t1}, z_{t2}, z_{t3})^\top$ are from a DSFM model specified by Borak et al. (2005)
- $T = 1052$ observations on z_t from January 4, 1999 to February 25, 2003, excluding days with no option trades
- z_t is investigated for unit root.
  For stationarity $(I(0))$, a VAR model for levels is analyzed.
  For integration $I(1)$, a VAR in first differences is appropriate

Unit Root Test Statistics

Series	$ADF - AIC$	lag order	$ADF - HQ$	lag order	KPSS	b
z_{t1}	-1.982 [0.295]	6	-2.241 [0.192]	2	1.402^{***}	21
Δz_{t1}	-15.199^{***} [0.000]	5	-23.582^{***} [0.000]	1	0.117	21
z_{t2}	-3.361^{**} [0.013]	8	-4.219^{***} [0.001]	4	2.232^{***}	24
Δz_{t2}	-12.599^{***} [0.000]	12	-15.646^{***} [0.000]	7	0.050	92
z_{t3}	-2.874^{**} [0.049]	7	-2.874^{**} [0.049]	7	0.855^{***}	25
Δz_{t3}	-13.855^{***} [0.000]	6	-13.855^{***} [0.000]	6	0.050	56

Table 1: *Unitroot tests*

Critical values for ADF test are -2.57 *(10%),* -2.86 *(5%) and* -3.44 *(1%) [Mackinnon, (1991)].*
Critical values for KPSS test are 0.347 (10%), 0.463 *(5%) and* 0.739 *(1%) [kwiatkowski (1992)].*
*** *and* ** *denote significance at 1% and 5% level respectively. p-values for ADF tests are in brackets.*
b is bandwidth for KPSS test determined from procedure of [Whitney Newey and Kenneth West(1994)].

Robustness check

To account for possible structural breaks, two subsamples 04.01.1999 – 31.07.2001 (655 obs.) and 01.08.2001 – 24.02.2003 (397obs.) are investigated.

- Unit root tests: ADF and KPSS propose a stationary z_{t2} and a nonstationary z_{t3} for first sample
- ADF test suggests stationarity for both series, while the KPSS test does reject the null of stationarity at the 5% level for second sample

Models for Loadings Dynamics

The dynamics underlying z_t is modelled by a VAR(p) process

- in level $z_t = \nu + A_1 z_{t-1} + \cdots + A_p z_{t-p} + u_t$
- in first difference $\Delta z_t = z_t - z_{t-1}$,
  $\Delta z_t = \nu + A_1 \Delta z_{t-1} + \cdots + \Delta A_p z_{t-p} + u_t$

ν is a $K \times 1$ vector of intercept parameters, $A_i, \;\; i = 1, \ldots, p$ are $K \times K$ parameter matrices, unobservable error term $u_t = (u_{t1}, \ldots, u_{tK})^\top$ with mean zero, time-invariant and non-singular covariance matrix $\Sigma_u = E[u_t u_t^\top]$

VAR Models diagnostics

Full sample (04.01.1999 – 25.02.2003)

- $p = 7$ for z_t and $p = 6$ for Δz_t reveal no autocorrelation

Sub-sample (04.01.1999 – 31.07.2001)

- lag length $p = 3$ reveals residuals with autocorrelation.
- lag length $p = 8$ reveals residuals free of autocorrelation

Evidence for non-normality and ARCH in the residuals is observed but left for further analysis

Impulse Response

Analysis of the inter-relation of model variables

- impulse response functions traces the effect of a shock to one endogenous variable on the other variables in the VAR system
- estimated residual correlation matrix with contemporaneous correlation

$$\widehat{P}_u = \begin{pmatrix} 1 & -0.49 & -0.23 \\ -0.49 & 1 & -0.10 \\ -0.23 & -0.10 & 1 \end{pmatrix} \qquad (2)$$

- orthogonalization by Cholesky decomposition to single out individual shock effect

Impulse Response

Starting by a fairly general model with $p = 7$ lags. Analysis depend on ordering of variables in the system

- innovation in z_{t1} has permanent negative effect on z_{t2} and a small positive effect on z_{t3}, which becomes insignificant after about 6 periods, Figure 8
- innovation in z_{t2} has permanent positive effect on itself but no significant effect with other variables.
  Similar result is obtained for a shock in z_{t3}

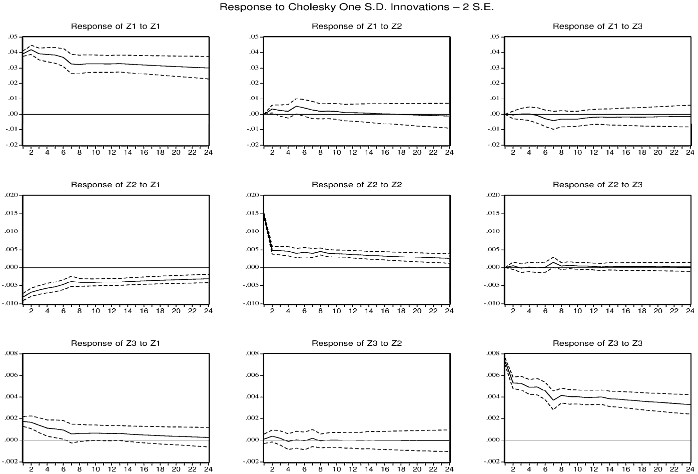

Figure 8: Impulse-Responses: VAR(7) for $z_t = (z_{t1}, z_{t2}, z_{t3})^\top$
Sample period: 04.01.1999 – 25.02.2003

Generalized Impulse Response

- the difference of conditional expectation given a one time shock occurs in series z_t
- coincide with the orthogonalized impulse responses if the residual covariance matrix Σ_u, is diagonal

Overall results from full sample and sub-sample are similar

[Pesaran, M.H. & Shin,Y. (1998)]

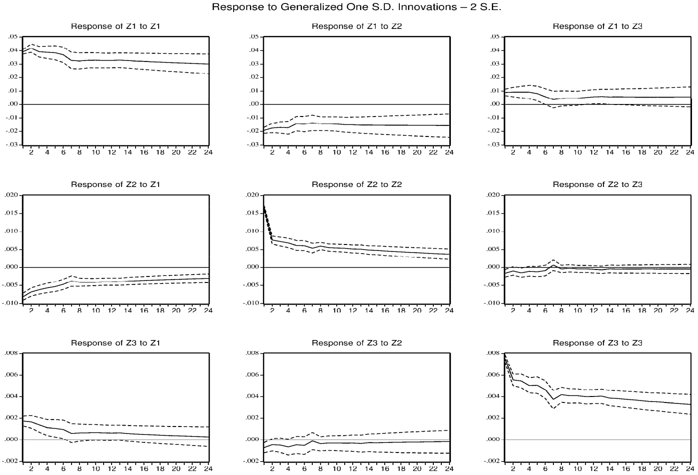

Figure 9: Generalized Impulse-Responses: VAR(7) for $z_t = (z_{t1}, z_{t2}, z_{t3})^\top$. Period: 04.01.1999 – 25.02.2003

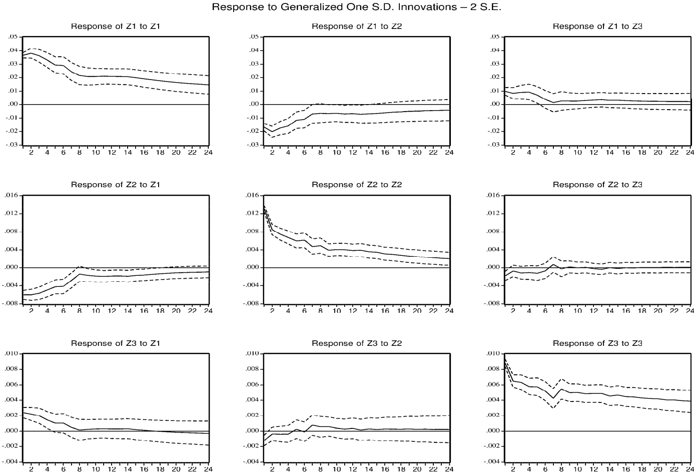

Figure 10: Generalized Impulse-Responses: VAR(8) for $z_t = (z_{t1}, z_{t2}, z_{t3})^\top$. Period: 04.01.1999 – 31.07.2001

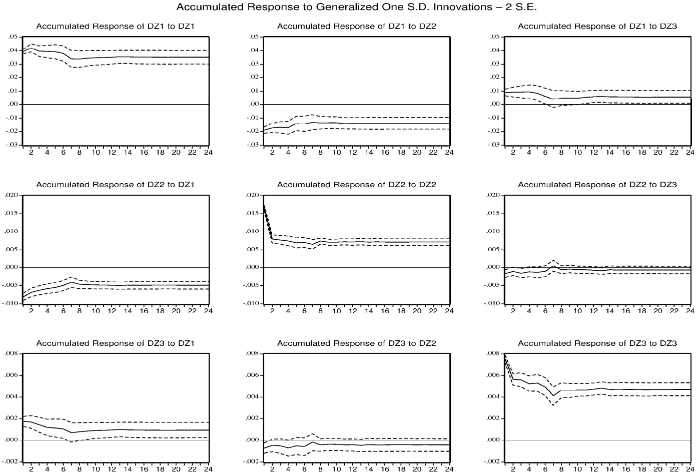

Figure 11: Generalized Impulse-Responses: VAR(6) for $\Delta z_t = (\Delta z_{t1}, \Delta z_{t2}, \Delta z_{t3})^\top$, from 04.01.1999 – 25.02.2003

Granger causality

Addressing the usefulness of each factor in forecasting the others.
Application of the Granger causality test

- testing zero restrictions of some VAR coefficients, which may have a non-standard asymptotic distribution when $I(1)$ variables are in the system.
- overfitting the VAR model by one lag to remove the singularity of the coefficient covariance matrix

[Granger (1969)]

H_0	Test result
$z_{t1} \nrightarrow z_{t2}, z_{t3}$	F(14,3072) = 4.53 (0.00)
$z_{t2} \nrightarrow z_{t1}, z_{t3}$	F(14,3072) = 1.66 (0.06)
$z_{t3} \nrightarrow z_{t1}, z_{t2}$	F(14,3072) = 0.86 (0.60)
$z_{t3} \nrightarrow z_{t1}$	$\chi^2(7) = 5.04$ (0.65)
$z_{t3} \nrightarrow z_{t2}$	$\chi^2(7) = 6.84$ (0.45)
$z_{t1} \nrightarrow z_{t3}$	$\chi^2(7) = 8.02$ (0.33)
$z_{t2} \nrightarrow z_{t3}$	$\chi^2(7) = 6.44$ (0.49)
$z_{t1}, z_{t2} \nrightarrow z_{t3}$	$\chi^2(14) = 12.41$ (0.57)

$\nrightarrow$ denotes 'does not Granger cause'. Results are based on model for z_t using $p = 7$ and full sample period 04.01.1999 - 25.02.2003. p-values in square brackets.

Results

- Granger non-causality of z_{t1} for z_{t2} and z_{t3} and non-causality of z_{t2} for z_{t1} and z_{t3} is rejected at the 10% significance level
- z_{t3} is neither Granger-caused by z_{t1} nor z_{t2} and Granger non-causality from z_{t1} to z_{t3} and from z_{t2} to z_{t3} cannot be rejected

z_{t3} does not influence the dynamics of z_{t1} and z_{t2} in terms of the VAR model

Vega-hedging of z_{t1} and z_{t2}

In DSFM the IV decomposition is given by:

$$\widehat{\sigma}_t = exp(\sum_{k=0}^{K} \widehat{z}_{t,k} \widehat{m}_k).$$

The sensitivities can be computed w.r.t. the factor loadings z_t! An understanding of the sensitivities is derived from the interpretations that z_{t1} and z_{t2} capture the systematic risk faced by an option investor

- $\frac{\partial}{\partial \widehat{z}_{t1}}$ is an **up-and-down shift vega** of the IVS
- $\frac{\partial}{\partial \widehat{z}_{t2}}$ is a **slope shift vega** of the IVS

Conclusion

- factor loadings describe the movements of IVS over time
- a VAR model for levels reveal significant interaction between first and second factor
- a positive shock in first factor has a negative permanent impact on the second and vice versa
- models in first differences and models for subsamples provide similar results for all model specifications

Outlook

- check co-movements of factor loadings as volatility risk indicators in association with movements in macroeconomic conditions like interest rates, exchange rates, oil prices etc.
- hedging of derivative positions and risk management of 'volatility derivatives' such as options on an implied volatility index
- extend modeling of factor loadings with consideration for ARCH effects

Alexander, C.
Principles of the Skew
RISK, 2001, **14**(1):S29–S32.

Black, F. and M. Scholes
The pricing of options and corporate liabilities
Journal of Political Economy, **81**:637–654, 1973.

Borak, S and Härdle, W & and Fengler, M
DSFM fitting of Implied Volatility Surfaces
SFB Discussion paper, Humboldt-Universität zu Berlin, 2005

Cont, R. and J. da Fonseca
The Dynamics of Implied Volatility Surfaces
Quantitative Finance , **2**(1):45–602, 2002.

Doornik, J. A. and Hansen, H.
A practical test of multivariate normality
unpublished paper, Nuffield College , 1994

Fengler,M., Härdle, W. and Mammen, E.
Semiparametric State Space Factor Models
CASE Discusion Paper, Humboldt-Universität zu Berlin, 2003.

Fengler, M.R., W. Härdle, and C. Villa
The dynamics of implied volatilities: A common principle components approach
Review of Derivatives Research , **6**:179–202, 2003.

Granger, C.W.J
Investigating causal relations by econometric models and cross-spectral methods
Econometrica , **37**:424–438, 1969.

Hafner R., and M. Wallmeier
The Dynamics of DAX Implied Volatilities
International Quarterly Journal of Finance, **1**(1):1–27, 2001.

Härdle, W., S. Klinke and M. Müller
XploRe - Learning Guide
Heildelberg: Springer Verlag, 2000

Jürgen, F., Härdle W., Hafner C.M.
Statistics of Financial Markets
Springer Verlag, 2004

Kwiatkowski, D., Peter C. B. Phillips, Peter Schmidt and Yongcheol Shin
Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root
Journal of Econometrics, **54**:159–178, 1992.

Lütkepohl, H.
Introduction to Multiple Time series Analysis
Berlin: Springer-Verlag, 1991

MacKinnon, J.G
Critical Values for Cointegration Tests in C. W. J. Granger & R. F. Engle (eds)
Long-Run Economic Relationships, Oxford University Press, Oxford.

Newey W. & West K.
Automatic Lag Selection in Covariance Matrix Estimation
Review of Economic Studies, **61**:631–653, 1994.

Pesaran, M.H. & Shin,Y.
Generalized impulse response analysis in linear multivariate models,
Economic Letters, **58**:17–29, 1994.

Reinhold Hafner
Stochactic Implied Volatility
A factor-Based Model
Springer velag Heildelberg, 2004

Skiadopoulos, G., S. Hodges, and L. Clewlow
The Dynamics of S&P 500 Implied Volatility Surface
Review of Derivatives Research, **3**:263–282, 1999.

Toda, H.Y. & Yamamoto, T
Statistical Inference in Vector Autoregressions with possibly integrated processes
Journal of Econometrics, **66**:225–250, 1995.