Leveraged ETF implied volatility paradox: a statistical study

Wolfgang Karl Härdle

Zhiwu Hong

Sergey Nasekin

Ladislaus von Bortkiewicz Chair of Statistics
Humboldt–Universität zu Berlin
http://lvb.wiwi.hu-berlin.de
http://irtg1792.hu-berlin.de
http://wise.xmu.edu.cn/english

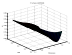
(L)ETFs

- Exchange-traded funds (ETFs): tracking returns on financial quantities and yielding the identical daily return, **e.g.**, SPDR S&P 500 ETF (SPY) tracks the S&P 500.

- Leveraged exchange-traded funds (LETFs): promising a fixed leverage ratio β w.r.t. a given underlying asset or index, **e.g.**,

LETF	β
ProShares Ultra S&P500 (SSO)	2
ProShares UltraPro S&P 500 (UPRO)	3
ProShares UltraShort S&P500 (SDS)	-2
ProShares UltraPro Short S&P 500 (SPXU)	-3

Table 1: LETFs with different β ▸ Illustration



Go with market

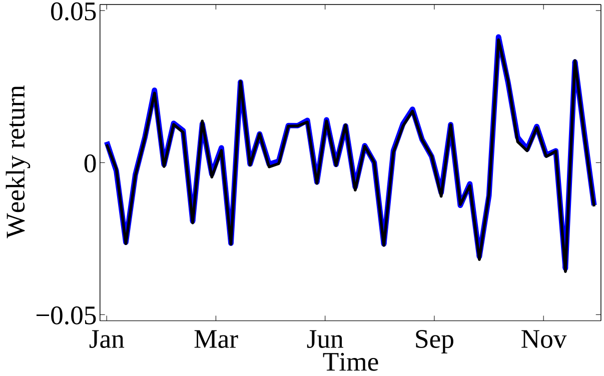


Figure 1: Weekly returns of ETF (SPY) and stock market (S&P 500) (20140101-20141230)

Leverage up/down

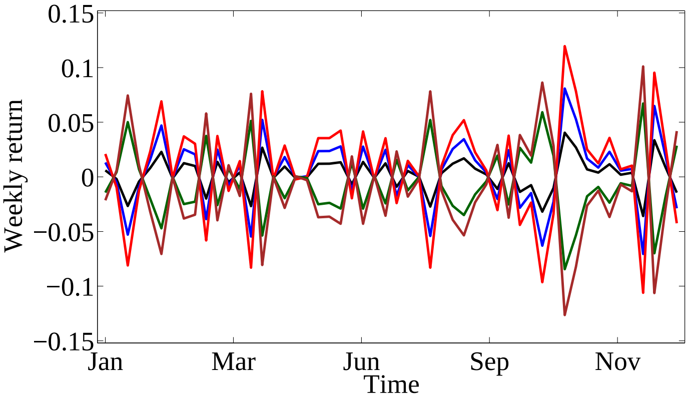


Figure 2: Weekly returns of LETFs (SSO, UPRO, SDS, SPXU) and stock market (S&P 500) (20140101-20141230)

Implied volatility paradox

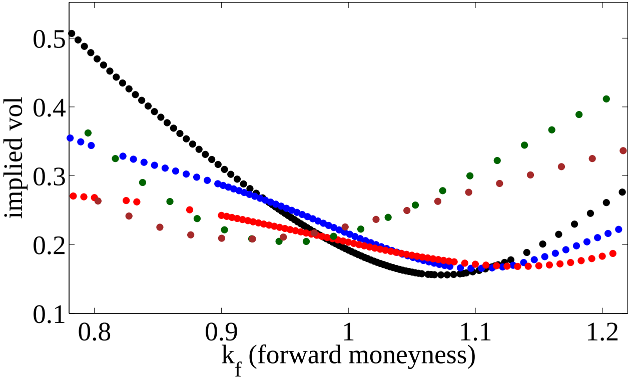


Figure 3: Implied volatility of (L)ETF options (SPY, SSO, UPRO, SDS, SPXU) with 21 days to maturity

Objectives

- introduce moneyness scaling technique
- study statistical significance of moneyness scaling
- identify LETF option price discrepancies using moneyness scaling
- introduce a dynamic model for IVS
- build a trading strategy based on possible arbitrage

Outline

1. Motivation ✓
2. Moneyness scaling
3. Confidence bands
4. DSFM model
5. Estimation Results
6. Trading strategy
7. Conclusions

LETFs and the Black-Scholes model

- asset price dynamics:

$$\frac{dS_t}{S_t} = rdt + \sigma dW_t^{\mathsf{Q}} \tag{1}$$

with interest rate r and volatility σ; W_t^{Q} standard Brownian motion under the risk-neutral measure P^{Q}

- (L)ETF dynamics:

$$\begin{aligned} \frac{dL_t}{L_t} &= \beta\left(\frac{dS_t}{S_t}\right) - \{(\beta - 1)r + c\}dt \\ &= (r - c)dt + \beta\sigma dW_t^{\mathsf{Q}} \end{aligned} \tag{2}$$

$0 \leq c \ll r$ (L)ETF expense ratio

Moneyness scaling (MS)

- with forward moneyness measure $\kappa_f \stackrel{\text{def}}{=} K/\{e^{(r-c)\tau} L_t\}$ and time to maturity τ:

$$\kappa_f{}^{(\beta_1)} = \exp\left\{-\frac{\beta_1}{2}(\beta_1 - \beta_2)\overline{\sigma}^2\tau\right\}(\kappa_f{}^{(\beta_2)})^{\frac{\beta_1}{\beta_2}}, \tag{3}$$

where $\overline{\sigma}$ is the average IV across all strikes

- with log-moneyness measure $LM \stackrel{\text{def}}{=} \log(K/L_t)$

$$LM^{\beta_1} = \frac{\beta_1}{\beta_2}\left[LM^{\beta_2} + \{r(\beta_2 - 1) + c_2\}\tau + \frac{\beta_2(\beta_2 - 1)}{2}\overline{\sigma}^2\tau\right] - \{r(\beta_1 - 1) + c_1\}\tau - \frac{\beta_1(\beta_1 - 1)}{2}\overline{\sigma}^2\tau \tag{4}$$

Data example

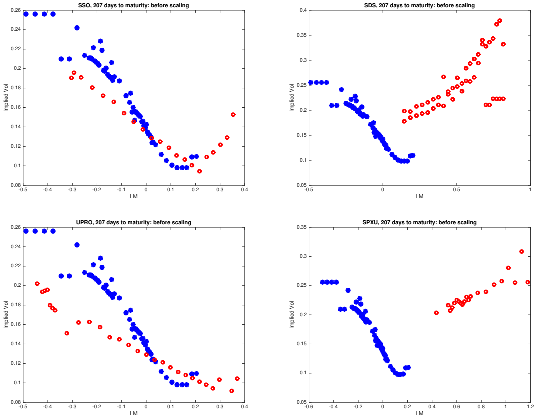


Figure 4: SPY and LETFs implied volatilities before scaling on June 23, 2015 with 207 days to maturity, plotted against their log-moneyness

Data example

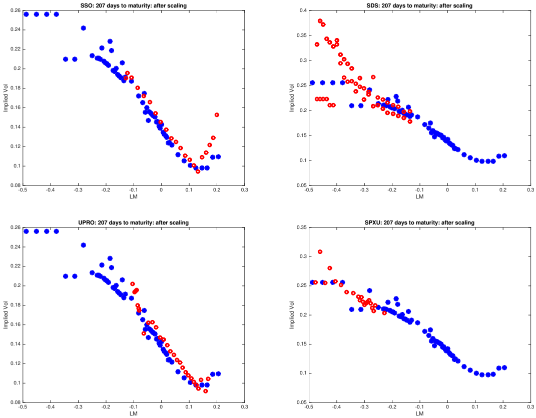


Figure 5: SPY and LETFs implied volatilities after moneyness scaling on June 23, 2015 with 207 days to maturity, plotted against their log-moneyness

Sustainability of MS effect

- is moneyness scaling effect statistically significant?
- how can one study MS sustainability?
- are there practical implications?

Confidence bands

1 Compute the estimate $\widehat{m}_h(X)$ by a local linear M-smoothing procedure with some kernel function and bandwidth h chosen by, e.g., cross-validation.

2 Given $\widehat{\varepsilon}_t \stackrel{\text{def}}{=} Y_t - \widehat{m}(X_t)$; $\widehat{m}$ from Step 1, do a bootstrap resampling from $\widehat{\varepsilon}_t$, that is, for each $t = 1, \ldots, T$, generate a random variable $\varepsilon_i^* \sim \widehat{F}_{\varepsilon|X}(z)$ and a re-sample

$$Y_t^* = \widehat{m}_g(X_t) + \varepsilon_t^*, \quad t = 1, \ldots, T \tag{5}$$

B times (bootstrap replications). with an "oversmoothing" bandwidth $g \gg h$ such as $g = \mathcal{O}(T^{-1/9})$.

Confidence bands

3 For each re-sample $\{X_t, Y_t^*\}_{t=1}^T$ compute $\widehat{m}_{h,g}^*$ using the bandwidth h and construct the random variable

$$d_b \stackrel{\text{def}}{=} \sup_{x \in B} \left[\frac{|\widehat{m}_{h,g}^*(x) - \widehat{m}_g(x)| \sqrt{\widehat{f}_X(x) \widehat{f}_{\varepsilon|X=x_t}(\varepsilon_t^*)}}{\sqrt{\widehat{\mathsf{E}}_{Y|X}\{\psi^2(\varepsilon_t^*)\}}} \right], \tag{6}$$

where $b = 1, \ldots, B$, $\psi(u) = \rho'(\cdot)$, ρ robust loss function

4 Calculate the $1 - \alpha$ quantile d_α^* of $d_1, \ldots, d_B$.

5 Construct

$$\widehat{m}_h(x) \pm \left[\frac{\sqrt{\widehat{\mathsf{E}}_{Y|X}\{\psi^2(\varepsilon_t^*)\}} d_\alpha^*}{\sqrt{\widehat{f}_X(x) \widehat{f}_{\varepsilon|X=x_t}(\varepsilon_t^*)}} \right] \tag{7}$$

Data analysis

- daily data in 20141117-20151117 for the LETFs SSO, UPRO, SDS
- M-smoother estimator of implied volatility Y given forward moneyness X
- times-to-maturity: 0.5, 0.6, 0.7 years
- $B = 1000$

Data analysis

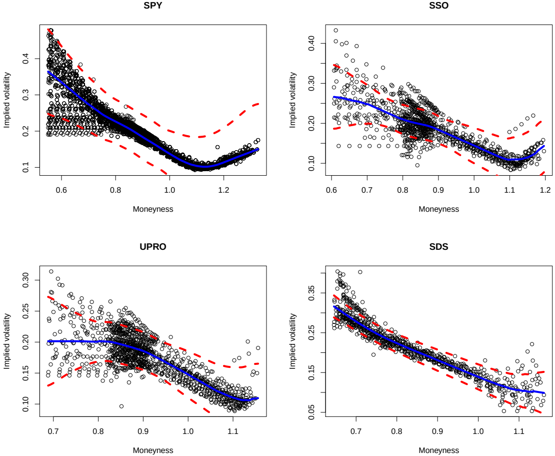


Figure 6: Fitted implied volatility and bootstrap uniform confidence bands for 4 (L)ETFs on S&P500; τ: 0.6 years

Data analysis

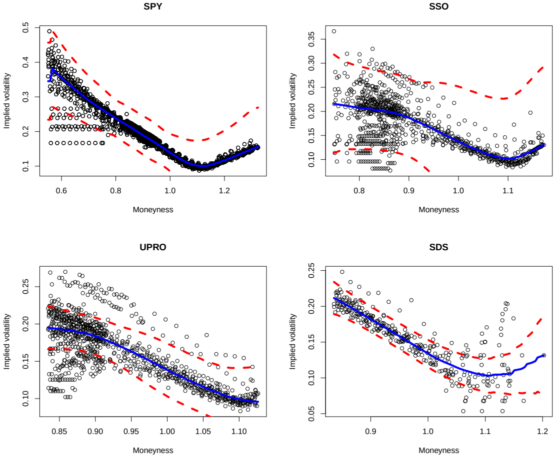


Figure 7: Fitted implied volatility and bootstrap uniform confidence bands for 4 (L)ETFs on S&P500; τ: 0.5 years

Data analysis

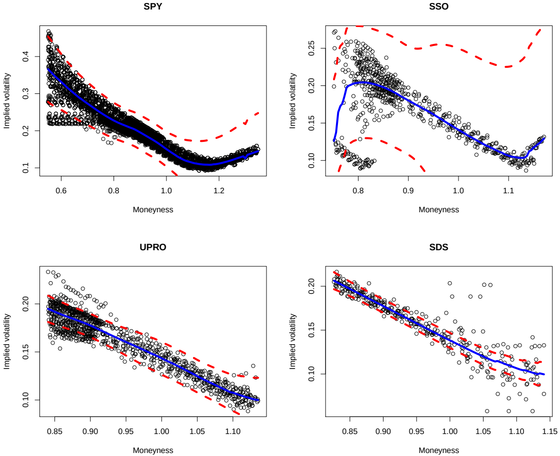


Figure 8: Fitted implied volatility and bootstrap uniform confidence bands for 4 (L)ETFs on S&P500; τ: 0.7 years

Combined bands

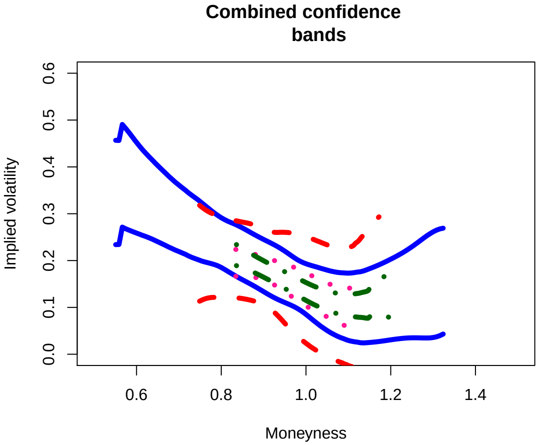


Figure 9: Combined uniform bootstrap confidence bands for SPY, SSO, UPRO and SDS after MS; τ: 0.5 years

Combined bands

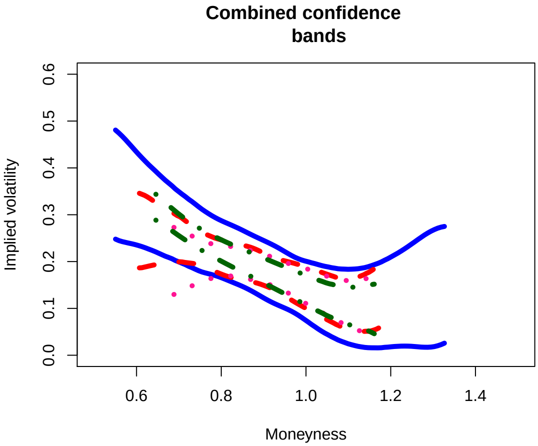


Figure 10: Combined uniform bootstrap confidence bands for SPY, SSO, UPRO and SDS after MS; τ: 0.6 years

Combined bands

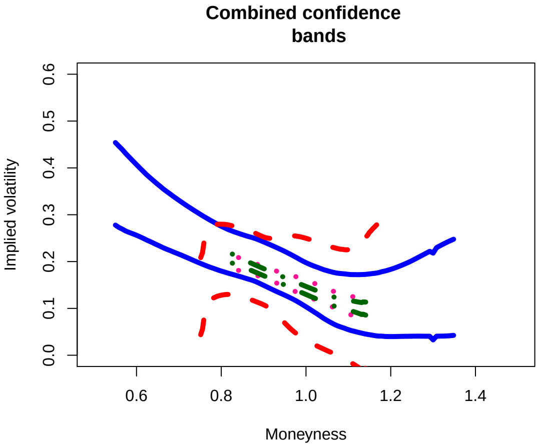


Figure 11: Combined uniform bootstrap confidence bands for SPY, SSO, UPRO and SDS after MS; τ: 0.7 years

Challenges for IV estimation

- how to model the IV surface?
- degenerated data design: IVS observations only for a small number of maturities
- observation grid does not cover desired estimation grid
  - the contracts are not traded for a particular strike
  - institutional arrangements at the futures' exchanges

Implied volatility in time

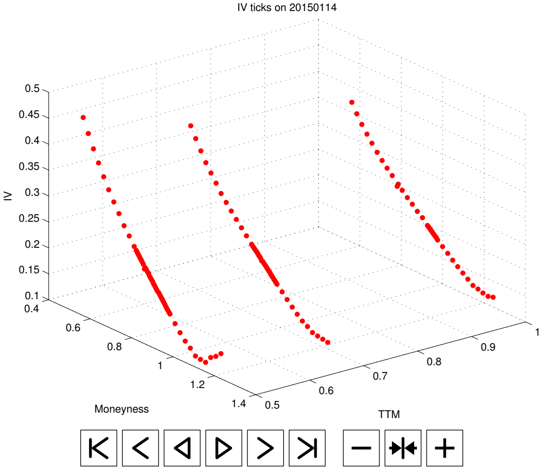


Figure 12: SPY ETF option IV ticks of 20150114-20150408

IV data design

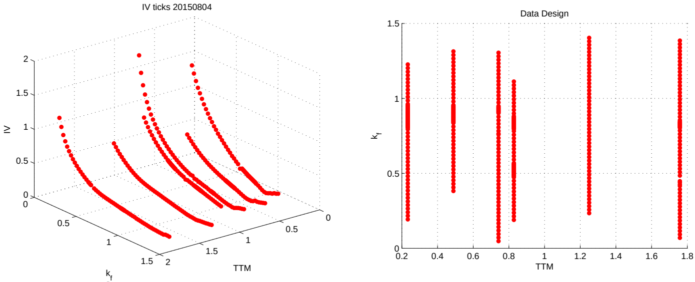


Figure 13: *Left panel*: SPY call IV observed on 20150408; *Right panel*: data design on 20150408

The dynamic semiparametric factor model for IVS

- define $\mathcal{J} \stackrel{\text{def}}{=} [\kappa_{min}, \kappa_{max}] \times [\tau_{min}, \tau_{max}]$, $Y_{i,j}$ implied volatility, $i = 1, \ldots, I$ time index, $j = 1, \ldots, J_i$ option intraday numbering on day i, $X_{i,j} \stackrel{\text{def}}{=} (\kappa_{i,j}, \tau_{i,j})^\top$, $Y_{i,j} \stackrel{\text{def}}{=} \sigma_{i,j}$ implied volatility; κ some measure of moneyness, e.g., log- or forward

- assume

$$Y_{i,j} = \mathcal{Z}_i^\top m(X_{i,j}) + \varepsilon_{i,j}, \tag{8}$$

  where $\mathcal{Z}_i = (1, Z_i^\top)$, $Z_i = (Z_{i,1}, \ldots, Z_{i,L})^\top$ unobservable L-dimensional process, $m = (m_0, \ldots, m_L)^\top$, real-valued functions m_l, $l = 1, \ldots, L+1$ defined on a subset of $\mathbb{R}^d$

- $X_{i,j}$, $\varepsilon_{i,j}$ are independent, $\varepsilon_{i,j} \sim (0, \sigma^2)$, $\sigma^2 < \infty$

DSFM for IVS

Approximate, Park et al. (2009):

$$\mathsf{E}(Y_i|X_i) = \mathcal{Z}_i^\top m(X_i) = \mathcal{Z}_i^\top \mathcal{A}\psi(X_i), \tag{9}$$

where

$$\psi(X_i) \stackrel{\text{def}}{=} \{\psi_1(X_i), \ldots, \psi_K(X_i)\}^\top \text{ space basis,}$$
$$\mathcal{A} : (L+1) \times K \text{ coefficient matrix}$$

Choose $\{\psi_k : 1 \le k \le K\}$ tensor B-spline basis ▸ Details, de Boor (2001)

K is the number of tensor B-spline sites s_k^τ, s_k^κ

Tensor B-spline basis

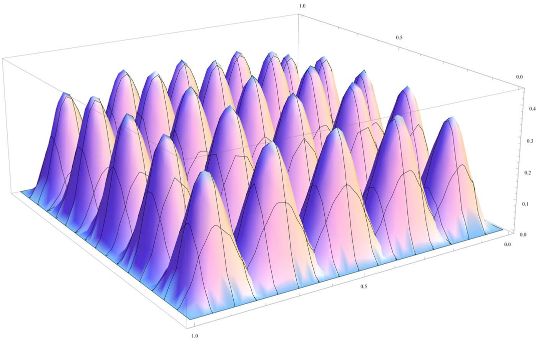


Figure 14: Tensor B-spline basis with 15 $\times$ 15 knots on $[0,1] \times [0,1]$, odd intervals

Estimation

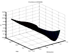
Define the least-squares estimators $\widehat{Z}_i = (\widehat{Z}_{i,1}, \ldots, \widehat{Z}_{i,L})^\top$, $\widehat{\mathcal{A}} = (\widehat{\alpha}_{l,k})_{l=0,\ldots,L;k=1,\ldots,K}$

$$(\widehat{Z}_i, \widehat{\mathcal{A}}) = \arg\min_{Z_i,\mathcal{A}} S(\mathcal{A}, Z), \tag{10}$$

where

$$S(\mathcal{A}, Z) \stackrel{\text{def}}{=} \sum_{i=1}^{I} \sum_{j=1}^{J_i} \left\{ Y_{i,j} - (1, Z_i^\top)\mathcal{A}\psi(X_{i,j}) \right\}^2 \tag{11}$$

Once $\widehat{\mathcal{A}}$ obtained, m can be estimated as $\hat{m} = \widehat{\mathcal{A}}\psi$



Identification

- the problem (10) can be solved via numeric algorithm Details
- under certain conditions Details, geometric convergence to a solution
- (10) has no unique solution: orthonormalize $\widehat{Z}_i$, $\hat{m}$ for better interpretation, see Fengler et al. (2007)

Why DSFM?

- can model and forecast the whole IV surface:

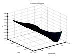
$$\widehat{IV}_{t;i,j} = \widehat{m}_{0;i,j} + \sum_{l=1}^{L} \widehat{\mathcal{Z}}_{l,t} \widehat{m}_{l;i,j}, \tag{12}$$

where

$$\widehat{m}_{l;i,j} = \sum_{i}^{|s^{\kappa}|} \sum_{j}^{|s^{\tau}|} \widehat{A}_{l;i,j} \psi_{i,k_{\kappa}}(\kappa_i) \psi_{j,k_{\tau}}(\tau_j), \tag{13}$$

k_{κ}, k_{τ} knot sequences, s^{κ}, s^{τ} site sets; $\widehat{\mathcal{A}}$ reshaped into a $|s^{\kappa}| \times |s^{\tau}| \times L$ array of L matrices $\widehat{A}$ of dimension $|s^{\kappa}| \times |s^{\tau}|$

- stochastic loadings $\widehat{\mathcal{Z}}_t$ have vector autoregressive (VAR) dynamics



Data overview

		Min.	Max.	Mean	Stdd.	Skewn.	Kurt.
SPY	TTM	0.26	1.05	0.76	0.19	−0.54	2.76
	Moneyness	0.05	1.43	0.48	0.17	−0.34	3.15
	IV	0.25	1.55	0.46	0.23	1.94	7.17
SSO	TTM	0.21	1.04	0.63	0.25	0.01	1.76
	Moneyness	0.18	1.69	0.63	0.29	0.92	3.61
	IV	0.25	1.34	0.41	0.11	1.91	10.81

Table 2: Summary statistics on SPY, SSO ETF options from 20140920 to 20150630 (in total $\sum_i J_t = 9828, 7619$ datapoints, respectively). Source: Datastream

Estimation

- estimation space $[\kappa_{min}, \kappa_{max}] \times [\tau_{min}, \tau_{max}]$ re-scaled (via marginal edf) to $[0, 1]^2$
- 9 knots in moneyness and 7 knots in maturity direction, cubic splines ($k = 3$) in both directions, so $K = 24$
- starting values for Z_i generated from a stable VAR process
  ▸ Details
- starting values for $\mathcal{A}$ randomly generated from $U(0, 1)$
- convergence tolerance for the Newton algorithm: $1e$-06

Model order selection

Select model order by explained variance $EV(L)$

$$EV(L) \stackrel{\text{def}}{=} 1 - \frac{\sum_{i=1}^{I} \sum_{j=1}^{J_i} \left\{ Y_{i,j} - \sum_{l=0}^{L} \widehat{Z}_{i,l} \hat{m}_l(X_{i,j}) \right\}^2}{\sum_{i=1}^{I} \sum_{j=1}^{J_i} (Y_{i,j} - \overline{Y})^2} \tag{14}$$

No. factors	$EV(L)$
3	0.912
4	0.916
5	0.924
6	0.925

Estimate $L = 3$ basis functions

Dynamics of $\widehat{Z}_i$

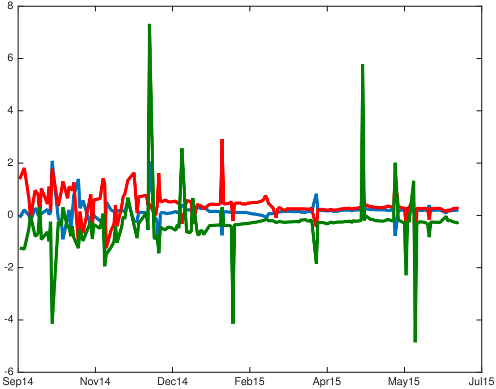


Figure 15: Time dynamics of $\widehat{Z}_{i,1}$, $\widehat{Z}_{i,2}$, $\widehat{Z}_{i,3}$

VAR modelling of $\widehat{Z}_i$

Model order n	AIC(n)	HQ(n)	SC(n)
1	-4.20^*	-4.10^*	-3.96^*
2	-4.13	-3.96	-3.72
3	-4.07	-3.83	-3.48
4	-4.03	-3.72	-3.27
5	-3.97	-3.59	-3.03

Table 3: The VAR model selection criteria. The smallest value is marked by an asterisk

VAR modelling of $\widehat{Z}_i$

- all roots of VAR(1) model lie inside the unit circle
- Portmanteau and Breusch-Godfrey LM test results with 12 lags fail to reject residual autocorrelation at 10% significance level

Estimated factor functions $\hat{m}_l$

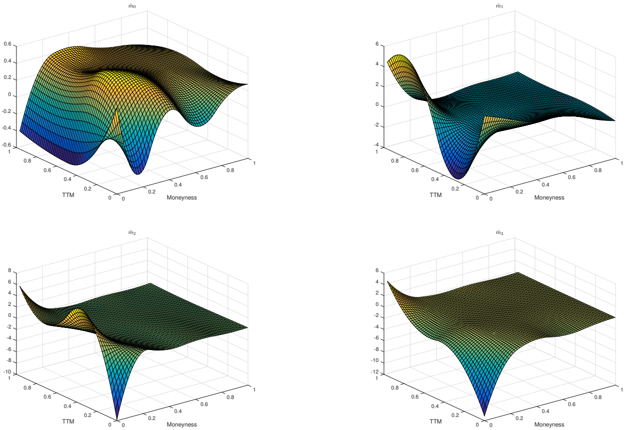


Figure 16: Factor functions $\hat{m}_0$, $\hat{m}_1$, $\hat{m}_2$, $\hat{m}_3$

Bias comparison

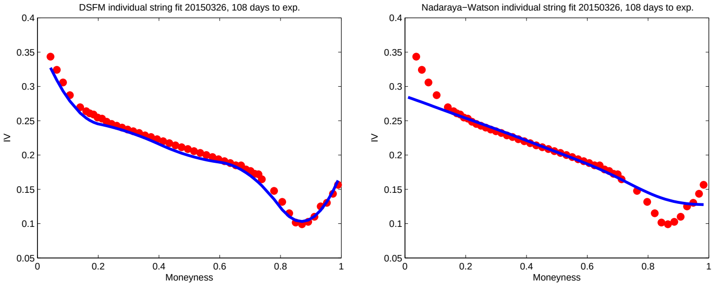


Figure 17: Bias comparison of the DSFM (*left panel*) and the Nadaraya-Watson estimator with $\hat{h} = (0.13, 0.12)^\top$, $\hat{h}$ by Scott's rule, for the 108 days to expiry data (red dots) on 20150326

Strategy motivation

- "trade-with-the-smile/skew" strategy adapted for the special case of ETF-LETF option IV discrepancy
- use the ETF option data to estimate the model (theoretical) smile of the leveraged counterpart and the information from the IV surface forecast to recognize the future (one-period-ahead) possible IV discrepancy
- delta-hedging is required

Strategy setup

Choose moving window width w; for each $t = w, \dots, T$ do:

1. Given $\beta_2 = 1$, β_1, re-scale LM^{β_2} according to the MS formula (4) to obtain $\widehat{LM}^{\beta_1}$, obtain the "model" moneyness coordinate for DSFM estimation.
2. Estimate (26) on $[\widehat{LM}^{\beta_1}_{min}, \widehat{LM}^{\beta_1}_{max}] \times [\tau^{SPY}_{min}, \tau^{SPY}_{max}]$ (re-scaled to $[0,1]^2$), obtain the IV surface estimates $\widehat{IV}_1, \dots, \widehat{IV}_t$.
3. Forecast $\widehat{IV}_{t+1}$ using the VAR structure of $\widehat{\mathcal{Z}}_t$ and factor functions $\widehat{m}_l$.

Strategy setup

4. Choose a specific IV "string" for some τ^* at time point t using SSO option data and calculate the marginally transformed $\widehat{LM}^{\beta_1}_{\tau^*}$ of the *true* SSO log-moneyness LM^{β_1} using the marginal distribution of $\widehat{LM}^{\beta_1}$.

5. Using $\widehat{LM}^{\beta_1}_{\tau^*}$, τ^* and $\widehat{IV}_{t+1}$, interpolate the "theoretical" IV $\widehat{IV}_{t+1}$ over the marginally re-scaled $[\tau^*, \tau^*] \times [\widehat{LM}^{\beta_1}_{min}, \widehat{LM}^{\beta_1}_{max}]$ to obtain "theoretical" values $\widehat{IV}_{t+1;LM^{\beta_1}_{\tau^*},\tau^*}$.

Strategy setup

6. Compare "theoretical" $\widehat{IV}_{t+1;LM^{\beta_1}_{\tau^*},\tau^*}$ with "true" $IV_{t;LM^{\beta_1}_{\tau^*},\tau^*}$ and construct a delta-hedged option portfolio: buy (long) an option with the absolute largest negative deviation from the "theoretical" IV (IV expected to fall) and sell (short) an option with the smallest positive deviation from the "theoretical" IV (IV expected to increase). Use the underlying SSO LETF asset to make the whole portfolio delta-neutral.
7. At time point $t+1$, terminate the portfolio, calculate profit/loss and repeat until time T.

Discrepancy 1

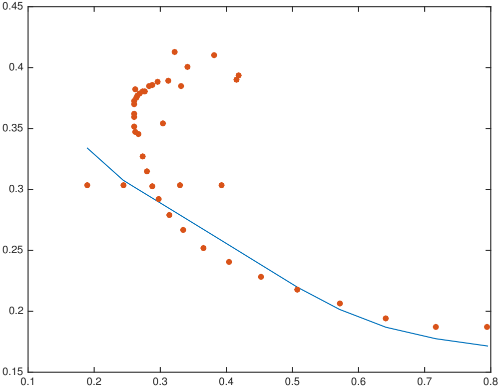


Figure 18: True SSO and MS-DSFM-modeled implied volatilities on 20141121

Discrepancy 2

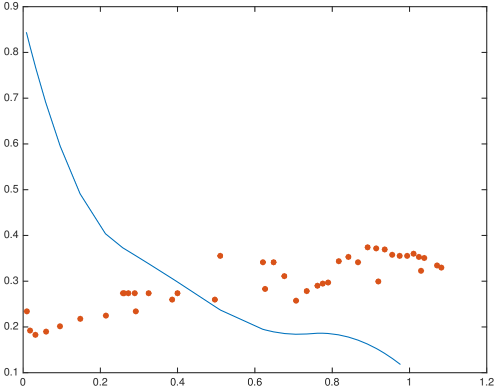


Figure 19: True SSO and MS-DSFM-modeled implied volatilities on 20150326

Financial gains

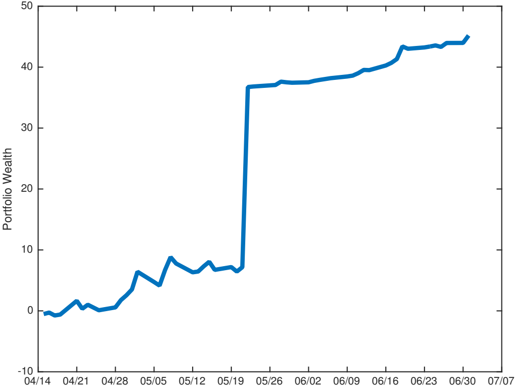


Figure 20: Option portfolios cumulative gains in 2015; $w = 100$, initial (long) portfolio 25.14

Conclusions

- moneyness scaling effect is not sustainable
- one can use use MS to detect arbitrage opportunities on the LETF option market
- the combined MS and DSFM approach allows to build profitable trading strategies on the LETF option market

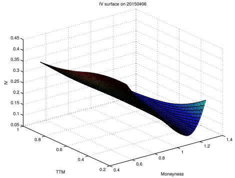
Leveraged ETF implied volatility paradox: a statistical study

Wolfgang Karl Härdle

Zhiwu Hong

Sergey Nasekin

Ladislaus von Bortkiewicz Chair of Statistics
Humboldt–Universität zu Berlin
http://lvb.wiwi.hu-berlin.de
http://irtg1792.hu-berlin.de
http://wise.xmu.edu.cn/english



References I

de Boor, C.
A Practical Guide to Splines. Revised Edition
Springer, 2001

Cai, T. T., Levine, M., Wang, L.
Variance Function Estimation in Multivariate Nonparametric Regression with Fixed Design
Journal of Multivariate Analysis, 100, 126-136 (2009)

Derman, E., Kani, I., & Zou, J. Z. (1996). The local volatility surface: Unlocking the information in index option prices. Financial analysts journal, 25-36.

References II

Fengler, M. R.
Arbitrage-free smoothing of the implied volatility surface
Quantitative Finance, 9(4), 417-428 (2009)

Hall, P., Horowitz, J.
A simple bootstrap method for constructing nonparametric confidence bands for functions
The Annals of Statistics, 41(4), 1892-1921 (2013)

Fengler, M. R., Härdle, W. K., & Mammen, E.
A semiparametric factor model for implied volatility surface dynamics
Journal of Financial Econometrics, 5(2), 189-218 (2007)

References III

Leung, T., & Sircar, R. (2014). Implied volatility of leveraged ETF options. Applied Mathematical Finance, (ahead-of-print), 1-27.

Fengler, M. R. (2005). Semiparametric modeling of implied volatility. Springer.

Park, B. U., Mammen, E., Härdle, W., Borak, S.
Time Series Modelling With Semiparametric Factor Dynamics
Journal of the American Statistical Association, Vol. 104, No. 485 (2009)

References IV

Prautzsch, H., Boehm, W., Paluszny, M.
Bézier and B-spline techniques
Springer, 2002

Ruppert, D., Wand, M. P.
Multivariate Locally Weighted Least Squares Regression
The Annals of Statistics, 22(3), 1346-1370 (1994)

Spokoiny, V.
Variance Estimation for High-Dimensional Regression Models
Journal of Multivariate Analysis, 82, 111-113 (2002)

S&P 500 and (L)ETFs Return

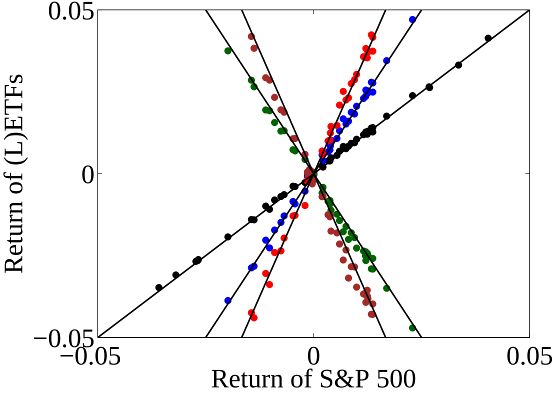


Figure 21: Weekly return relationship of S&P 500 and (L)ETFs (SPY, SSO, UPRO, SDS, SPXU) (20140101-20141230) Back

Local Polynomial Smoothing (LPS) for IVS

$$\min_{\alpha(x_1,x_2)\in\mathbb{R}^5} \sum_{i=1}^{n} \{y_i - \alpha_0 - \alpha_1(x_{1i} - x_1) - \alpha_2(x_{2i} - x_2) - \alpha_3(x_{1i} - x_1)^2$$

$$-\alpha_4(x_{2i} - x_2)^2 - \alpha_5(x_{1i} - x_1)(x_{2i} - x_2)\}^2 \frac{K(\frac{x_1 - x_{1i}}{h_1})}{h_1} \frac{K(\frac{x_2 - x_{2i}}{h_2})}{h_2},$$

where $K(u) = \frac{3}{4}(1 - u^2)I(|u| < 1)$ Epanechnikov kernel;

$$\frac{\partial y}{\partial x_1}\Big|_{(x_{1i},x_{2i})} = \alpha_1, \quad \frac{\partial y}{\partial x_2}\Big|_{(x_{1i},x_{2i})} = \alpha_2, \quad \frac{\partial^2 y}{\partial {x_1}^2}\Big|_{(x_{1i},x_{2i})} = 2\alpha_3$$

Return to "Simulation example"

BS formula and its derivatives for LETF options

Recall (??), the BS is

$$C_{BS}^{\beta} = e^{-c\tau} L_t \Phi(d_1) - e^{-r\tau} K \Phi(d_2) \tag{15}$$

where $\Phi(\cdot)$ is the cdf of $N(0,1)$.

$$\begin{aligned}
\frac{\partial C_{BS}^{\beta}}{\partial T} &= \frac{\partial C^{BS}}{\partial T} + \frac{\partial C^{BS}}{\partial \sigma} |\beta| \frac{\partial \widehat{\sigma}}{\partial T} \qquad (16) \\
&= \frac{e^{-c\tau} L_t |\beta| \widehat{\sigma} \Phi(d_1)}{2\sqrt{\tau}} - c e^{-c\tau} L_t \Phi(d_1) + r e^{-r\tau} K \Phi(d_2) \\
&\quad + \frac{\partial C^{BS}}{\partial \sigma} |\beta| \frac{\partial \widehat{\sigma}}{\partial T}
\end{aligned}$$

BS formula and its derivatives for LETF option

$$
\begin{aligned}
\frac{\partial C_{BS}^{\beta}}{\partial K} &= \frac{\partial C^{BS}}{\partial K} + \frac{\partial C^{BS}}{\partial \sigma} |\beta| \frac{\partial \widehat{\sigma}}{\partial K} \\
&= -e^{-r\tau}\Phi(d_2) + \frac{\partial C^{BS}}{\partial \sigma} |\beta| \frac{\partial \widehat{\sigma}}{\partial K}
\end{aligned} \tag{17}
$$

$$
\begin{aligned}
\frac{\partial^2 C_{BS}^{\beta}}{\partial K^2} &= \frac{\partial^2 C^{BS}}{\partial K^2} + 2\frac{\partial^2 C^{BS}}{\partial K \partial \widehat{\sigma}} |\beta| \frac{\partial \widehat{\sigma}}{\partial K} \\
&\quad + \frac{\partial^2 C^{BS}}{\partial \sigma^2} |\beta|^2 \left(\frac{\partial \widehat{\sigma}}{\partial K}\right)^2 + \frac{\partial C^{BS}}{\partial \sigma} |\beta| \frac{\partial^2 \widehat{\sigma}}{\partial K^2} \\
&= \frac{\partial C^{BS}}{\partial \sigma} \left\{ \frac{1}{|\beta| K^2 \widehat{\sigma} \tau} + \frac{2 d_1}{K \widehat{\sigma} \sqrt{\tau}} \frac{\partial \widehat{\sigma}}{\partial K} + \frac{|\beta| d_1 d_2}{\widehat{\sigma}} \left(\frac{\partial \widehat{\sigma}}{\partial K}\right)^2 + |\beta| \frac{\partial^2 \widehat{\sigma}}{\partial K^2} \right\}
\end{aligned} \tag{18}
$$

BS formula and its derivatives for LETF option

Substitute (17)-(18) into (**??**), obtain the Dupire formula in terms of IV and its derivatives:

$$\widehat{\sigma}^2_{K,T} = \frac{\frac{\widehat{\sigma}}{2\tau} + \frac{\partial \widehat{\sigma}}{\partial T} + (r-c)K\frac{\partial \widehat{\sigma}}{\partial K}}{\frac{1}{2}K^2\left\{\frac{1}{K^2\widehat{\sigma}\tau} + \frac{2|\beta|d_1}{K\widehat{\sigma}\sqrt{\tau}}\frac{\partial \widehat{\sigma}}{\partial K} + \frac{|\beta|^2 d_1 d_2}{\widehat{\sigma}}\left(\frac{\partial \widehat{\sigma}}{\partial K}\right)^2 + |\beta|^2\frac{\partial^2 \widehat{\sigma}}{\partial K^2}\right\}}$$

Back

Moneyness scaling

Given the general solution of (2):

$$L_T = L_t \exp\left\{(r-c)(T-t) - \frac{\beta^2}{2}\int_t^T \sigma_s^2 ds + \beta \int_t^T \sigma_s dW_s^*\right\}, \tag{19}$$

write (19) for $L_T^{(\beta_1)}$, $L_T^{(\beta_2)}$, obtain

$$\frac{L_T^{(\beta_1)}}{e^{(r-c)\tau} L_t^{(\beta_1)}} = \exp\left(-\frac{\beta_1^2}{2}\int_0^\tau \sigma_s^2 ds + \beta_1 \int_0^\tau \sigma_s dW_s\right) \tag{20}$$

$$\frac{L_T^{(\beta_2)}}{e^{(r-c)\tau} L_t^{(\beta_2)}} = \exp\left(-\frac{\beta_2^2}{2}\int_0^\tau \sigma_s^2 ds + \beta_2 \int_0^\tau \sigma_s dW_s\right) \tag{21}$$

where σ_s is the instantaneous volatility at time s.

Moneyness scaling

From (21) follows:

$$\int_t^T \sigma_s dW_s^* = \frac{\log\left(\frac{L_T^{(\beta_2)}}{e^{(r-c)\tau} L_t^{(\beta_2)}}\right) + \frac{\beta_2^2}{2}\int_0^\tau \sigma_s^2 ds}{\beta_2} \tag{22}$$

Substitute (22) into (20) to eliminate the stochastic term $\int_t^T \sigma_s dW_s^*$, obtain:

$$\frac{L_T^{(\beta_1)}}{e^{(r-c)\tau} L_t^{(\beta_1)}} = \exp\left\{-\frac{\beta_1}{2}(\beta_1 - \beta_2)\int_0^\tau \sigma_s^2 \mathrm{d}s\right\}\left\{\frac{L_T^{(\beta_2)}}{e^{(r-c)\tau} L_t^{(\beta_2)}}\right\}^{\frac{\beta_1}{\beta_2}} \tag{23}$$

Moneyness scaling

Take logs and expectations conditioned on $K^{(\beta_1)} = L_T^{(\beta_1)}$ and $K^{(\beta_2)} = L_T^{(\beta_2)}$, obtain

$$\log({k_f}^{(\beta_1)}) = -\frac{\beta_1}{2}(\beta_1 - \beta_2)\mathrm{E}^* \left(\int_0^\tau \sigma_s^2 \mathrm{d}s \,\middle|\, K^{(\beta_1)} = L_T^{(\beta_1)}, K^{(\beta_2)} = L_T^{(\beta_2)} \right) + \frac{\beta_1}{\beta_2} \log({k_f}^{(\beta_2)})$$

Assuming constant σ and exponentiating, one obtains (3)

Return to "Moneyness scaling"

Tensor product B-splines

Define $U \stackrel{\text{def}}{=} \{\sum_i \alpha_i N_{i,h,s} : \alpha_i \in \mathbb{R}, i \in \mathbb{Z}\}$,
$V \stackrel{\text{def}}{=} \left\{\sum_j \beta_j N_{j,k,t} : \beta_j \in \mathbb{R}, j \in \mathbb{Z}\right\}$, then the tensor product
B-spline is $b \stackrel{\text{def}}{=} \sum_{i,j} \gamma_{i,j} N_{i,j}$, $\gamma_{i,j} \in \mathbb{R}$, $w \in U \otimes V$, where

$$N_{i,j}(x,y) \stackrel{\text{def}}{=} N_{i,h,s}(x) N_{j,k,t}(y),$$

$$N_{i,k,t}(x) \stackrel{\text{def}}{=} \left(\frac{x - t_i}{t_{i+k} - t_i}\right) N_i^{k-1}(x) + \left(\frac{t_{i+k+1} - x}{t_{i+k+1} - t_{i+1}}\right) N_{i+1}^{k-1}(x),$$

with the starting point

$$N_{i,0,t}(x) \stackrel{\text{def}}{=} \begin{cases} 1, & \text{if } t_i \leq x < t_{i+1}, \\ 0, & \text{otherwise,} \end{cases}$$

here $k \in \mathbb{N}$, t_i infinite set of knots Back to "DSFM for IVS"

Numerical differentiation

Use Taylor expansion for $\hat{\sigma}(\tau+h,\kappa)$, $\hat{\sigma}(\tau-h,\kappa)$, $\hat{\sigma}(\tau,\kappa+h)$, $\hat{\sigma}(\tau,\kappa-h)$, obtain the following approximations, h small:

- $$\frac{\partial\widehat{\sigma}}{\partial\tau}=\frac{\hat{\sigma}(\tau+h,\kappa)-\hat{\sigma}(\tau-h,\kappa)}{2h}$$
- $$\frac{\partial\widehat{\sigma}}{\partial\kappa}=\frac{\hat{\sigma}(\tau,\kappa+h)-\hat{\sigma}(\tau,\kappa-h)}{2h}$$
- $$\frac{\partial^2\widehat{\sigma}}{\partial\kappa^2}=\frac{\hat{\sigma}(\tau,\kappa+h)-2\hat{\sigma}(\tau,\kappa)+\hat{\sigma}(\tau,\kappa-h)}{h^2}$$

Return to "Backing out LVS"

Tensor B-spline derivatives

- a B-spline surface $b(x, y)$ can be represented in Bézier form, Prautzsch et al. (2002)

$$b(x, y) = \sum_i \sum_j \beta_{ij} B_i^n(x) B_j^k(y), \tag{24}$$

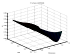
where B_i^n are Bernstein polynomials ▸ Details

- the partial derivatives of (24) are given by

$$\frac{\partial^{q+r} b(x, y)}{\partial x^q \partial y^r} = \frac{n!k!}{n!k! - qr} \sum_i \sum_j \Delta^{01} \Delta^{q,r-1} \beta_{ij} B_i^{n-q}(x) B_j^{k-r}(y), \tag{25}$$

where the forward difference $\Delta^{qr} \beta_{ij} = \beta_{i+q,j+r} - \beta_{i,j}$

▸ Return to "Tensor B-spline derivatives"

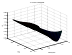


Bernstein polynomials

Bernstein polynomials of degree n are given by

$$B_i^n(u) = \begin{pmatrix} n \\ i \end{pmatrix} u^i (1-u)^{n-i},$$

where $i = 0, \ldots, n$ Return to "Tensor B-spline derivatives"



Convergence conditions

Initial choice (α^0, Z^0) such that, Park et al. (2009):

- A1 it holds that $\sum_{i=1}^{I} Z_i^0 = 0$; $\sum_{i=1}^{I} Z_i^0 {Z_i^0}^\top$ and the Hessian from (10) at (α^0, Z^0), $\mathcal{H}(\alpha^0, Z^0)$ are invertible
- A2 there exists a version $(\widehat{\alpha}, \widehat{Z})$ with $\sum_{i=1}^{I} \widehat{Z}_i = 0$ such that $\sum_{i=1}^{I} \widehat{Z}_i {Z_i^0}^\top$ is invertible. Also, $\hat{\alpha}_l = (\hat{\alpha}_{l1}, \ldots, \hat{\alpha}_{lK})^\top$, $l = 0, \ldots, L$ are linearly independent

Return to "Identification"

Numeric algorithm I

The first-order conditions for (10):

$$\frac{\partial S(\mathcal{A}, Z)}{\partial \alpha} = 2\sum_{i=1}^{I}\left\{(\Psi_i\Psi_i^\top)\otimes(\mathcal{Z}_i\mathcal{Z}_i^\top)\right\}\alpha - 2\sum_{i=1}^{I}(\Psi_i Y_i)\otimes\mathcal{Z}_i, \tag{26}$$

$$\begin{aligned}\frac{\partial S(\mathcal{A}, Z)}{\partial Z} &= 2(\mathcal{Z}_1^\top\mathcal{A}\Psi_1\Psi_1^\top A^\top - Y_1^\top\Psi_1^\top A^\top, \ldots, \mathcal{Z}_I^\top\mathcal{A}\Psi_I\Psi_I^\top A^\top \\ &\quad - Y_I^\top\Psi_I^\top A^\top),\end{aligned} \tag{27}$$

where A is $\mathcal{A}$ without 1st row, $\Psi_i \stackrel{\text{def}}{=} \{\psi(X_{i,1}), \ldots, \psi(X_{i,J_i})\}$, $\alpha \stackrel{\text{def}}{=} \text{vec}(\mathcal{A})$ Return to "Identification"

Numeric algorithm II

The second-order conditions for (10):

$$\frac{\partial^2 S(\mathcal{A}, Z)}{\partial \alpha^2} = 2\sum_{i=1}^{I}\left\{(\Psi_i\Psi_i^\top)\otimes(\mathcal{Z}_i\mathcal{Z}_i^\top)\right\}, \tag{28}$$

$$\frac{\partial^2 S(\mathcal{A}, Z)}{\partial Z^2} = \begin{pmatrix} A\Psi_1\Psi_1^\top A^\top & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A\Psi_I\Psi_I^\top A^\top \end{pmatrix}, \tag{29}$$

$$\frac{\partial^2 S(\mathcal{A}, Z)}{\partial \alpha \partial Z} = 2\{F_1(\alpha, Z), \ldots, F_I(\alpha, Z)\}, \tag{30}$$

where

$F_i(\alpha, Z) \stackrel{\text{def}}{=} (\Psi_i\Psi_i^\top A^\top)\otimes\mathcal{Z}_i + (\Psi_i\Psi_i^\top\mathcal{A}^\top\mathcal{Z}_i)\otimes\mathcal{I} - (\Psi_i Y_i)\otimes\mathcal{I}$,

$\mathcal{I} = (0, I_L)$, I_L is $L\times L$ identity matrix Return to "Identification"

Numeric algorithm III

Collect the FOCs (26)-(27) and the SOCs (28)-(30) into the Newton iteration for (10):

$$x_{k+1} = x_k - \mathcal{H}^{-1}(x_k)\nabla(x_k), \tag{31}$$

where $x_k \stackrel{\text{def}}{=} \begin{pmatrix} \alpha^{(k)} \\ Z^{(k)} \end{pmatrix}$, $\mathcal{H}^{-1}(x_k) \stackrel{\text{def}}{=} \left.\begin{pmatrix} \frac{\partial^2 S(\mathcal{A},Z)}{\partial \alpha^2} & \frac{\partial^2 S(\mathcal{A},Z)}{\partial \alpha \partial Z} \\ \frac{\partial^2 S(\mathcal{A},Z)}{\partial \alpha \partial Z}^{\top} & \frac{\partial^2 S(\mathcal{A},Z)}{\partial Z^2} \end{pmatrix}\right|_{x_k}$,

$$\nabla(x_k) \stackrel{\text{def}}{=} \left.\begin{pmatrix} \frac{\partial S(\mathcal{A},Z)}{\partial \alpha} \\ \frac{\partial S(\mathcal{A},Z)}{\partial Z} \end{pmatrix}\right|_{x_k}$$

Return to "Identification"

Stable vector autoregressive process

VAR(p) process

$$y_t = c + A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t, \quad t \in \mathbb{Z}, \tag{32}$$

where $y_t \in \mathbb{R}^k$ random vector, $A_i \in \mathbb{R}^{k\times k}$ fixed coefficient matrices, $c \in \mathbb{R}^k$ fixed vector of intercept terms, $u_t \in \mathbb{R}^k$ innovation process, $\mathrm{E}\, u_t = 0$, $\mathrm{E}\, u_t u_s^\top = 0$, $s \neq t$, $\Sigma_u \overset{\text{def}}{=} \mathrm{E}\, u_t u_t^\top$ is called *stable* if

$$\det(I_k - A_1 z - \cdots - A_p z^p) \neq 0 \;\; \text{for} \;\; |z| \leq 1,$$

i.e., the reverse characteristic polynomial of (32) has no roots inside and on the complex unit circle Return to "Estimation"

Local linear estimator I

The local linear estimator obtained from, see Ruppert and Wand (1994):

$$\min_{\alpha \in \mathbb{R}, \beta \in \mathbb{R}^p} \sum_{i=1}^{n} \left\{ Y_i - \alpha - \beta^\top (X_i - x) \right\}^2 K_H(X_i - x), \tag{33}$$

where $(X_i^\top, Y_i)^\top$, $i = 1, \ldots, n$ are iid random vectors, $Y_i \in \mathbb{R}$, $X_i \in \mathbb{R}^d$; H is a $d \times d$ bandwidth matrix, K d-variate kernel; $\int K(u)du = 1$, $K_H(u) \stackrel{\text{def}}{=} |H|^{-1/2} K(|H|^{-1/2} u)$

▸ Return to "Confidence bands: Construction I"

Local linear estimator II

The local linear estimator has the form:

$$\widehat{m}(x;H) = e_1^\top (X_x^\top W_x X_x)^{-1} X_x^\top W_x Y, \tag{34}$$

$Y \stackrel{\text{def}}{=} (Y_1, \ldots, Y_n)^\top$, $W_x \stackrel{\text{def}}{=} \operatorname{diag}\{K_H(X_1 - x), \ldots, K_H(X_n - x)\}$,
e_1 first column of a $(d+1) \times (d+1)$ identity matrix, and

$$X_x \stackrel{\text{def}}{=} \begin{bmatrix} 1 & (X_1 - x)^\top \\ \vdots & \vdots \\ 1 & (X_n - x)^\top \end{bmatrix}$$

Return to "Confidence bands: Construction I"

Asymptotics

Under the conditions (A1)-(A3) from Ruppert and Wand (1994) it holds

$$\begin{aligned} &\mathsf{E}\{\widehat{m}(x;H) - m(x)|X_1,\dots,X_n\} \\ &\qquad = 0.5\mu_2(K)\mathrm{tr}\{H\mathcal{H}_m(x)\} + \mathcal{O}_P\{\mathrm{tr}(H)\} \end{aligned}$$

and

$$\begin{aligned} &\mathrm{Var}\{\widehat{m}(x;H) - m(x)|X_1,\dots,X_n\} \\ &\qquad = \{n^{-1}|H|^{-1/2}\|K\|_2^2/f_X(x)\}v(x)\{1+\mathcal{O}_P(1)\}, \end{aligned}$$

where $v(x) \stackrel{\text{def}}{=} \sigma^2(x)$, $\mathcal{H}_m(x)$ the Hessian of a "sufficiently smooth" d-variate function at x, $\mu_2(K)$ such that $\int uu^\top K(u)du = \mu_2(K)I_d$; $m(x) \stackrel{\text{def}}{=} \mathsf{E}(Y|X=x)$

Return to "Confidence bands: Construction II"

Kernel smoother

Return to "Uniform confidence bands I"

Variance function estimator

▸ Return to "Confidence bands: Bootstrap II"