A comprehensive look at commodity volatility forecasting

Loïc Maréchal

University of Neuchâtel

December 19, 2019

Motivation

- High frequency Realized Volatility (RV) measures deliver better forecasts than parametric models (Andersen and Bollerslev, 1998a).
  - RV as dependent variable improves forecast accuracy of other models (stoch. vol. and (G)ARCH).
  - RV as predictor additionally improves the forecast power of many models.
  - Non-parametric simple measure. We just need high frequency data.
- If RV modeling works for many assets (stocks, currencies, bonds), does it still hold for commodities?
  - Commodity volatility stylized facts depart from the rest of the assets.
  - Identified commodity volatility factors are unique.
  - RV has not been tested on commodity futures (to my knowledge).

Contribution

- Examine three key commodity volatility factors with high frequency (RV) measures.
- Add two standard RV models (HAR-RV and VAR-RV) to improve out of sample forecast power.
- Deliver many potential risk management applications for commodity futures. I provide a methodology for optimal hedging ratio determination with RV and high frequency covariances.

Results

- Three commodity volatility factors are significant determinants of the commodity RV for the nine commodities tested.
- Explanatory power and out of sample performance increases further in HAR-RV and VAR-RV models.
- The optimal hedging ratio forecast computed with high frequency data is superior to alternative methods

- RV as dependent variable.
  - Andersen and Bollerslev (1998b): regress RV on the forecast of a GARCH(1,1) in Mincer-Zarnowitz (Mincer and Zarnowitz, 1969) regressions vastly improves the R^2.
  - Better because as the sampling frequency of RV increases to ∞, it tends to the integrated volatility (IV). An AR(1)-RV yields a better forecast than squared returns (Andersen et al., 2007.)
- RV as independent variable.
  - Corsi (2009): Heterogeneous AutoRegressive RV (HAR-RV), i.e. past RV averaged over different intervals.
  - Andersen et al. (2003) find that a VAR using RV performs better than GARCH-type models.
- Realized Betas (Andersen et al., 2006). More stable and less forecast error.

Literature review: Commodities

- Stylized fact of commodities: positive skewness (Gorton and Rouwenhorst, 2006), "leverage effect" (Carpantier, 2010) that is inverse relative to stocks and bonds (Black, 1976).

- Three commodity volatility factors.
  - Theory of storage (Brennan, 1958; Kaldor, 1939; Working, 1933): inverse and convex relationship between inventories and volatility. Haugom et al. (2014) extend it: both low and high inventories lead to high volatility.
  - Samuelson hypothesis (Samuelson, 1965, 1976): inverse relationship between time to maturity and volatility.
  - Uncertainty resolution (Anderson and Danthine, 1983): e.g. volatility shock when the crop supply is known. Seasonal components as proxy.

Data

- Minute prices for nine commodity futures' term structure of three sectors obtained from the Barchart Java API. From May 2008 to Jan 2019.
- Following the literature standard methodology, contracts are rolled 10 business days before maturity.

Data - Contracts

Ticker	Trading venue	Underlying	Unit	Maturity
C	CBT	Corn	bu (5,000)	HKNUZ
S	CBT	Soybeans	bu (5,000)	FHKNQUX
W	CBT	Chicago wheat	bu (5,000)	HKNUZ
CL	NYMEX/ICE	WTI crude oil	bbl (1,000)	FGHJKMNQUVXZ
HO	NYMEX	Heating oil	gal (42,000)	FGHJKMNQUVXZ
NG	NYMEX/ICE	Natural gas	MMBtu (10,000)	FGHJKMNQUVXZ
GC	CMX	Gold	oz (100)	GJMQVZ
HG	COMEX	Copper	lb (25,000)	FGHJKMNQUVXZ
SI	CMX	Silver	oz (5,000)	FHKNUZ

Maturity month code: F = January, G = February, H = Mars, J = April, K = May, M = June, N = July, Q = August, U = September, V = October, X = November, Z = December.

- Computation of RV.

$$RV_t(\Delta) = \sum_{j=1}^{1/\Delta} \left(r^2_{\Delta, t+j\times\Delta} \right),$$

where $1/\Delta$ is the number of observations in one day and r^2_Δ represents the squared returns sampled at the period Δ.

- Scaled basis (annualized).

$$SL^m_{c,t} = \frac{f^{m+1}_{c,t} - f^m_{c,t}}{\delta(m)} \times 250,$$

where $\delta(m)$ indicates the time difference in days between the two consecutive maturities.

- Conditional slope.

$$SL^+_{c,t} = max(SL_{c,t}, 0) \text{ and } SL^-_{c,t} = min(SL_{c,t}, 0)$$

		squared returns		*RV*	
Ticker (contract)	Contango %	AR(1)	t-statistic	AR(1)	t-statistic
C (corn)	14	0.03	1.36	0.42	24.02
CL (WTI crude oil)	14	0.29	15.56	0.81	72.13
GC (gold)	12	0.18	9.41	0.60	38.32
HG (copper)	35	0.31	16.79	0.64	42.52
HO (heating oil)	22	0.18	9.28	0.78	64.31
NG (natural gas)	12	0.18	9.58	0.48	27.97
S (soybeans)	43	0.04	1.89	0.43	24.86
SI (silver)	8	0.18	9.48	0.48	27.80
W (wheat)	3	0.09	4.59	0.48	27.89

Commodity factors are good RV predictors (1)

$$RV_{c,t} = \alpha_{1,c} \times SL_{c,t-1} + \alpha_{2,c} \times DTM_{c,t-1} + \sum_{i=3}^{i=14} \alpha_{i,c} \times DM_{i,t-1} + \epsilon_{c,t}$$

	(C)	(CL)	(GC)	(HG)	(HO)	(NG)	(S)	(SI)	(W)
	$RV_{c,t}$								
	Panel A: continuous term structure								
term structure	0.23	11.82***	−12.79***	8.84***	5.97***	2.56***	0.26	−1.12	6.14***
	(0.59)	(37.76)	(−6.06)	(6.32)	(11.00)	(11.05)	(1.05)	(−0.58)	(8.43)
dtm $\times 10^{-7}$	37.60***	14.54***	24.16***	37.50***	20.08***	29.09***	26.09***	42.36***	27.01***
	(23.56)	(8.88)	(25.20)	(24.78)	(12.85)	(12.92)	(20.80)	(22.59)	(20.74)
seasonality (+)	Jul***	Sep***	Sep***	Oct***	Feb***	Jan***	Aug***	Nov***	Aug***
Observations	2,653	2,657	2,639	2,648	2,657	2,657	2,653	2,580	2,653
Adjusted R^2	0.80	0.85	0.79	0.77	0.82	0.83	0.81	0.75	0.88
Note:									*p<0.1; **p<0.05; ***p<0.01

Commodity factors are good RV predictors (2)

$$RV_{c,t} = \alpha^+_{1,c} \times SL^+_{c,t-1} + \alpha^-_{1,c} \times SL^-_{c,t-1} + \alpha_{2,c} \times DTM_{c,t-1} + \sum_{i=3}^{i=14} \alpha_{i,c} \times DM_{i,t-1} + \epsilon_{c,t}$$

	(C)	(CL)	(GC)	(HG)	(HO)	(NG)	(S)	(SI)	(W)
					$RV_{c,t}$				
					Panel B: conditional term structure				
contango	6.09***	12.34***	31.34***	36.95***	17.82***	4.31***	12.79***	10.54***	6.60***
	(4.57)	(37.53)	(9.68)	(12.74)	(22.51)	(17.48)	(6.02)	(4.25)	(8.71)
backwardation	−1.29**	3.38*	−47.00***	−6.85***	−5.54***	−5.51***	−0.32	−26.51***	−13.64
	(−2.52)	(1.95)	(−16.72)	(−3.46)	(−7.12)	(−10.04)	(−1.21)	(−6.75)	(−1.51)
dtm $\times 10^{-7}$	37.72***	15.86***	24.37***	35.76***	15.84***	28.73***	28.55***	40.98***	27.05***
	(23.72)	(9.60)	(26.82)	(24.03)	(10.72)	(13.37)	(21.74)	(21.97)	(20.79)
seasonality (+)	Jul***	Nov***	Oct***	Oct***	Feb***	Jan***	Aug***	Nov***	Aug***
Observations	2,653	2,657	2,639	2,648	2,657	2,657	2,653	2,580	2,653
Adjusted R^2	0.80	0.85	0.81	0.78	0.84	0.85	0.81	0.76	0.88

Note: *p<0.1; **p<0.05; ***p<0.01

HAR-RV

$$RV_{c,t} = \beta_{1,c} \times RV^{D}_{c,t-1} + \beta_{2,c} \times RV^{W}_{c,t-1} + \beta_{3,c} \times RV^{M}_{c,t-1} + \boldsymbol{\beta_{v,c}} \times \boldsymbol{X_{c,t-1}} + \epsilon_{c,t}$$

	(C)	(CL)	(GC)	(HG)	(HO)	(NG)	(S)	(SI)	(W)
					$RV_{c,t}$				
$RV^{D}_{c,t-1}$	0.22***	0.29***	0.14***	0.08***	0.27***	0.10***	0.14***	0.13***	0.13***
	(9.77)	(12.46)	(6.38)	(3.56)	(11.91)	(4.15)	(6.15)	(5.66)	(5.79)
$RV^{W}_{c,t-1}$	0.07**	0.31***	0.17***	0.39***	0.23***	0.24***	0.08**	0.24***	0.28***
	(2.13)	(8.50)	(4.12)	(8.95)	(6.71)	(6.55)	(2.30)	(5.77)	(7.26)
$RV^{M}_{c,t-1}$	0.06	0.30***	0.46***	0.38***	0.38***	0.30***	0.39***	0.32***	0.28***
	(1.46)	(8.28)	(11.51)	(9.26)	(12.11)	(7.40)	(9.43)	(7.26)	(6.28)
Observations	2,649	2,657	2,623	2,639	2,657	2,657	2,649	2,528	2,649
Adjusted R^2	0.81	0.93	0.87	0.87	0.93	0.87	0.84	0.81	0.90

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

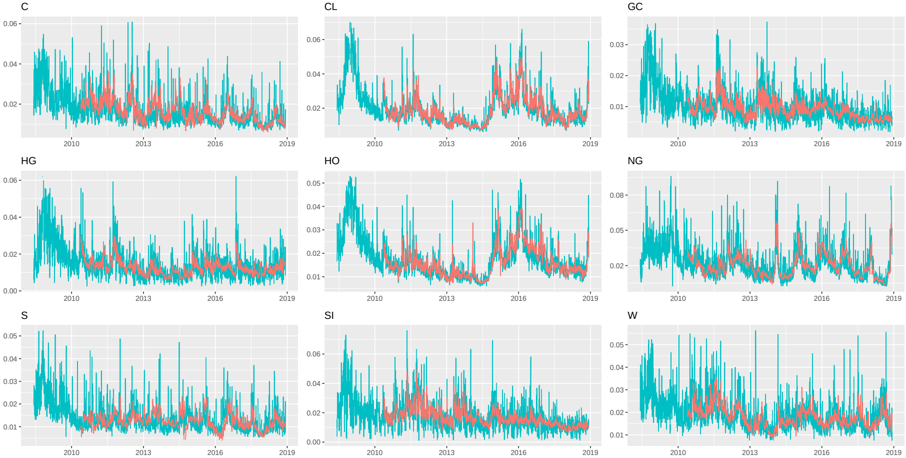
VAR-HAR-RV

$$RV_{c,t} = \gamma_{1,c} \times RV^{agriculture}_{c,t-1} + \gamma_{2,c} \times RV^{energy}_{c,t-1} + \gamma_{3,c} \times RV^{metal}_{c,t-1} + \boldsymbol{\gamma_{v,c}} \times \boldsymbol{Z_{c,t-1}} + \epsilon_{c,t}$$

	(C)	(CL)	(GC)	(HG)	(HO)	(NG)	(S)	(SI)	(W)
					$RV_{c,t}$				
$RV^{D,agriculture}_{t-1}$	0.29***	0.03*	−0.01	0.01	0.02	−0.01	0.05*	0.04	0.01
	(5.67)	(1.84)	(−0.95)	(0.73)	(1.00)	(−0.42)	(1.71)	(1.28)	(0.20)
$RV^{D,energy}_{t-1}$	0.04*	−0.01	0.01	0.01	−0.01	0.10**	0.02	−0.05**	0.01
	(1.88)	(−0.34)	(0.75)	(0.46)	(−0.60)	(2.37)	(1.47)	(−2.18)	(0.56)
$RV^{D,metal}_{t-1}$	0.12***	0.10***	0.06***	0.11***	0.07***	0.02	0.08***	0.28***	0.06***
	(4.66)	(5.19)	(3.43)	(4.26)	(4.65)	(0.74)	(4.11)	(4.74)	(3.20)
Observations	2,649	2,653	2,619	2,635	2,653	2,653	2,649	2,524	2,649
Adjusted R^2	0.82	0.93	0.87	0.87	0.94	0.87	0.84	0.81	0.90

Note: *p<0.1; **p<0.05; ***p<0.01

VAR-HAR-RV rolling forecast and RV



VAR-HAR-RV out of sample tests

	C	CL	GC	HG	HO	NG	S	SI	W
	Panel A: Mincer-Zarnowitz F-test, MSE and MAE								
GARCH(1,1)									
MZ F-test	260.83	48.64	65.56	41.70	61.01	34.29	335.54	29.09	49.78
MAE %	0.52	0.40	0.29	0.42	0.34	0.67	0.43	0.59	0.47
MSE $\times 10^{-4}$	0.86	0.32	0.16	0.34	0.29	1.34	0.43	0.71	0.47
VAR-HAR-RV									
MZ F-test	26.86	5.04	10.01	15.12	12.46	7.92	28.11	8.90	8.10
MAE %	0.39	0.30	0.26	0.37	0.27	0.54	0.32	0.53	0.40
MSE $\times 10^{-4}$	0.38	0.21	0.14	0.25	0.18	0.86	0.27	0.59	0.40
	Panel B: Modified Diebold-Mariano test								
VAR-HAR-RV $>$ GARCH(1,1)									
HLN-DM(1) F-test	8.32	7.80	2.20	3.01	4.34	5.98	-1.01	0.28	3.17
	[0.00]	[0.00]	[0.01]	[0.00]	[0.00]	[0.00]	[0.84]	[0.39]	[0.00]
HLN-DM(2) F-test	4.07	5.66	1.75	1.90	3.02	4.60	0.11	0.28	-0.21
	[0.00]	[0.00]	[0.04]	[0.03]	[0.00]	[0.00]	[0.46]	[0.39]	[0.58]

RV - Realized beta application for OHR

- Ederington (1979) defines the optimal hedging ratio (OHR) as,

$$OHR_t = \frac{\mathrm{Cov}\left(r^s_{c,t}, r^f_{c,t}\right)}{\mathrm{Var}\left(r^f_{c,t}\right)}$$

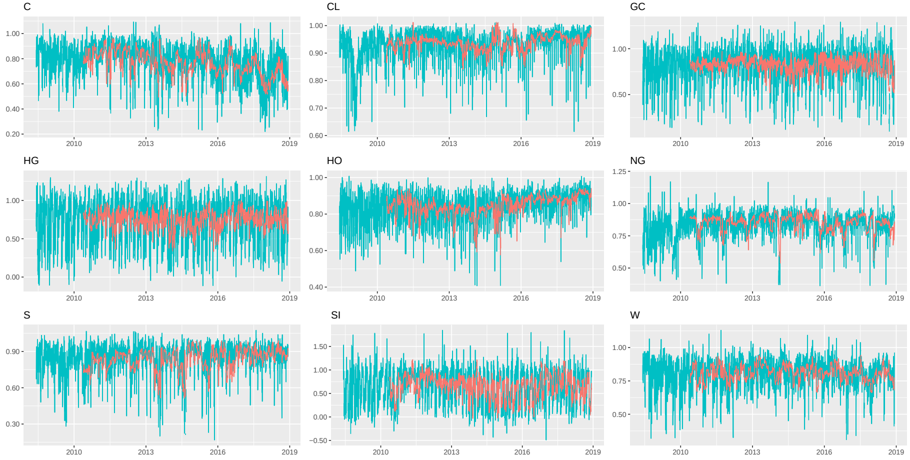
- Using the realized beta methodology of Andersen et al. (2006),

$$ROHR_t = \frac{\mathrm{RCov}\left(r^s_{c,t}, r^f_{c,t}\right)}{\mathrm{RVar}\left(r^f_{c,t}\right)}$$

- Which becomes,

$$\widehat{ROHR}_t = \frac{\widehat{\mathrm{RCov}}_{\mathrm{t}}\left(r^s_{c,t}, r^f_{c,t}\right)}{\widehat{\mathrm{RVar}}_{\mathrm{t}}\left(r^f_{c,t}\right)}$$

VAR-HAR-RV rolling forecast and ROHR (RBeta)



Conclusion and future research

- As for stocks and bonds, RV is a good predictor or future volatility (and optimal hedging ratio).
- In the light of RV, traditional commodity volatility factors differ:
  - It is the magnitude of the slope and not its sign which matters.
  - The Samuelson hypothesis is rejected, I find the opposite: The more time to maturity, the more volatility!
  - There is a strong seasonal component in the RV determination.
- Research extension:
  - Economic valuation of the volatility: can RV extended model forecast implied volatility?
  - Does the β stability mitigates the transaction costs involved in an optimal commodity futures hedging?

Thank you!

loic.marechal@unine.ch