Does Anything Beat 5-Minute RV?

A Comparison of Realized Measures Across a Panel of Assets

Lily Liu
Duke

Andrew Patton
Duke

Kevin Sheppard
Oxford

October 2012

Description of the problem

- Accurate estimates of return volatility are used in many applications: derivatives pricing, asset allocation, risk management, etc.
- The last 15 years has witnessed a profusion of new and improved ways to estimate volatility using high frequency data: **"realized measures"**
  - Realized volatility, Two-scales realized volatility, Realized kernels, Realized range, etc.
- The range of assets for which we have high frequency data is growing
  - US equities, international equities, FX, fixed income securities, commodities, derivatives

★ How should we choose a realized measure for a given data set? Is there one that works best in all/most applications?

Objectives of this project

1. Do any of the new, sophisticated, estimators significantly out-perform a simple realized volatility computed using 5-minute data?
   - 5-min RV turns out to be hard, but not impossible, to beat.
2. What are the characteristics of a "good" realized measure?
   - Sampling frequency, sampling scheme, functional form, etc.
3. Are there patterns in the performance of different realized measures across asset classes?
   - Does one estimator work better for equities, and another for FX?

Answering our research questions: Lots of data

- **Realized measures:** We compute around 350 different measures of daily volatility, across six different classes of realized measures
  - Realized volatility, autocorrelation-adjusted RV, two-scales RV, multi-scales RV, realized kernels, realized range, quantile RV
- **Asset returns:** 31 different asset price series across five asset classes
  - Individual equities (high and low liquidity), computed equity indices, exchange rates, interest rates, index futures
- **Sample period:** January 2000 to December 2010, so $T \approx 2700$ days.
- **Sampling frequencies:** From 1 second to 15 minutes, so we use $n \approx\in [25, 25000]$ intra-daily observations.

Outline of the presentation

1. The realized measures under analysis (brief)
2. Methods for comparing realized measures
3. Main results:
    1. Guidelines on sampling frequency, sampling scheme, etc
    2. Does anything beat 5-min RV?
    3. The *set* of best realized measures
    4. Out-of-sample forecast comparisons
4. Summary and conclusions

Quadratic variance of a price process

- Consider a general jump-diffusion model for the log-price of an asset:

$$dp(t) = \mu(t)\,dt + \sigma(t)\,dW(t) + \kappa(t)\,dN(t)$$

  - μ is the drift, σ is the (stochastic) volatility, W is a B.M., κ is the jump size, and N is a counting measure for the jumps.

- Quadratic variation over the period $[t, t+1]$ is:

$$\begin{aligned} QV_{t+1} &\equiv \operatorname*{plim}_{n\to\infty} \sum_{j=1}^{n} r_{t+j/n}^2 \\ \text{where} \quad r_{t+j/n} &\equiv p_{t+j/n} - p_{t+(j-1)/n} \end{aligned}$$

- Realized variance (RV) is the sample analog of QV:

$$RV_{t+1} \equiv \sum_{j=1}^{n} r_{t+j/n}^2$$

Sampling frequency, sampling scheme

- **Sampling frequencies:** 1 sec, 5 sec, 1 min, 5 min, 15 min
- **Sampling schemes:**
  1. *Calendar time:* Sample prices every m minutes
  2. *Tick time:* Sample prices every s observations
- **Sub-sampling:** use all possible "grids" of prices if sampling lower than 1 second
- **Price series:** Transaction prices or mid-quotes
- **Total:** 42 versions of each realized measure.

Classes of realized measures I

- New realized measures have been proposed to provide robustness to various types of market microstructure effects (bid-ask bounce, stale quotes, mis-reported prices) and to improve the efficiency of estimates of volatility.
- We consider six broad classes of realized measures.

1. **Realized volatility:** simple sum of squared high-frequency returns
2. **RV with optimal sampling (RVbr):** Bandi and Russell (2008, REStud)
3. **Autocorrelation adjusted RV (RVac1):** Like RV, but incorporates possible first-order autocorrelation in high frequency returns. French, Schwert and Stambaugh (1987, JFE), Zhou (1996, JBES), Hansen and Lunde (2006, JBES)

Classes of realized measures II

4. **Two-scales and Multi-scales RV (TSRV, MSRV):** Use a combination of high and lower frequencies to estimate the volatility and the noise (to remove it). Zhang, Mykland and Aït-Sahalia (2005, JASA) and Zhang (2006, Bernoulli)

5. **Realized kernels (RK):** Generalization of RVac1 to handle more lags and various shapes of autocorrelation function, Barndorff-Nielsen, Hansen, Lunde and Shephard (2011, Ecta)

6. **Maximum-likelihood RV (MLRV):** Uses maximum-likelihood estimation, assuming MA(1) structure for observed returns to account for MMS noise, Aït-Sahalia, Mykland, and Zhang (2005, RFS)

7. **Realized range RV (RRV):** Uses sum of squared high-low ranges for intra-daily periods rather than sum of squared returns, Christensen and Podolskij (2007, JoE)

- **Total**: 398 realized measures per asset.

Jump-robust realized measures I

- In the forecasting application we will also consider some "jump robust" estimators of volatility

$$\begin{aligned} dp(t) &= \mu(t)\,dt + \sigma(t)\,dW(t) + \kappa(t)\,dN(t) \\ QV_t &= \underbrace{\int_{t-1}^{t} \sigma^2(\tau)\,d\tau}_{IV_t} + \underbrace{\sum_{t-1<\tau\leq t} \kappa^2(\tau)}_{JV_t} \end{aligned}$$

- We consider four classes of jump-robust realized measures.

1. **Bi-power variation (BPV):** Sum of adjacent absolute returns, Barndorff-Nielsen and Shephard (2004, JFEC)

Jump-robust realized measures II

2. **Quantile-based RV (QRV):** Using relation between quantile and volatility to get new estimator, Christensen, Oomen and Podolskij (2010, JoE)

3. **Nearest neighbor truncated RV:** The MinRV and MedRV estimators use min or median of blocks of 2 or 3 returns, Andersen, Dobrev and Schaumburg (2008, JoE)

4. **Truncated RV (TRV):** Sum of squared returns, truncating "large" returns, Mancini (2001, 2009, Scan. J. Stats)

- **Total:** In the forecasting application we have a total of 398+206=604measures of asset price volatility.

Outline of the presentation

1. The realized measures under analysis (brief)
2. **Methods for comparing realized measures**
3. Main results:
    1. Guidelines on sampling frequency, sampling scheme, etc
    2. Does anything beat 5-min RV?
    3. The *set* of best realized measures
    4. Out-of-sample forecast comparisons
4. Summary and conclusions

Comparing realized measures

- We compare competing realized measures using two approaches:

1. **Forecast accuracy**, when combined with a simple and widely-used volatility forecasting model (the "HAR" model of Corsi, 2009). Horizons from 1 to 50 days.
2. **Estimation accuracy**, for the latent quadratic variation on a given day, using the method of Patton (2011, JoE). Notation:

$$\begin{aligned} \text{True QV} &= QV_t \\ \text{"Proxy" for QV} &= \widetilde{QV}_t \\ \text{Forecast of QV} &= \widehat{QV}_{t+h|t} \\ \text{Realized measure} &= M_{it} \end{aligned}$$

Accuracy of a realized measure

- In all cases we need to choose a penalty for error in the realized measure. Two common choices:

$$\text{MSE} \quad L(\theta, M) = (\theta - M)^2$$

$$\text{QLIKE} \quad L(\theta, M) = \frac{M}{\theta} - \log \frac{M}{\theta} - 1$$

- We focus on QLIKE as it has better power properties.
  - The fact that is relies only on the ratio (M/θ) provides some automatic normalization, which is helpful.

Comparing estimation accuracy I

- We use the data-based ranking method for realized measures proposed in Patton (2011, JoE) to compare estimation accuracy.

- This method overcomes the fact that QV is unobservable, even ex post, by using a (finite-sample) *unbiased proxy* for QV. i.e., one that satisfies

$$E\left[\widetilde{QV}_t | \mathcal{F}_{t-1}, QV_t\right] = QV_t$$

  - Examples of such a proxy: daily RV, 15-min RV, 5-min RV. The proxy can be noisy, but must be reasonably assumed to be *unbiased* (so unaffected by microstructure effects)

Comparing estimation accuracy II

- Then, exploiting the fact that QV is very persistent from day to day, we use a one-period lead of the low-freq RV to break the dependence between the proxy error and the error in the realized measures under analysis, so

$$Cov\left[\left(\widetilde{QV}_t - QV_t\right) , \left(M_{it} - QV_t\right) \ | \mathcal{F}_{t-1}, QV_t \right] = 0$$

- Finally, we use a loss function L that is "robust" to the use of a noisy proxy
  - There are many such loss functions. MSE and QLIKE are two examples.

Comparing estimation accuracy III

- Then, we can show that

$$E[L(\widetilde{QV}_t, M_{it})] \lessgtr E[L(\widetilde{QV}_t, M_{jt})] \Leftrightarrow E\left[L\left(QV_t, M_{it}\right)\right] \lessgtr E\left[L\left(QV_t, M_{jt}\right)\right.$$

- The ranking on the RHS is infeasible, but we can estimate the ranking on the LHS, and under standard (long-span) assumptions:

$$\sqrt{T}\left(\frac{1}{T}\sum_{t=1}^{T} L(\widetilde{QV}_t, M_{it}) - E\left[L\left(QV_t, M_{it}\right)\right]\right) \xrightarrow{d} N\left(0, \Omega\right)$$

- This enables us to use existing methods for comparing forecasts:
  - **Pair-wise comparisons**: Diebold-Mariano (1995, JBES), West (1996, Ecta), Giacomini-White (2006, Ecta)
  - **Multiple comparisons**: White (2000, Ecta), Romano-Wolf (2005, Ecta), Hansen, Lunde and Nason (2011, Ecta)

Comparing *forecast* accuracy

- We can also compare realized measures through the accuracy of forecasts based on them
- This of course requires a foreasting *model*, and we use the "heterogeneous autoregressive" (HAR) model of Corsi (2009, JFEC):

$$\widetilde{QV}_{t+h} = \beta_{0h} + \beta_{Dh} M_t + \beta_{Wh} \frac{1}{5} \sum_{k=0}^{4} M_{t-k} + \beta_{Mh} \frac{1}{22} \sum_{k=0}^{21} M_{t-k} + \varepsilon_t$$

- This model relates QV at period $t+h$ to the realized measure over the most recent 22 observations, breaking these into three components (daily, weekly and monthly)
  - This captures "long memory"-like effects, but is simpler to estimate

Outline of the presentation

1. The realized measures under analysis (brief)
2. Methods for comparing realized measures
3. **Main results:**
    1. Guidelines on sampling frequency, sampling scheme, etc
    2. Does anything beat 5-min RV?
    3. The *set* of best realized measures
    4. Out-of-sample forecast comparisons
4. Summary and conclusions

Data description

- 31 assets:
  1. Individual stocks (US and UK)
  2. FX futures
  3. Interest rate futures
  4. Equity index futures
  5. Computed equity indices
- Sample Period: Jan 2000 – Dec 2010, T $\approx$ 2700 days
- Transaction prices and quote prices
- Data source: Thomson Reuter's Tick History
- Data are cleaned using the results in Barndorff-Nielsen, Hansen, Lunde and Shephard (2009, EJ)

Data description

Col 1: Avg vol (%), Col 2, 3: Avg seconds b/w Trades, Quotes

Individual Equities			
KO	19	7.6	2.6
SYY	22	12.5	3.4
IFF	24	26.6	5.4
MSFT	25	2.7	1.5
LSI	49	15.6	3.8
DGE	24	15.8	3.6
SAB	28	23.6	3.8
VOD	29	7.0	2.3
RSA	39	28.1	6.4
SDR	46	52.4	8.7
Index futures			
JNI	15	3.5	0.9
ES	15	0.5	0.2
FFI	16	1.9	0.5
STXE	18	2.0	0.7
FDX	18	1.5	0.8

Interest Rate Futures			
TU (2yr)	1.4	7.6	0.5
FV (5yr)	3.5	3.0	0.3
TY (10yr)	5.2	1.9	0.3
US	8.1	2.4	0.4
FGBS	1.3	9.0	1.9
FGBL	4.6	2.7	1.0
Currency futures			
BP	6.7	2.9	0.4
URO	6.9	1.4	0.3
JY	7.3	3.1	0.4
CD	8.4	4.1	0.6
AD	9.3	4.9	0.5
Computed Indices			
N225	15	48.1	
SPX	16	15.9	
FTSE	16	4.9	
STOXX50E	19	15.2	
DAX	19	2.9	

Plain, simple ranking of realized measures

- We implement the 398 realized measures, and using the methods above we can obtain a ranking based on average, unconditional, accuracy.
- We present below the top 5 for each asset class, averaging ranks within asset classes
  - Rank correlation within asset classes are:

Individual stocks	: 0.67
FX futures	: 0.87
Bond futures	: 0.85
Equity index futures	: 0.75
Computed equity indices	: 0.84

Top 5 estimators for each asset class

Indiv. Equities	Int. Rate Fut	Currency Fut
RKth2, 5s tick, mq	RRVm5, 5s tick, mq	TSRV, 1s cal, tr, ss
RKbart, 5s tick, mq	RRVm5, 5s tick, mq, ss	TSRV, 1s cal, tr
RKnfp, 1s tick, mq	RRVm10, 1s cal, mq, ss	MSRV, 1s tick, mq, ss
RKbart, 1s tick, mq	RRVm10, 1s cal, mq	MLRV, 1s cal, mq
RKnfp, 1s tick, tr	RRVm10, 1s tick, mq, ss	MLRV, 1s cal, mq, ss

Index Fut	Comp. Index
RV, 1m tick, tr, ss	RVac1, 1m tick
RVac1, 1m tick, tr, ss	RVac1, 1m cal
RV, 1m tick, tr	RKth2, tick-by-tick
MSRV, 5s cal, tr, ss	RKcub, tick-by-tick
RKbart, 1s cal, tr	RKbart, 1m tick

Pair-wise comparisons of realized measures

- We now try to understand the characteristics of a "good" realized measure. We compare them on three binary dimensions:
  1. Calendar-time vs. Tick-time sampling
  2. Transaction prices vs. Quote prices
  3. Sub-sampled vs. not
- The tables below present the proportion (across 31) assets of t-statistics for these comparisons are significantly positive minus the proportion that are significantly negative. (Negative values favor the first approach)

Calendar-time vs. Tick-time sampling

Calendar time preferred for higher frequencies; Tick time for lower frequencies

	1s	5s	1m	5m	15m
RV	-84	-74	0	23	35
RVac1	-84	-68	-3	42	29
RK	-13	6	48	48	39
MSRV	-45	-19	29	42	32
TSRV	35	-52	-3	32	42
MLRV	-81	-45	6	42	19
RRV	-61	3	55	77	81
BR	0				

Transaction prices vs. Quote prices

Transaction prices generally preferred

	1t	1s	5s	1m	5m	15m
RV	73	73	-19	-62	-81	-81
RVac1	-38	-4	-38	-54	-42	-42
RK	-4	27	-15	-65	-96	-88
MSRV	-42	-23	-4	77	50	-23
TSRV	-65	-92	-42	-12	27	4
MLRV	-31	77	23	-50	-46	4
RRV	-23	69	19	23	0	23
BR	8					

Not sub-sampled vs. Sub-sampled

Not-subsampled preferred for high frequencies; Sub-sampling helps for lower frequencies

	1s	5s	1m	5m	15m
RV	3	6	6	29	52
RVac1	-58	-39	29	84	94
MSRV	-3	0	10	19	0
TSRV	0	0	16	55	35
MLRV	0	3	6	65	77
RRV	0	-26	-19	-42	-58
BR	6				

Does anything beat 5-minute RV?

- We attempt to answer the question in the title of the paper
- We want to compare simple 5min RV with *all* of the 397 other realized measures, controlling for the fact that we are doing multiple comparisons
- We do so using the step-wise testing method of Romano and Wolf (2005, Ecta), which builds on the "reality check" of White (2000, Ecta).
- This approach considers the $S = 397$ hypotheses

$$H_0^{(s)} : E\left[L\left(\theta_t, M_{0t}\right)\right] = E\left[L\left(\theta_t, M_{st}\right)\right], \quad s = 1, 2, ..., S$$

and identifes the subset of these than can be rejected, controlling the family-wise error rate.

Table 5.
Number of estimators that are significantly different from RV5min in Romano-Wolf Tests

Proxy:	**Worse**					**Better**					**Total Estimators**
	RV Daily	*RV 15min*	*RV 1min*	*MSRV 1min*	*RKth2 1min*	*RV Daily*	*RV 15min*	*RV 1min*	*MSRV 1min*	*RKth2 1min*	
KO	161	231	219	240	237	0	0	0	0	0	396
LSI	160	265	257	272	278	0	0	0	0	0	395
MSFT	243	285	272	288	290	0	0	0	0	0	396
IFF	127	238	254	259	252	0	0	0	0	0	391
SYY	129	210	206	190	190	0	0	0	0	0	392
DGE	157	318	335	231	247	0	0	0	0	0	398
VOD	179	279	351	211	212	0	0	0	0	0	397
SAB	126	322	278	312	316	0	0	0	0	0	398
SDR	116	301	295	274	277	0	0	0	0	0	394
RSA	141	291	362	165	202	0	0	0	0	0	397
TU	204	180	194	166	187	0	0	0	0	0	397
FV	192	237	220	221	236	0	0	0	0	0	398
TY	188	229	213	211	225	0	9	24	28	23	398
US	202	247	241	243	254	0	0	0	0	0	397
FGBL	183	269	266	267	268	0	0	0	0	0	398
FGBS	310	367	131	363	343	0	0	0	0	0	398
CD	120	177	178	177	178	0	0	0	0	0	398
AD	102	171	173	180	181	0	0	0	0	0	398
BP	134	166	170	165	166	0	0	0	0	0	398
URO	149	167	172	172	172	0	0	0	0	0	398
JY	139	172	178	175	172	0	0	0	0	0	398
STXE	177	60	183	280	284	0	0	0	0	0	398
JNI	250	324	331	317	318	0	0	0	0	0	394
FDX	142	145	145	182	181	0	0	0	0	0	398
FFI	150	183	182	184	185	0	0	0	0	0	398
ES	159	204	204	204	206	0	0	0	0	0	398
SPX	156	169	169	155	163	0	0	0	7	1	199
STOXX50E	123	170	168	143	166	0	0	0	0	0	199
DAX	122	148	155	147	152	0	0	0	0	0	199
FTSE	153	175	172	129	169	0	0	0	0	0	199
N225	143	159	161	161	160	0	0	0	0	0	197

Note: Results from when a potential proxy has significantly different mean from RVdaily are displayed in lighter color.

Does anything beat 5min RV?

- Many, many estimators are significantly worse than 5min RV. Very few (≈ 0) are significantly better.
- Is this a problem of power?
  - The fact that many are rejected as *worse* is reassuring
  - We also try with more accurate proxies (RV15min, RV5min) and find little difference
  - We also try Hansen's (2005) refinement of the reality check, designed to boost power, and find no change
- Beyond the answer to the question, we can use these results to gain further insights into these measures:

Proportion of measures that are signif worse than RV5min

All 31 assets

All 31 Assets

	1t	1s	5s	1m	5m	15m
RV	70	55	39	18	17	71
RVac1	30	41	27	19	49	73
RK	11	15	18	50	87	91
MSRV	21	24	13	43	93	87
TSRV	75	39	71	97	98	96
MLRV	28	38	22	22	84	78
RRV	25	35	27	22	66	95
BR	18					

Proportion of measures that are signif worse than RV5min

Individual equities

Individual Equities

	1t	1s	5s	1m	5m	15m
RV	65	62	54	25	0	40
RVac1	40	58	39	14	29	60
RK	0	14	3	28	71	86
MSRV	20	44	17	13	84	84
TSRV	80	49	61	91	96	92
MLRV	30	58	32	0	71	61
RRV	25	50	37	5	39	86
BR	11					

Proportion of measures that are signif worse than RV5min

Interest rate futures

Interest Rate Futures

	1t	1s	5s	1m	5m	15m
RV	75	59	33	8	52	100
RVac1	36	41	19	46	81	96
RK	40	18	55	98	99	97
MSRV	25	15	8	96	98	89
TSRV	58	15	58	100	100	94
MLRV	33	34	17	83	100	85
RRV	13	19	13	50	100	98
BR	31					

Proportion of measures that are signif worse than RV5min

FX futures

Currency Futures

	1t	1s	5s	1m	5m	15m
RV	70	36	10	0	0	65
RVac1	0	6	0	0	40	58
RK	0	0	1	41	93	88
MSRV	0	0	0	25	98	78
TSRV	60	24	70	100	98	100
MLRV	0	6	0	0	80	80
RRV	0	0	0	1	58	100
BR	0					

Proportion of measures that are signif worse than RV5min

Equity index futures

Index Futures

	1t	1s	5s	1m	5m	15m
RV	60	44	30	0	23	95
RVac1	10	33	15	10	55	83
RK	10	3	10	64	100	95
MSRV	20	8	5	58	98	93
TSRV	90	49	95	100	100	98
MLRV	20	28	5	15	93	95
RRV	25	28	15	16	79	100
BR	5					

Proportion of measures that are signif worse than RV5min

Computed equity indices

Computed Indices

	1t	1s	5s	1m	5m	15m
RV	100	100	100	80	20	85
RVac1	80	100	100	30	60	85
RK	5	65	45	18	80	98
MSRV	60	100	71	50	95	94
TSRV	100	100	100	100	100	100
MLRV	80	100	100	20	85	85
RRV	100	100	100	75	80	100
BR	75					

Estimating the *set* of best realized measures I

- Taking 5min RV as the estimator under the "null" hypothesis might give it undue preferential treatment
- An alternative method for comparing many realized measures is the "model confidence set" of Hansen, Lunde and Nason (2011, Ecta)
- This method provides the subset of measures that contains the unknown best estimator with some given level of probability
    - It is a natural extension of a confidence interval for a single parameter

Estimating the *set* of best realized measures II

- We apply this method and found that the estimated MCS contains between 3 and 143 realized measures (1% to 40% of all estimators) across the 31 assets.
  - On average, the MCS contained 40 estimators, around 11% of the total
  - Individual equities and equity indices have the largest MCSs (around 17% of all estimators)
  - Equity index futures and interest rate futures have the smallest MCSs (around 5% of all estimators)
- Below we summarize these results by reporting the proportion (across assets) of MCSs that include a given realized measure at a given frequency

Proportion of measures that are in the 90% MCS

All 31 assets

All 31 Assets

	1t	1s	5s	1m	5m	15m
RV	4	3	15	30	18	1
RVac1	7	6	18	27	8	0
RK	18	29	26	6	0	0
MSRV	9	24	21	3	0	0
TSRV	0	11	4	0	0	0
MLRV	9	15	22	15	0	0
RRV	15	11	17	20	2	0
BR	9					

Proportion of measures that are in the 90% MCS

Individual equities

Indiv. Equities

	1t	1s	5s	1m	5m	15m
RV	10	9	13	49	33	3
RVac1	20	11	19	46	16	1
RK	34	58	54	4	0	0
MSRV	20	18	38	0	0	0
TSRV	0	11	10	0	0	0
MLRV	25	11	22	29	0	0
RRV	30	11	18	45	5	0
BR	17					

Proportion of measures that are in the 90% MCS

Interest rate futures

Int. Rate Futures

	1t	1s	5s	1m	5m	15m
RV	0	0	0	21	8	0
RVac1	0	0	8	8	2	0
RK	0	10	2	0	0	0
MSRV	0	24	21	0	0	0
TSRV	0	22	4	0	0	0
MLRV	0	0	23	0	0	0
RRV	4	10	22	6	0	0
BR	0					

Proportion of measures that are in the 90% MCS

FX futures

Currency Futures

	1t	1s	5s	1m	5m	15m
RV	0	3	25	23	15	0
RVac1	0	15	33	15	8	0
RK	10	21	19	1	0	0
MSRV	0	39	15	5	0	0
TSRV	0	12	0	0	0	0
MLRV	0	36	35	10	0	0
RRV	5	20	18	15	0	0
BR	14					

Proportion of measures that are in the 90% MCS

Equity index futures

Index Futures

	1t	1s	5s	1m	5m	15m
RV	0	0	30	25	0	0
RVac1	0	3	20	10	0	0
RK	0	10	3	0	0	0
MSRV	0	26	8	0	0	0
TSRV	0	0	0	0	0	0
MLRV	0	21	18	0	0	0
RRV	15	8	19	0	0	0
BR	5					

Proportion of measures that are in the 90% MCS

Computed equity indices

Computed Indices

	1t	1s	5s	1m	5m	15m
RV	0	0	0	5	25	0
RVac1	0	0	0	50	5	0
RK	50	10	28	48	0	0
MSRV	20	0	0	20	0	0
TSRV	0	0	0	0	0	0
MLRV	0	0	0	40	0	0
RRV	0	0	0	0	5	0
BR	0					

Summary so far

- When 5min RV is taken as the benchmark realized measure, it is very hard to beat.
- When we treat all measures symmetrically, we find the following are most often in the MCS:
  - 1min RV
  - TSRV and MSRV on 1sec data
  - Realized kernels on 1sec data
- Measures that do particularly poorly include:
  - *Any* measure using 15-min data (except RV and RVac1)
  - TSRV, MSRV, RK, MLRV and RRV on 5-min data
- These results hold also when using a *more accurate RV* (15-min, 5-min RV) or *non-RV proxies* (1-min MSRV and RKth2).

Conditional comparisons of realized measures

- The ranking method described above can also be used to obtain *conditional rankings* of realized measures. For example:

$$L(\widetilde{QV}_t, M_{0t}) - L(\widetilde{QV}_t, M_{jt}) = \beta_0 + \beta_1 Z_{t-1} + e_t$$

  where Z is some conditioning variable

- We consider panel regressions of this form, using lagged volatility and lagged liquidity (using the bid-ask spread) as conditioning variables
- We compare a subset of the better measures so far with RV5min

RV5min vs Other, conditional on level of volatility

RVdaily even worse when vol is high; Same for most on computed indices

- t-statistics on the coefficient on lagged volatility:

"Other" Estimator	Daily RV	RV_1m	RVac1_1m	MSRV_5s	RKth2_5s
All assets	**-5.71**	-1.54	**3.46**	**-3.87**	-1.84
Individual Equities	**-3.08**	**2.68**	0.87	1.03	1.18
Interest Rate Futures	**-2.00**	-1.27	**4.69**	-1.23	-0.61
Currency Futures	-1.52	-0.98	-0.09	-0.93	-0.73
Index Futures	**-3.75**	-0.73	1.26	-1.89	-1.71
Computed Indices	**-4.91**	**-3.48**	-0.09	**-4.62**	**-2.39**

RV5min vs Other, conditional on level of liquidity

RV1min and MSRV do worse when liquidity dries up

- *t*-statistics on the coefficient on lagged liquidity:

"Other" Estimator	Daily RV	RV_1m	RVac1_1m	MSRV_5s	RKth2_5s
All assets	-0.73	**-3.49**	-1.23	**-2.43**	-0.91
Individual Equities	0.34	**-6.07**	-1.53	**-4.10**	-1.53
Interest Rate Futures	**3.22**	0.57	-0.68	0.59	1.21
Currency Futures	-1.41	-0.62	0.32	-0.79	-0.45
Index Futures	**-3.47**	**-2.47**	-0.28	**-2.21**	-1.46

Outline of the presentation

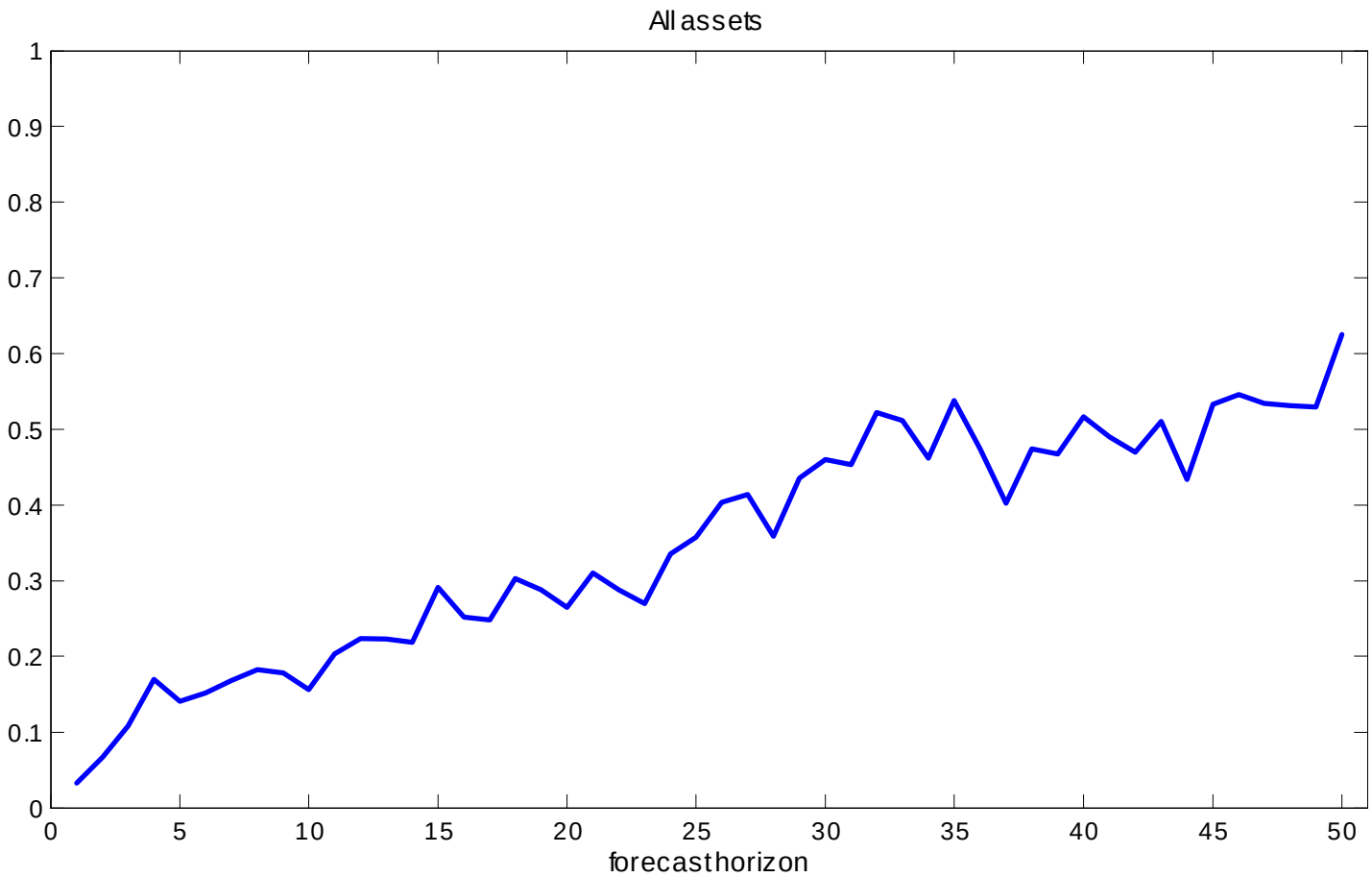
1. The realized measures under analysis (brief)
2. Methods for comparing realized measures
3. Main results:
    1. Guidelines on sampling frequency, sampling scheme, etc
    2. Does anything beat 5-min RV?
    3. The *set* of best realized measures
    4. **Out-of-sample forecast comparisons**
4. Summary and conclusions

Out-of-sample forecasting with realized measures

- Finally, we compare our set of realized measures in an out-of-forecasting experiment.
- We use the HAR model (described earlier), estimated using the most recent 500 days of data, and re-estimate the model for each horizon and each day of the sample.
- We consider forecast horizons from 1-50 days.
- Below we show the size the MCS as the horizon grows, and then we zoom in on the first 5 horizons

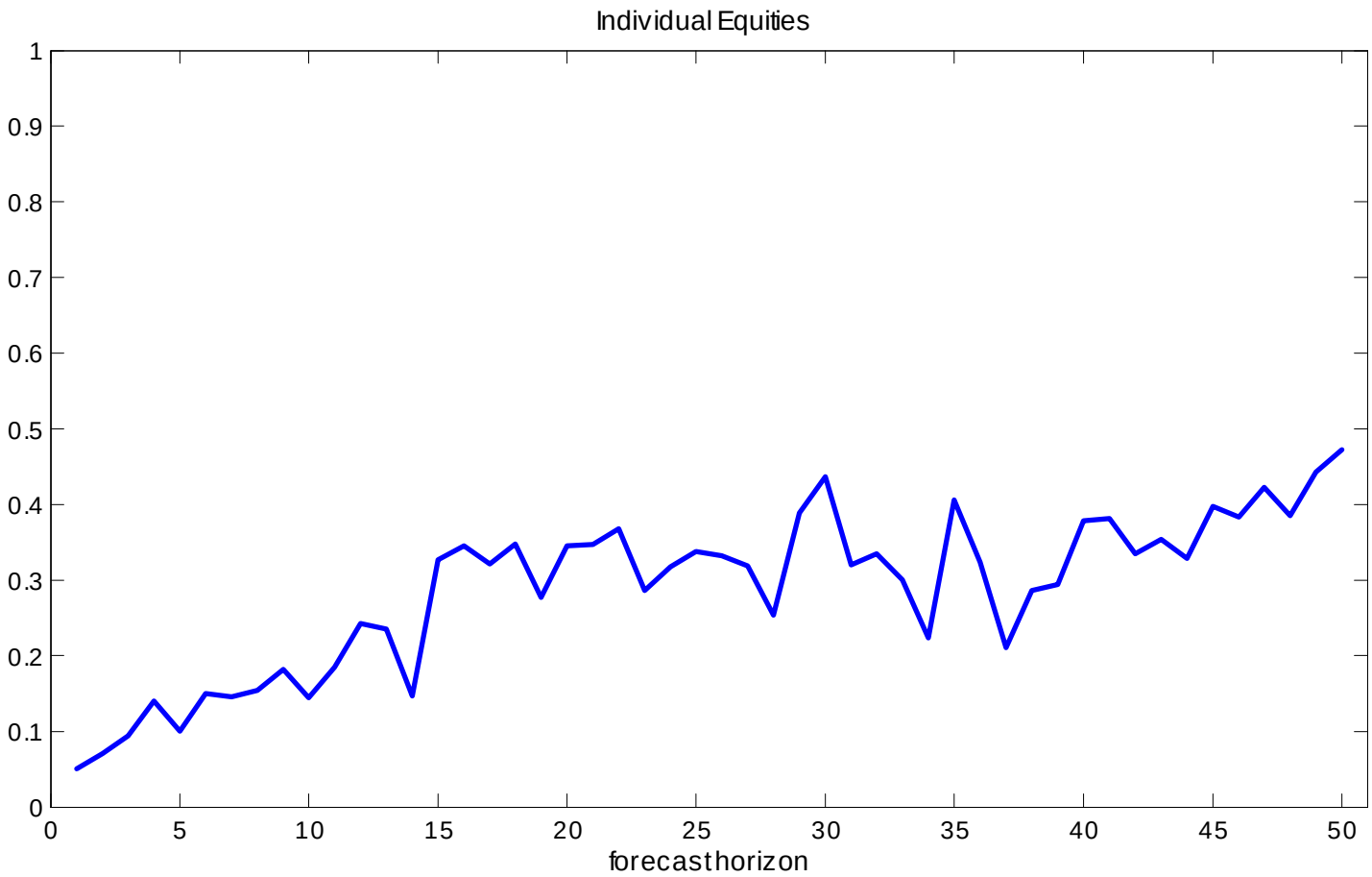
Proportion of measures in the 90% MCS, across horizons

All 31 assets



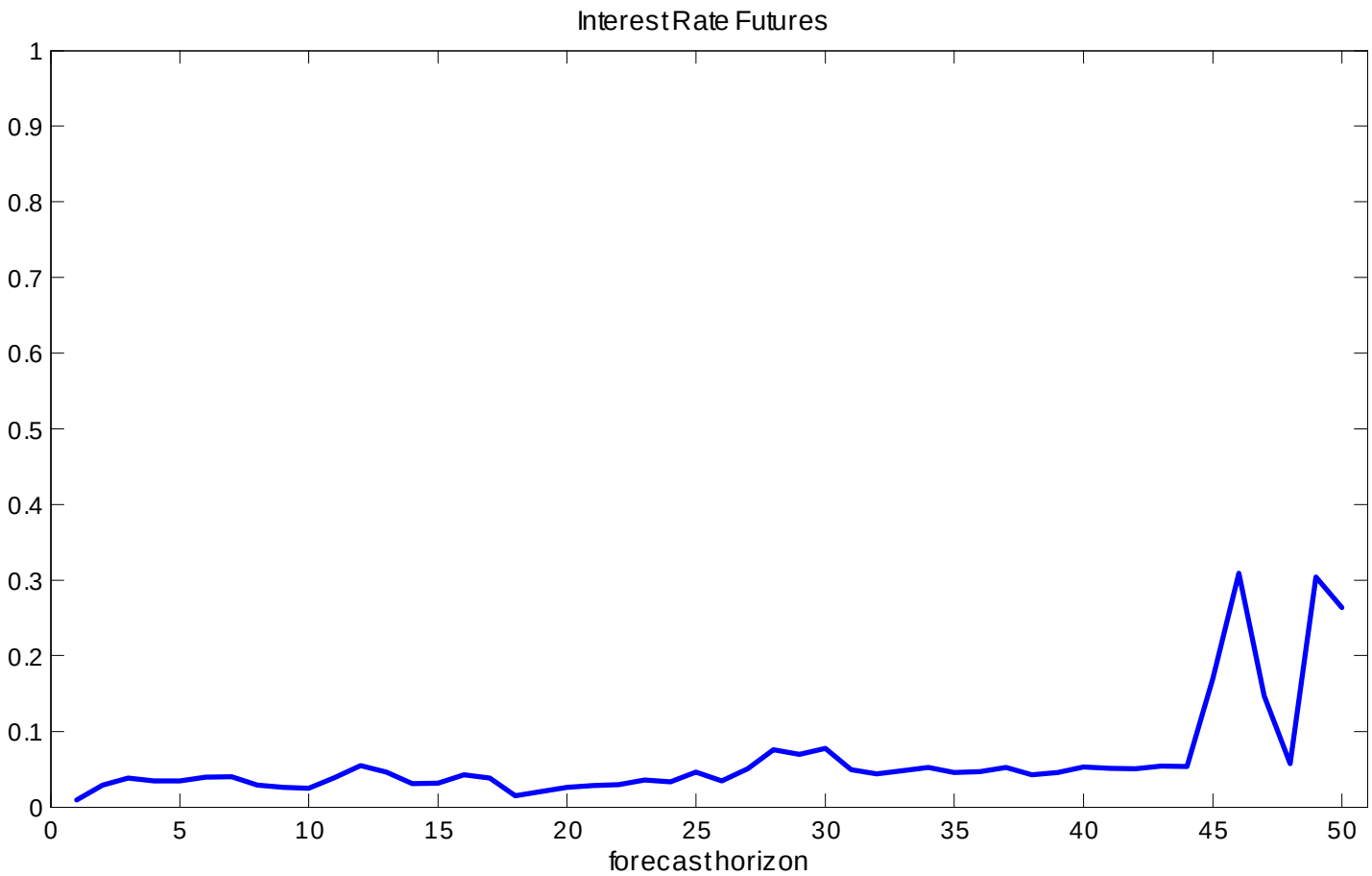
Proportion of measures in the 90% MCS, across horizons

Individual equities



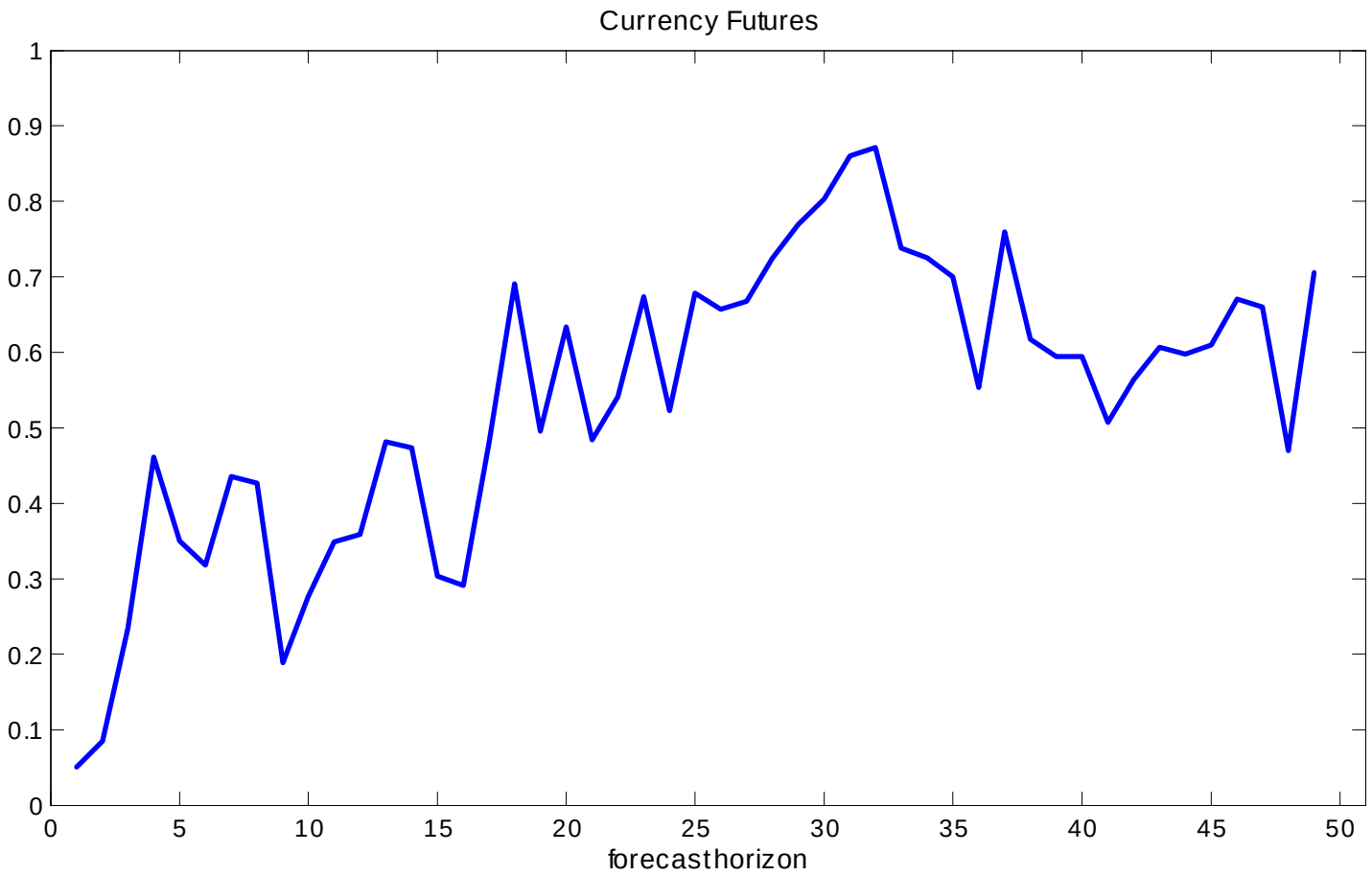
Proportion of measures in the 90% MCS, across horizons

Interest rate futures



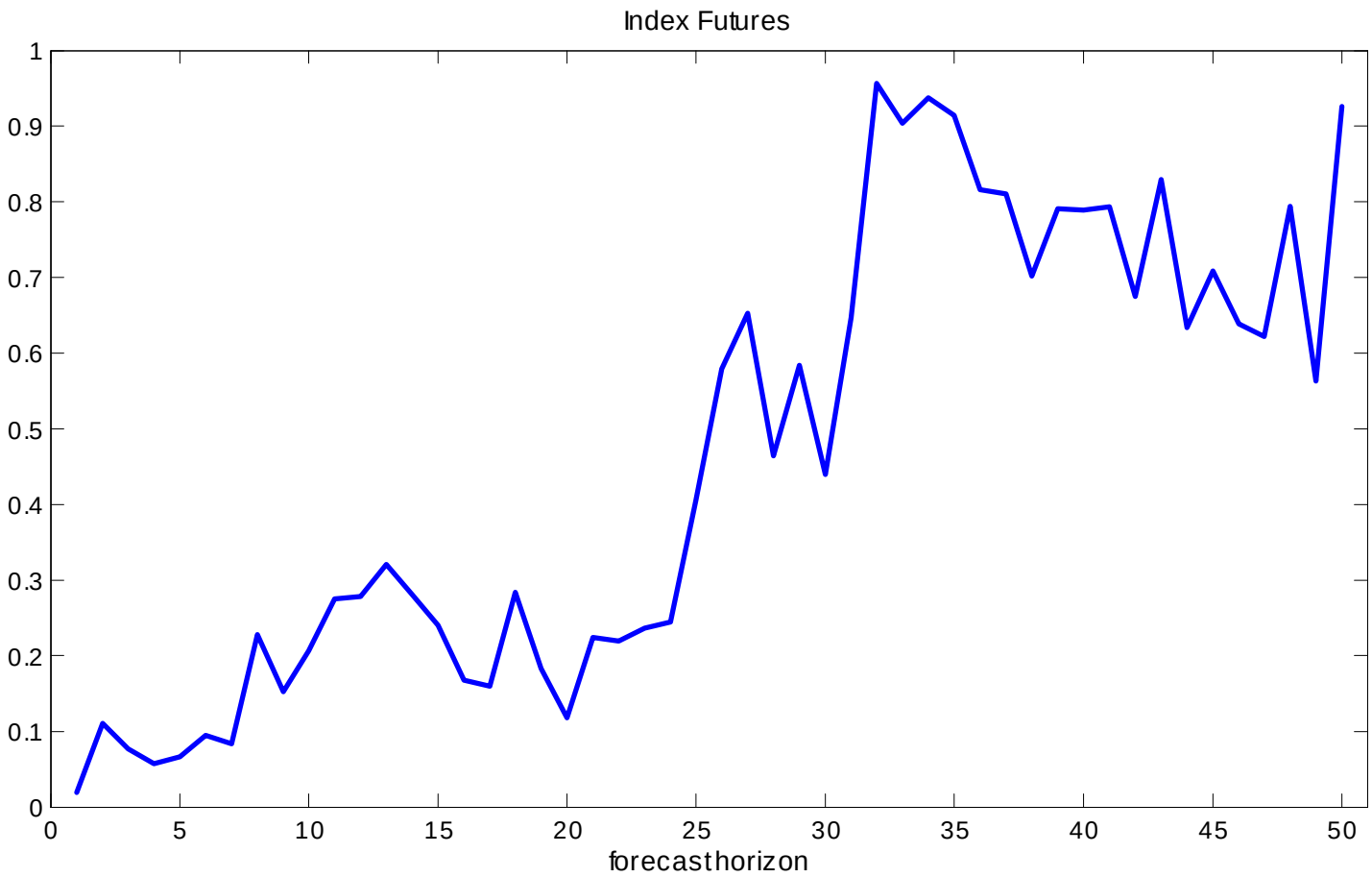
Proportion of measures in the 90% MCS, across horizons

FX futures



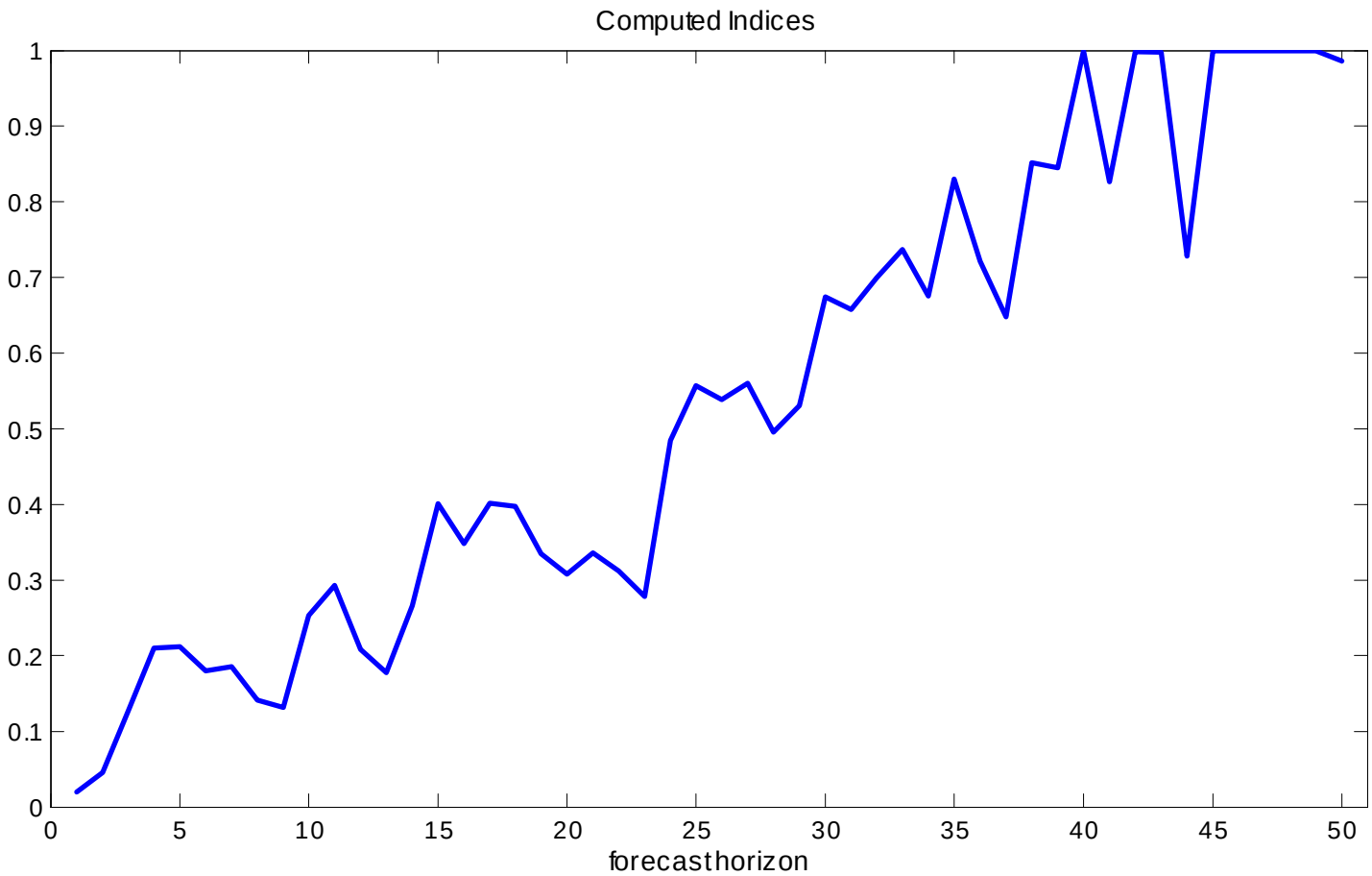
Proportion of measures in the 90% MCS, across horizons

Equity index futures



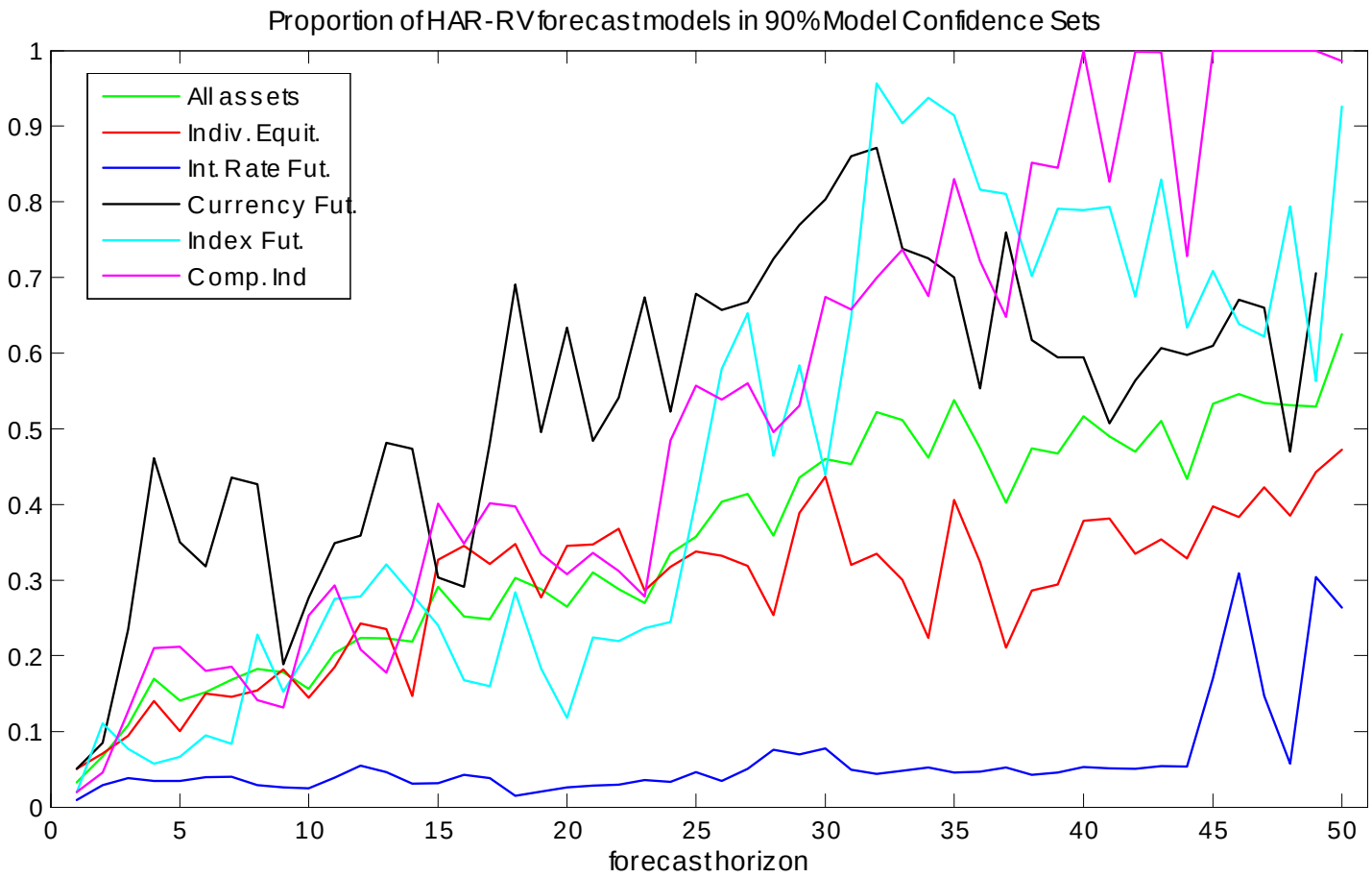
Proportion of measures in the 90% MCS, across horizons

Computed equity indices



Proportion of measures in the 90% MCS, across horizons

All 31 assets



Proportion of measures in the 90% MCS, h=1,2,...,5

All 31 assets

All 31 Assets

	1t	1s	5s	1m	5m	15m
RV	11	14	19	29	38	41
RVac1	15	17	19	31	36	30
RK	36	32	36	40	35	17
MSRV	29	23	25	29	29	12
TSRV	11	14	18	32	36	15
MLRV	21	18	23	32	37	14
RRV	18	16	20	39	43	31
BR	38					
BPV	9	10	15	28	40	49
minRV	9	10	15	26	34	42
medRV	10	9	14	25	40	47
QRV	10	9	17	36	64	54
TrunRV	14	3	17	43	63	66

Proportion of measures in the 90% MCS, h=1,2,...,5

Individual equities

Individual Equities

	1t	1s	5s	1m	5m	15m
RV	10	10	13	31	48	47
RVac1	19	12	18	37	38	38
RK	43	38	45	51	40	19
MSRV	27	22	29	30	32	7
TSRV	13	14	22	38	33	15
MLRV	21	12	20	38	39	12
RRV	18	11	13	44	56	35
BR	39					
BPV	8	6	7	21	33	46
minRV	9	6	7	21	28	38
medRV	8	8	7	21	34	44
QRV	8	0	8	14	51	49
TrunRV	11	4	7	15	37	52

Proportion of measures in the 90% MCS, h=1,2,...,5

Interest rate futures

Interest Rate Futures

	1t	1s	5s	1m	5m	15m
RV	2	5	9	1	7	8
RVac1	2	8	5	2	8	8
RK	6	2	3	9	11	8
MSRV	15	12	5	7	13	11
TSRV	0	7	6	4	10	12
MLRV	5	8	8	3	10	13
RRV	1	6	2	5	15	22
BR	8					
BPV	0	0	0	7	23	37
minRV	0	0	1	10	20	23
medRV	0	0	3	8	23	31
QRV	0	1	9	38	71	63
TrunRV	10	-	10	59	78	81

Proportion of measures in the 90% MCS, h=1,2,...,5

FX futures

Currency Futures

	1t	1s	5s	1m	5m	15m
RV	10	26	38	65	76	80
RVac1	22	35	40	63	79	54
RK	67	66	72	80	76	45
MSRV	28	43	55	72	59	37
TSRV	14	30	34	61	79	38
MLRV	26	42	44	65	79	42
RRV	30	35	56	81	80	67
BR	58					
BPV	16	29	34	73	85	89
minRV	18	25	33	68	76	85
medRV	22	20	33	69	79	88
QRV	20	20	45	84	91	91
TrunRV	20	5	42	90	98	99

Proportion of measures in the 90% MCS, h=1,2,...,5

Equity index futures

Index Futures

	1t	1s	5s	1m	5m	15m
RV	14	12	15	20	23	26
RVac1	6	14	10	21	25	13
RK	31	21	26	23	20	2
MSRV	42	16	9	11	11	0
TSRV	10	2	6	21	25	1
MLRV	24	14	12	23	21	0
RRV	22	14	12	24	18	4
BR	52					
BPV	2	3	25	21	27	30
minRV	6	5	23	18	18	28
medRV	4	5	17	16	27	31
QRV	4	13	14	33	49	27
TrunRV	0	0	11	41	60	46

Proportion of measures in the 90% MCS, h=1,2,...,5

Computed equity indices

Computed Indices

	1t	1s	5s	1m	5m	15m
RV	28	30	33	33	30	47
RVac1	28	30	37	38	28	37
RK	32	31	31	24	22	12
MSRV	44	34	30	32	37	0
TSRV	28	24	20	37	52	8
MLRV	40	30	46	36	41	0
RRV	28	33	40	43	36	22
BR	41					
BPV	40	-	-	26	48	56
minRV	20	-	-	12	42	44
medRV	28	-	0	14	56	48
QRV	32	0	10	26	82	34
TrunRV	32	0	14	37	68	60

Summary and conclusion

- Across 31 assets, 11 years, 350+ realized measures, we find:

1. If 5-min RV is taken as the benchmark measure, it is very hard to beat by *any* measure
2. If no benchmark is specified, the best estimators appear to be:
    - RV on 1-min data, Realized kernels and TSRV on 1-sec data
3. For forecasting, 5-min *truncated* RV appears to provide best results
4. The gains from more sophisticated realized measures are more apparent for more liquid assets (currency & equity index futures), less so for less liquid assets (individ equities & computed indices)
5. For measures based on 5-minute data, tick-time sampling and "sub-sampling" generally lead to improved accuracy.