Supplemental Appendix for

Better the Devil You Know: Improved Forecasts from Imperfect Models

by Dong Hwan Oh and Andrew J. Patton

November 2021

Table S1: Out-of-sample forecast performance for GARCH-X models

	Method details				*Forecast performance*		
Rank	**Model**	**StateVar**	**Bwidth**	**Window**	**AvgLoss**	**GW stat**	**MCS**
1	GARCH-X	time,RV	0.9999,0.4	full	0.293	-9.329	✓
2	GARCH-X	RV	0.41	full	0.294	-9.382	✓
3	GARCH-X	time,FFR	0.98,0.39	full	0.309	-5.017	✓
4	GARCH-X	time,VIX	0.98,0.89	full	0.313	-4.013	✓
5	GARCH-X	time	0.98	full	0.313	-4.653	✓
6	GARCH	time,RV	0.9995,0.34	full	0.320	-4.890	×
7	GARCH-X	-	-	250	0.324	-4.999	×
8	GARCH	RV	0.37	full	0.325	-4.239	×
9*	GARCH-X	time,10Y-2Y	0.9825,0.18	full	0.329	-1.828	×
10	GARCH	time,VIX	0.995,0.28	full	0.333	-2.294	×
11	GARCH-X	-	-	500	0.334	-5.330	×
12	GARCH-X	-	-	1000	0.335	-7.398	×
13	GARCH	VIX	0.32	full	0.349	-1.363	×
14	GARCH-X	VIX	2.63	full	0.351	-8.426	×
15	GARCH-X	10Y-2Y	0.26	full	0.358	-0.191	×
16	GARCH-X	FFR	0.44	full	0.359	-0.018	×
17	GARCH-X	-	-	full	0.359	★	×
18	GARCH	time	0.995	full	0.371	1.428	×
19	GARCH	-	-	500	0.375	2.049	×
20	GARCH	-	-	250	0.376	1.711	×
21	GARCH	time,10Y-2Y	0.9975,0.25	full	0.380	2.191	×
22	GARCH	time,FFR	0.9975,0.49	full	0.381	2.632	×
23	GARCH	-	-	1000	0.382	2.892	×
24	GARCH	FFR	1.81	full	0.400	4.623	×
25	GARCH	10Y-2Y	2.6	full	0.400	4.724	×
26	GARCH	-	-	full	0.402	4.844	×

Notes: This table presents measures of forecast performance over the out-of-sample period (January 2011 to June 2021) from GARCH and GARCH-X models estimated using either QML (non-local), or local QML. All GARCH-X models use VIX^2 as the extra variable. The rows are ordered by average OOS QLIKE loss, reported in the third-last column. The local method with the best performance in the validation sample (the second half of the estimation sample) is marked in the first column with $*$. The local estimators use the state variable(s) given in the third column and bandwidth parameter(s) from the fourth column, which are selected using the validation sample. The fifth column reports the window of data used in estimation, where "full" implies the entire in-sample period (2737 observations). The penultimate column reports Giacomini-White t-statistics of each model relative to the benchmark method (marked with ★), which is taken as the non-local method using the full estimation window, with negative t-statistics indicating lower average loss. The final column includes a check mark if a given method is included in the 95% model confidence set, and a cross otherwise.

Table S2: Out-of-sample forecast performance for HAR-X models

	Method details				*Forecast performance*		
Rank	**Model**	**StateVar**	**Bwidth**	**Window**	**AvgLoss**	**GW stat**	**MCS**
1	HAR-X	RV	0.79	full	0.232	-7.001	✓
2*	HAR-X	time,RV	0.9975,0.8	full	0.232	-6.843	✓
3	HAR-X	VIX	0.63	full	0.236	-6.722	×
4	HAR-X	time,10Y-2Y	0.9875,0.8	full	0.241	-6.085	×
5	HAR-X	time,VIX	0.9925,0.73	full	0.245	-5.723	×
6	HAR	time,VIX	0.999,0.62	full	0.246	-5.256	×
7	HAR-X	-	-	250	0.248	-5.576	×
8	HAR-X	time	0.995	full	0.248	-5.433	×
9	HAR	VIX	1.8	full	0.252	-4.829	×
10	HAR	time	0.995	full	0.252	-4.909	×
=10	HAR	time,RV	0.995,∞	full	0.252	-4.909	×
=10	HAR	time,10Y-2Y	0.995,∞	full	0.252	-4.909	×
=10	HAR	time,FFR	0.995,∞	full	0.252	-4.909	×
14	HAR	10Y-2Y	1.91	full	0.253	-4.789	×
15	HAR	RV	2.86	full	0.253	-4.743	×
16	HAR	-	-	full	0.253	-4.757	×
17	HAR	-	-	500	0.253	-4.785	×
18	HAR	FFR	2.3	full	0.253	-4.743	×
19	HAR	-	-	250	0.255	-4.742	×
20	HAR-X	time,FFR	0.99,0.32	full	0.263	-3.563	×
21	HAR-X	-	-	500	0.273	-3.097	×
22	HAR	-	-	1000	0.300	-0.533	×
23	HAR-X	-	-	1000	0.307	-0.782	×
24	HAR-X	-	-	full	0.325	★	×
25	HAR-X	10Y-2Y	1.96	full	0.351	2.564	×
26	HAR-X	FFR	1.62	full	0.372	3.734	×

Notes: This table presents measures of forecast performance over the out-of-sample period (January 2011 to June 2021) from HAR and HAR-X models estimated using either QML (non-local), or local QML. All HAR-X models use VIX^2 as the extra variable. The rows are ordered by average OOS QLIKE loss, reported in the third-last column. The local method with the best performance in the validation sample (the second half of the estimation sample) is marked in the first column with $*$. The local estimators use the state variable(s) given in the third column and bandwidth parameter(s) from the fourth column, which are selected using the validation sample. The fifth column reports the window of data used in estimation, where "full" implies the entire in-sample period (2737 observations). The penultimate column reports Giacomini-White t-statistics of each model relative to the benchmark method (marked with ★), which is taken as the non-local method using the full estimation window, with negative t-statistics indicating lower average loss. The final column includes a check mark if a given method is included in the 95% model confidence set, and a cross otherwise.

Table S.3: Out-of-sample forecast performance for GARCH-FZ models

	Method details			*Forecast performance*		
Rank	**StateVar**	**Bwidth**	**Window**	**AvgLoss**	**GW stat**	**MCS**
1	RV	1.96	full	-3.862	-4.136	✓
2	-	-	1000	-3.861	-0.619	✓
3	VIX	1.67	full	-3.860	-2.148	✓
4	10Y-2Y	2.72	full	-3.856	-1.249	×
5	-	-	full	-3.855	★	×
=5	FFR	∞	full	-3.855	0.000	×
7	time	0.995	full	-3.846	0.508	×
8	-	-	500	-3.836	0.876	×
9	time,RV	0.99,2.02	full	-3.830	1.071	×
10*	time,VIX	0.9925,1.21	full	-3.829	1.253	×
11	time,FFR	0.9925,2.24	full	-3.828	1.221	×
12	time,10Y-2Y	0.9925,1.04	full	-3.825	1.333	×
13	-	-	250	-3.812	1.308	×

Notes: This table presents measures of forecast performance over the out-of-sample period (January 2011 to June 2021) from GARCH-FZ models estimated using either M estimation or local M estimation and the FZ0 loss function in Equation (31). The rows are ordered by average OOS FZ0 loss, reported in the third-last column. For a given model, the local method with the best performance in the validation sample (the second half of the estimation sample) is marked in the first column with $*$. The local estimators use the state variable(s) given in the second column and bandwidth parameter(s) from the third column, which are selected using the validation sample. The fourth column reports the window of data used in estimation, where "full" implies the entire in-sample period (2737 observations). The penultimate column reports Giacomini-White t-statistics of each model relative to the benchmark method (marked with ★), which is taken as the non-local method using the full estimation window, with negative t-statistics indicating lower average loss. The final column includes a check mark if a given method is included in the 95% model confidence set, and a cross otherwise.