Supplement to "Quantile Forecasting with a Large Number of Potential Predictors: A Hybrid Averaging Approach"[1]

By Jan G. De Gooijer[a] and Dawit Zerom[b]

[a] *University of Amsterdam and* [b] *California State University Fullerton*

In this supplementary document, we first provide some summary statistics of the equity premium and the 14 macroeconomic variables (Section A). Then, in Section B, we present estimates of the slope coefficients on the 14 macroenomic variable using models (12)–(15) in the main text. Section C is similar to Section 5.1 in the main text, but now we report forecasting results for the full sample period using a rolling window scheme.

A Summary statistics

Table 1 reports summary statistics for the equity risk premium and 14 macroeconomic variables. The sample mean monthly equity risk premium is 0.51%. The values of the sample kurtosis larger than three indicate that outliers exist in the equity risk premium series R_t, and in the macroeconomic variables EP, DE, RVOL, NTIS, TBL, LTR, DFY, DFR, and INFL.

B Parameter estimates of the macroeconomic variables

Table 2 shows parameter estimates of the macroeconomic variables for the mean and variance equations for models (12)–(15) in the main text. The variables RVOL, and LTR are statistically significant at the 5% nominal significance level across all mean equations. The variables RVOL, DFY, DFR, and INFL are significant at the 5% nominal level in the variance equations of PM-EGARCHZ (model (13)) and TVM-EGARCHZ (model (15)). These latter results are different from those reported by Cenesizoglu and Timmermann (2012, Table 1) for the period 1956:01–2010:12, who noted that DFR and TBL are significant at the 5% nominal level in the variance equations.

C Empirical results: Rolling forecasting scheme

We generate one-step ahead quantile forecasts of R_{N+1} with a rolling window forecasting scheme. Our first forecast origin/base is 1965:12 and hence the first sample covers the period $t \in$[1951:01, 1964:12] ($n = 168$). The rolling forecasting scheme is based on a fixed window of 180 observations, giving rise to a total of 612 conditional quantile forecasts.

[1] This paper can be viewed as an extension of the work by De Gooijer and Zerom (2019).

C.1 Variable selection

Table 3 shows that at quantile levels $\tau = 0.1$ and 0.9 very few covariates are selected by the hybrid conditional quantile approach using SCAD(L) and SCAD(AD) as shrinkage methods. This in contrast to the variable selection results presented in Table 1 of the main text with a recursive window scheme. Moreover, we see that only in six cases technical indicators are included in the set of selected variables, almost all of them at $\tau = 0.5$. Further, there is hardly any overlap between the sets of variables selected at $\tau = 0.1$ and $\tau = 0.9$. Again, this is different from the results presented in Table 1 of the main text. Finally, we see that the hybrid conditional quantile approach does not identify important quantiles from the set of 28 marginal conditional quantiles associated with the TVM model $\big($model (11)$\big)$.

C.2 Forecasting performance

Table 4 is similar to Table 2 in the main text, but forecasts are now based on the rolling window scheme. In six out of nine cases, taken across Panels A–C, the bold typed entries show that the lowest ratios occur for pairwise combinations between the hybrid conditional quantiles and $\widehat{Q}^{(\mathrm{EW})}_{\tau,1}$ with $\tau = 0.1$ and 0.9. The remaining three cases are for combinations with $\widehat{Q}^{(\mathrm{EW})}_{\tau=0.5,2}$ (column 7). For $\tau = 0.1$ and $\tau = 0.9$, all reported p-values are quite small, irrespective of the adopted shrinkage method and the time period. Once more, these findings strongly indicate that the hybrid averaging approach achieves large improvements in forecast accuracy over parametric EW models in the tails of the conditional quantile distribution.

Table 5 presents p-values of the test statistic $\overline{\mathrm{DM}}_\tau$. We see that for all values of τ and across all time periods, there is no clear difference between the four shrinkage methods, i.e. there are no rejections of the null hypothesis of equal forecasting performance at the 5% nominal level. Recall that finding the tuning parameter λ_n for SCAD(L) and SCAD(AD) is based on minimizing QBIC, whereas CV is used as a selection criterion for ada-LASSO(L) and ada-LASSO(AD). QBIC is known to control the proliferation of overfit without much losing the sensitivity of detecting significant covariates in a much better way than CV. Hence, if sparsity of the covariate effects is preferred over non-sparsity, SCAD(L) and SCAD(AD) are recommended. On the other hand, the ada-LASSO penalty methods provide better out-of-sample forecasting performance.

References

Cenesizoglu, T. and Timmermann, A. (2012). Do return prediction models add economic value? *Journal of Banking & Finance*, 36, 2974–2987. http://doi.org/10.1016/j.jbankfin.2012.06.008.

De Gooijer, J.G. and Zerom, D. (2019). Semiparametric quantile averaging in the presence of high-dimensional predictors. *International Journal of Forecasting*, 35. Available at http://ssrn.com/abstract=3057523.

Table 1: Summary statistics; Time period 1951:01 – 2016:12 (792 observations).

Variable	Mean	Std.dev.	Median	Min	Max	Skewness	Kurtosis
$R_t \times 100$	0.51	4.18	0.85	-24.84	14.87	-0.65	5.42
DP	-3.52	0.41	-3.48	-4.52	-2.60	-0.25	2.41
DY	-3.52	0.41	-3.47	-4.53	-2.61	-0.25	2.43
EP	-2.79	0.42	-2.83	-4.84	-1.90	-0.79	5.99
DE	-0.73	0.30	-0.69	-1.24	1.38	2.57	18.51
RVOL	0.14	0.05	0.14	0.05	0.32	0.84	3.94
BM	0.52	0.25	0.50	0.12	1.21	0.57	2.65
NTIS	0.01	0.02	0.02	-0.06	0.05	-0.97	3.87
TBL	4.33	3.09	4.21	0.01	16.30	0.86	4.10
LTY	6.05	2.75	5.54	1.75	14.82	0.82	3.20
LTR	0.53	2.77	0.30	-11.24	15.23	0.50	6.19
TMS	1.71	1.40	1.64	-3.65	4.55	-0.14	2.86
DFY	0.97	0.44	0.85	0.32	3.38	1.79	7.57
DFR	0.02	1.39	0.06	-9.75	7.37	-0.33	9.81
INFL	0.29	0.35	0.26	-1.92	1.79	0.14	6.06

Table 2: Parameter estimates of the macroeconomic variables for models (12)-(15) in the main text; Time period 1951:01 – 2016:12 (792 observations).

	TVM	TVM-EGARCH	TVM-EGARCHZ	PM-EGARCHZ	TVM-EGARCHZ
Variable	Mean eq.	Mean eq.	Mean eq.	Variance eq.	Variance eq.
DP	0.059	0.045	0.049	0.000	-0.006
DY	0.064	0.045*	0.052	0.000	-0.006
EP	0.039	0.045	0.053	0.001	-0.005
DE	0.027	0.022	0.025	-0.001	-0.005
RVOL	0.077*	0.132*	0.079*	0.184*	0.128*
BM	0.020	0.014	0.014	0.001	-0.001
NTIS	-0.017	-0.700*	-0.041	-0.024*	-0.020
TBL	-0.083*	-0.058	-0.102*	0.035	0.041*
LTY	-0.056	-0.035	-0.085*	0.055	0.049*
LTR	0.090*	0.091*	0.093*	-0.050	-0.056*
TMS	0.072*	0.054	0.055	-0.009	-0.012
DFY	0.018	0.035	0.032	0.096*	0.079*
DFR	0.046	0.009	-0.012	-0.120*	-0.122*
INFL	-0.019	-0.052	-0.082*	0.045*	0.046*

Note: * denotes significant at the 5% nominal level.

Table 3: Selected covariates by the hybrid quantile average approach, with SCAD(L) and SCAD(AD) as shrinkage methods, using jointly one-step ahead conditional quantiles obtained from the four parametric models (12)–(15) in the main text and one-step ahead marginal conditional quantile forecasts from the SP method. Rolling window scheme.

Shrinkage method	Selected model/method	Quantile levels $\tau = 0.1$	$\tau = 0.5$	$\tau = 0.9$
			Panel A: Full prediction sample	
SCAD(L)	(12)	DFR		
	(14)		TMS	
	SP	MA(1,12)	DY, EP, RVOL, TBL, LTR, VOL(1,9)	DFY
SCAD(AD)	(12)	DFR	EP	
	(14)		DY, TMS, MA(1,12)	
	SP		RVOL, TBL, LTR	
			Panel B: Contraction	
SCAD(L)	(12)	NTIS		
	(13)			TBL
	(14)	DY, TBL, TMS, DFY	TBL, TBS	DY, TBL, TMS
	SP	INFL	BM, TBL, LTR, TMS	DP, DY, BM
SCAD(AD)	(12)	NTIS		
	(13)	DY		BM, TBL
	(14)	DY, TBL, TMS, DFY	BM, TBL, TMS	DY, DE, RVOL
	SP		LTR, TMS, VOL(2,12)	
			Panel C: Expansion	
SCAD(L)	(12)	NTIS	EP	
	(13)	DFR	TMS	
	SP	TMS	EP, RVOL, TBL, LTR, VOL(1,9)	DFY
SCAD(AD)	(12)		EP	
	(13)	DFR		
	(14)		EP, TBL, TMS	
	SP		DE, RVOL, TBL, LTR, VOL(1,9)	

Table 4: Comparing four shrinkage-based hybrid (H) conditional quantile averaging methods with the parametric EW quantile forecasts $\widehat{Q}_{\tau,i}^{(\mathrm{EW})}$ for the four parametric models (12)–(15) in the main text and as indexed by the subscript i ($i = 1, \ldots, 4$). Rolling window scheme.

Shrinkage	$\tau = 0.1$				$\tau = 0.5$				$\tau = 0.9$			
method	$\widehat{Q}_{\tau,1}^{(\mathrm{EW})}$	$\widehat{Q}_{\tau,2}^{(\mathrm{EW})}$	$\widehat{Q}_{\tau,3}^{(\mathrm{EW})}$	$\widehat{Q}_{\tau,4}^{(\mathrm{EW})}$	$\widehat{Q}_{\tau,1}^{(\mathrm{EW})}$	$\widehat{Q}_{\tau,2}^{(\mathrm{EW})}$	$\widehat{Q}_{\tau,3}^{(\mathrm{EW})}$	$\widehat{Q}_{\tau,4}^{(\mathrm{EW})}$	$\widehat{Q}_{\tau,1}^{(\mathrm{EW})}$	$\widehat{Q}_{\tau,2}^{(\mathrm{EW})}$	$\widehat{Q}_{\tau,3}^{(\mathrm{EW})}$	$\widehat{Q}_{\tau,4}^{(\mathrm{EW})}$
					Panel A: Full prediction sample							
SCAD(L)	0.823	0.887	0.886	0.883	0.965	0.946	0.954	0.960	0.828	0.897	0.851	0.911
	(**.000**)	(**.000**)	(**.001**)	(**.000**)	(.310)	(**.032**)	(.103)	(.118)	(**.020**)	(.076)	(**.022**)	(.135)
SCAD(AD)	0.799	0.860	0.860	0.856	0.967	0.948	0.956	0.962	0.822	0.889	0.844	0.904
	(**.000**)	(**.000**)	(**.001**)	(**.000**)	(.346)	(.033)	(.112)	(.130)	(**.020**)	(.072)	(**.021**)	(.126)
ada-LASSO(L)	0.761	0.820	0.819	0.815	0.815	**0.800**	0.807	0.811	0.761	0.824	0.782	0.837
	(**.000**)	(**.000**)	(**.000**)	(**.000**)	(.113)	(.056)	(.079)	(.082)	(**.005**)	(**.001**)	(**.001**)	(**.008**)
ada-LASSO(AD)	**0.756**	0.815	0.814	0.811	0.823	0.807	0.814	0.819	**0.741**	0.803	0.762	0.816
	(**.000**)	(**.000**)	(**.000**)	(**.000**)	(.219)	(.113)	(.158)	(.166)	(**.003**)	(**.000**)	(**.000**)	(**.001**)
					Panel B: Contraction							
SCAD(L)	0.742	0.877	0.864	0.860	0.953	0.919	0.934	0.942	0.801	0.863	0.877	0.875
	(.061)	(.206)	(.212)	(.262)	(.415)	(.265)	(.333)	(.361)	(.057)	(**.017**)	(**.001**)	(**.032**)
SCAD(AD)	0.725	0.857	0.844	0.840	0.932	0.899	0.914	0.922	0.786	0.846	0.860	0.858
	(.055)	(.181)	(.185)	(.227)	(.220)	(**.040**)	(.095)	(.099)	(.058)	(**.016**)	(**.002**)	(**.028**)
ada-LASSO(L)	**0.627**	0.740	0.729	0.726	0.702	**0.678**	0.688	0.694	0.723	0.779	0.791	0.790
	(**.001**)	(**.002**)	(**.002**)	(**.002**)	(**.044**)	(**.027**)	(**.035**)	(**.035**)	(**.039**)	(**.000**)	(**.002**)	(**.001**)
ada-LASSO(AD)	0.628	0.742	0.731	0.727	0.763	0.736	0.748	0.754	**0.718**	0.774	0.786	0.785
	(**.002**)	(**.004**)	(**.003**)	(**.002**)	(.262)	(.149)	(.200)	(.212)	(**.037**)	(**.000**)	(**.001**)	(**.000**)
					Panel C: Expansion							
SCAD(L)	0.858	0.899	0.902	0.898	0.973	0.959	0.965	0.970	0.843	0.914	0.856	0.930
	(.129)	(.168)	(.191)	(.189)	(.521)	(.275)	(.382)	(.382)	(.155)	(.397)	(.108)	(.427)
SCAD(AD)	0.834	0.874	0.876	0.873	0.983	0.968	0.975	0.979	0.836	0.906	0.849	0.921
	(.273)	(.309)	(.333)	(.326)	(.519)	(.441)	(.477)	(.479)	(.101)	(.377)	(.076)	(.415)
ada-LASSO(L)	0.804	0.842	0.845	0.841	0.849	0.837	0.842	0.846	0.772	0.837	0.784	0.851
	(**.005**)	(**.001**)	(**.011**)	(**.001**)	(.423)	(.336)	(.375)	(.377)	(**.003**)	(**.048**)	(**.002**)	(.090)
ada-LASSO(AD)	**0.798**	0.836	0.839	0.835	0.841	**0.829**	0.834	0.838	**0.748**	0.811	0.759	0.824
	(**.006**)	(**.003**)	(**.004**)	(**.002**)	(.314)	(.237)	(.271)	(.272)	(**.000**)	(**.010**)	(**.000**)	(**.025**)

Notes: i) The entries are ratios $\rho_\tau(e_{i,t}^{(\mathrm{H})})/\rho_\tau(e_{i,t}^{(\mathrm{EW})})$ averaged over 50 replications. Embolded entries show the lowest ratios for each τ value and each panel; ii) The values in parentheses are p-values of the test statistic $\overline{\mathrm{D}}_{\tau,h=1}(i,i')$ $(i,i' = 1,\ldots,4)$. Embolded p-values indicate rejection of the null hypothesis at the 5% nominal significance level.

Table 5: Paired comparison of the hybrid quantile forecasts using four shrinkage methods.

τ	$\mathbb{H}_0: \mathrm{M}_1$ $\mathbb{H}_1: \mathrm{M}_2$	$\mathbb{H}_0: \mathrm{M}_1$ $\mathbb{H}_1: \mathrm{M}_3$	$\mathbb{H}_0: \mathrm{M}_1$ $\mathbb{H}_1: \mathrm{M}_4$	$\mathbb{H}_0: \mathrm{M}_2$ $\mathbb{H}_1: \mathrm{M}_3$	$\mathbb{H}_0: \mathrm{M}_2$ $\mathbb{H}_1: \mathrm{M}_4$	$\mathbb{H}_0: \mathrm{M}_3$ $\mathbb{H}_1: \mathrm{M}_4$
			Panel A: Full prediction sample			
0.1	0.933	0.901	0.952	0.926	0.909	0.817
0.5	0.366	0.862	0.745	0.869	0.756	0.244
0.9	0.517	0.957	0.998	0.949	0.995	0.941
			Panel B: Recession			
0.1	0.702	0.596	0.688	0.543	0.629	0.813
0.5	0.498	0.847	0.595	0.940	0.678	0.103
0.9	0.910	0.710	0.755	0.655	0.700	0.685
			Panel C: Expansion			
0.1	0.518	0.622	0.784	0.546	0.607	0.738
0.5	0.489	0.578	0.679	0.579	0.669	0.725
0.9	0.493	0.763	0.882	0.833	0.932	0.931

Note: The entries are p-values of the test statistic $\overline{\mathrm{DM}}_{\tau,h=1}$ for six pairwise combinations of four shrinkage methods SCAD(L), SCAD(AD), ada-LASSO(L), and ada-LASSO(AD) denoted by the short hand notation $\mathrm{M}_1, \ldots, \mathrm{M}_4$, respectively. Rolling window scheme.