Supplement to “On LASSO for Predictive Regression”

Ji Hyung Lee Zhentao Shi Zhan Gao

(for online publication only)

S1 Parameters in Simulations

We use the dataset of Welch and Goyal (2008) to empirically determine the covariance matrices, Σ_ε, and the AR(1) coefficient Φ, used in the data generation process (DGP) in our simulation exercises. According to Figure 1, the long-term return of government bonds (`ltr`), inflation (`infl`) and stock variance (`svar`) are determined as I(0) regressors, denoted as z_{i1}, z_{i2} and z_{i3} in this section. All other 9 regressors, the earning price ratio (`ep`), book-to-market ratio (`bm`), treasury bill rates (`tbl`), default yield spread (`dfy`), net equity expansion (`ntis`), dividend yield (`dy`), dividend price ratio (`dp`), default return spread (`dfr`) and term spread (`tms`), are viewed as I(1) processes, denoted as $x_{i1}, \ldots, x_{i9}$.

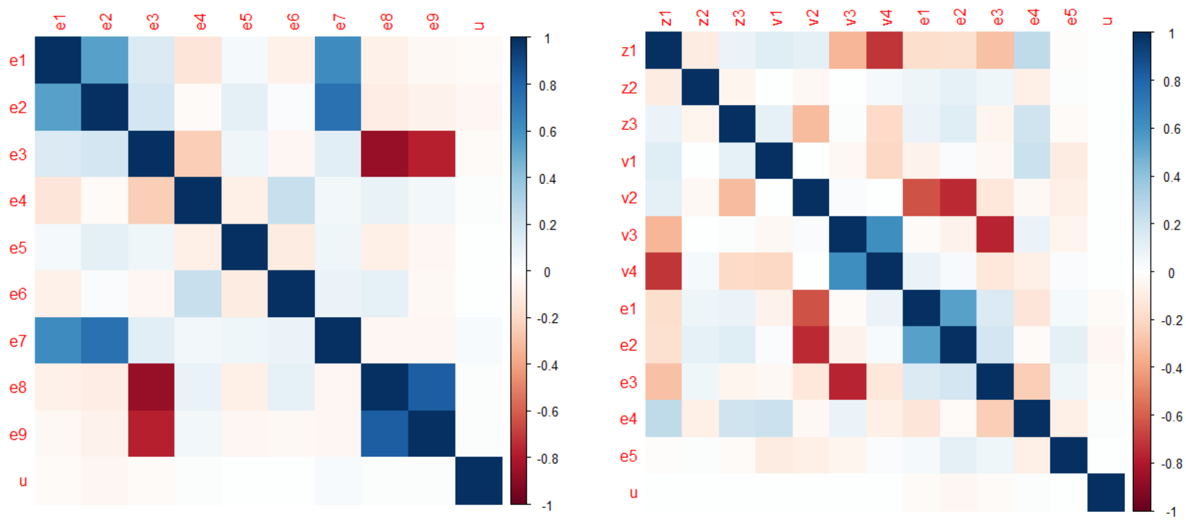
We impose cointegration relations among (`dp`, `dy`) and (`dfr`, `tms`) to obtain estimated cointegration residuals v_{i1}, v_{i2}, v_{i3}, v_{i4}. The error term in the predictive regression equation u_i is obtained as the OLS residuals of regressing the one-month equity premium on all 12 regressors. Then we can estimate $\varepsilon_{i\cdot}^{(z)}$, $\varepsilon_{i\cdot}^{(v)}$ and $\varepsilon_{i\cdot}^{(e)}$ as the AR(1) regression residuals of $z_{i\cdot}$, $v_{i\cdot}$, $x_{i\cdot}$, respectively, and $\varepsilon_i^{(u)} = u_i$.

In DGP 1, the covariance matrix Σ is calculated using $\varepsilon_{il}^{(e)}$, $l = 1, 2, \cdots, 9$, and $\varepsilon_i^{(u)}$ with the variance normalized to be 1. In DGPs 2–4, the covariance matrix Σ_ε, is then calculated using $\varepsilon_{il}^{(z)}$, $l = 1, 2, 3$, $\varepsilon_{il}^{(v)}$, $l = 1, 2, 3, 4$ and $\varepsilon_{il}^{(e)}$, $l = 1, 2, \cdots 5$, with Σ_{zu} and Σ_{vu} imposed to be 0 and variance normalized to be 1. The values of the correlation matrix are represented by the heat map in Figure S1. Denote ϕ as the vector containing estimated AR(1) coefficients of z_{i1}, z_{i2}, z_{i3}, v_{i1}, v_{i2}, v_{i3} and v_{i4}. In DGP 1 we set Φ is a zero matrix, while in DGPs 2–4 we specify $\Phi = \mathrm{diag}\left((\phi', \mathbf{0}_5')'\right)$.

S2 Number of Selected Predictors in Empirical Application

We report the average number of selected predictors from the dataset of Welch and Goyal (2008). The tuning parameters are determined by cross validation. As we can see in the left panel of Table S1, across different rolling window widths and forecast horizons the conventional LASSO methods select at least one predictor with at least 80% of chance, and Alasso/TAlasso do so with at least 38%. The right panel of Table S1 shows the average number of selected predictors.

Figure S1: Correlation matrix heat map



(S1.a) DGP 1

(S1.b) DGPs 2–4

S3 Additional Simulation

We modify DGP 2 by letting $\beta_n^* = (1/\sqrt{n}, 0.1, 0, 0, 0)$, i.e. we fix one coefficient associated with an I(1) predictor as a constant. The corresponding simulation results are summarized in Table S2.

S4 Additional Figures

We take a closer look at Figure 2 in the paper by dividing the whole time span into 10-year sub-periods, as shown in Figures S2–S6. There are good times and bad times for all methods. For example, in the period around 1992 in the $h = 1$ and 10-year rolling window case, Alasso/TAlasso manages to track trend while conventional LASSO fails. In terms of RMPSE and MPAE, Table 4 shows the adaptive algorithms perform better in terms of forecasting accuracy on average.

References

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies 21*(4), 1455–1508.

Table S1: Number of selected predictors in the empirical application

		Fraction of the number of selected predictors larger than 1				Average number of selected predictors			
	h	Alas.	TAlas.	Plas.	Slas.	Alas.	TAlas.	Plas.	Slas.
10-year rolling window	1/12	0.7468	0.7338	0.7295	0.8849	2.7237	2.6187	3.9698	4.3856
	1/4	0.6883	0.6811	0.8384	0.8788	2.2179	2.1659	5.1414	5.1962
	1/2	0.7609	0.7464	0.9232	0.9739	2.2710	2.1797	6.7377	7.0783
	1	0.7734	0.7705	0.9605	0.9868	2.2661	2.2018	6.9327	7.9576
	2	0.3839	0.3780	0.8155	0.9554	1.3750	1.3571	5.4821	6.8021
	3	0.3864	0.3848	0.9788	0.9758	1.4515	1.4318	6.7970	7.8152
15-year rolling window	1/12	0.8268	0.8063	0.8787	0.9559	2.9228	2.7591	4.6709	5.7748
	1/4	0.9558	0.9321	0.9179	0.9400	2.8957	2.7725	5.0742	5.5450
	1/2	0.9333	0.9190	0.9587	0.9508	2.7270	2.6063	6.0429	7.2143
	1	0.8718	0.8574	0.9760	1.0000	2.5849	2.4038	6.7468	7.6987
	2	0.3905	0.3905	0.8758	0.9101	1.4101	1.4069	5.3284	6.5343
	3	0.5000	0.5000	0.9883	0.9917	1.6567	1.6533	7.3417	8.5967

Table S2: MPSE and variable screening in simulations

(a) MPSE (Relative to OLS)

n	Oracle	OLS	Alas.	TAlas.	Plas.	Slas.
80	0.8651	1.0000	0.9437	**0.9407**	0.9483	0.9596
120	0.9034	1.0000	0.9539	**0.9505**	0.9674	0.9763
200	0.9445	1.0000	0.9746	**0.9719**	0.9888	0.9937
400	0.9748	1.0000	0.9844	**0.9820**	0.9958	1.0032
800	0.9875	1.0000	0.9909	**0.9906**	0.9982	1.0066

(b) Variable screening success rates

	SR				SR_1				SR_2			
n	Alas.	TAlas.	Plas.	Slas.	Alas.	TAlas.	Plas.	Slas.	Alas.	TAlas.	Plas.	Slas.
80	0.738	**0.772**	0.622	0.521	0.869	0.855	0.939	**0.964**	0.645	**0.713**	0.396	0.204
120	0.780	**0.817**	0.611	0.529	0.922	0.913	0.974	**0.983**	0.678	**0.748**	0.352	0.205
200	0.825	**0.865**	0.595	0.536	0.965	0.961	0.993	**0.994**	0.725	**0.796**	0.310	0.208
400	0.873	**0.915**	0.573	0.545	0.990	0.990	**0.999**	**0.999**	0.790	**0.862**	0.268	0.221
800	0.903	**0.946**	0.555	0.552	0.998	0.998	**1.000**	**1.000**	0.836	**0.910**	0.238	0.233

Note: Panel (a) compares the out-of-sample prediction accuracy in terms of the mean prediction squared error (MPSE) $E\big[(y_n - \widehat{y}_n)^2\big]$ where that of OLS is normalized to be 1. Panel (b) compares the variable screening performance in terms of success rates SR, SR_1 and SR_2 defined in (28) and (29). The bold number is for the best performance in each category of measurement.

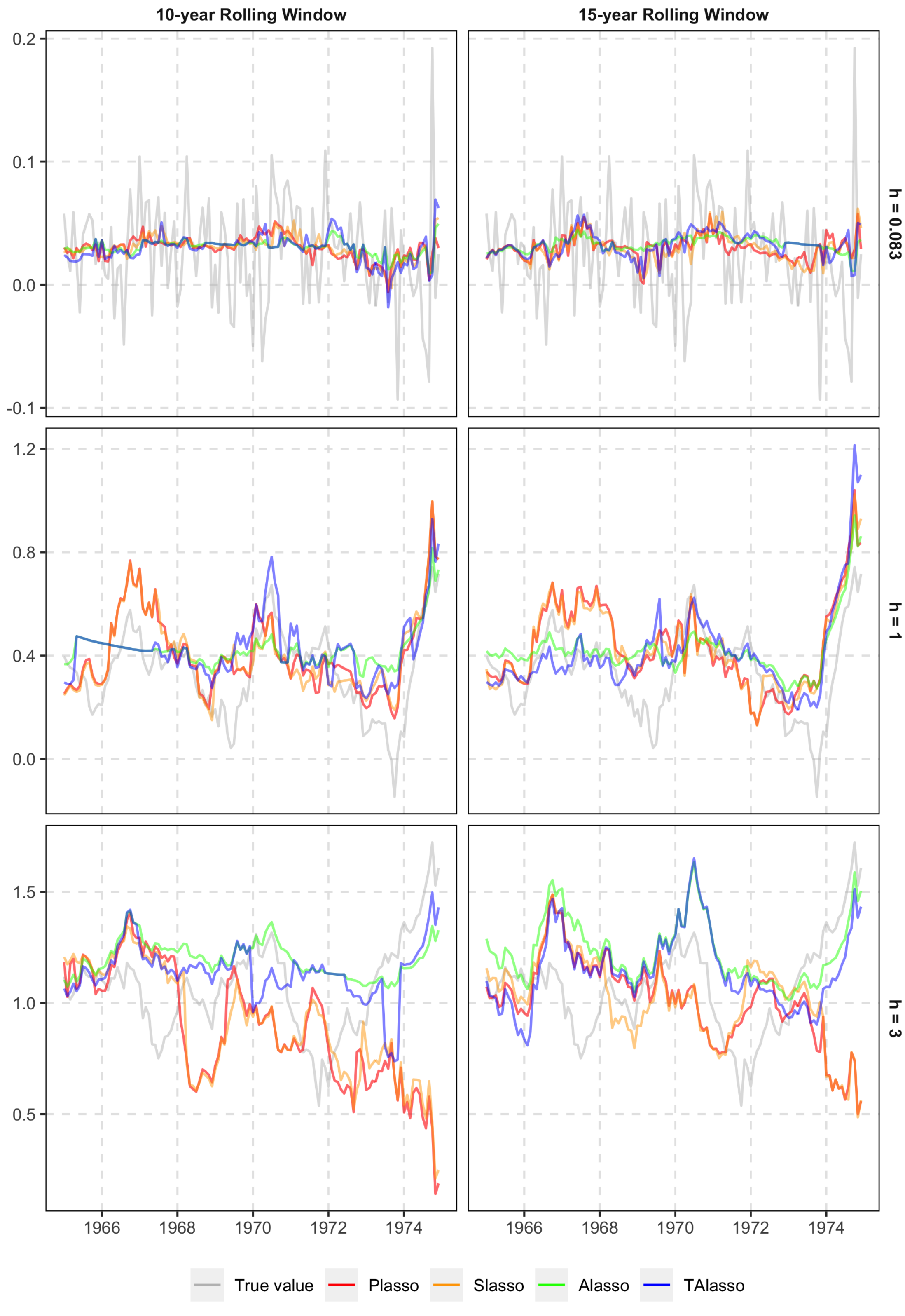


Figure S2: Realized return versus predicted returns: 1965–75

Figure S3: Realized return versus predicted returns: 1975–85

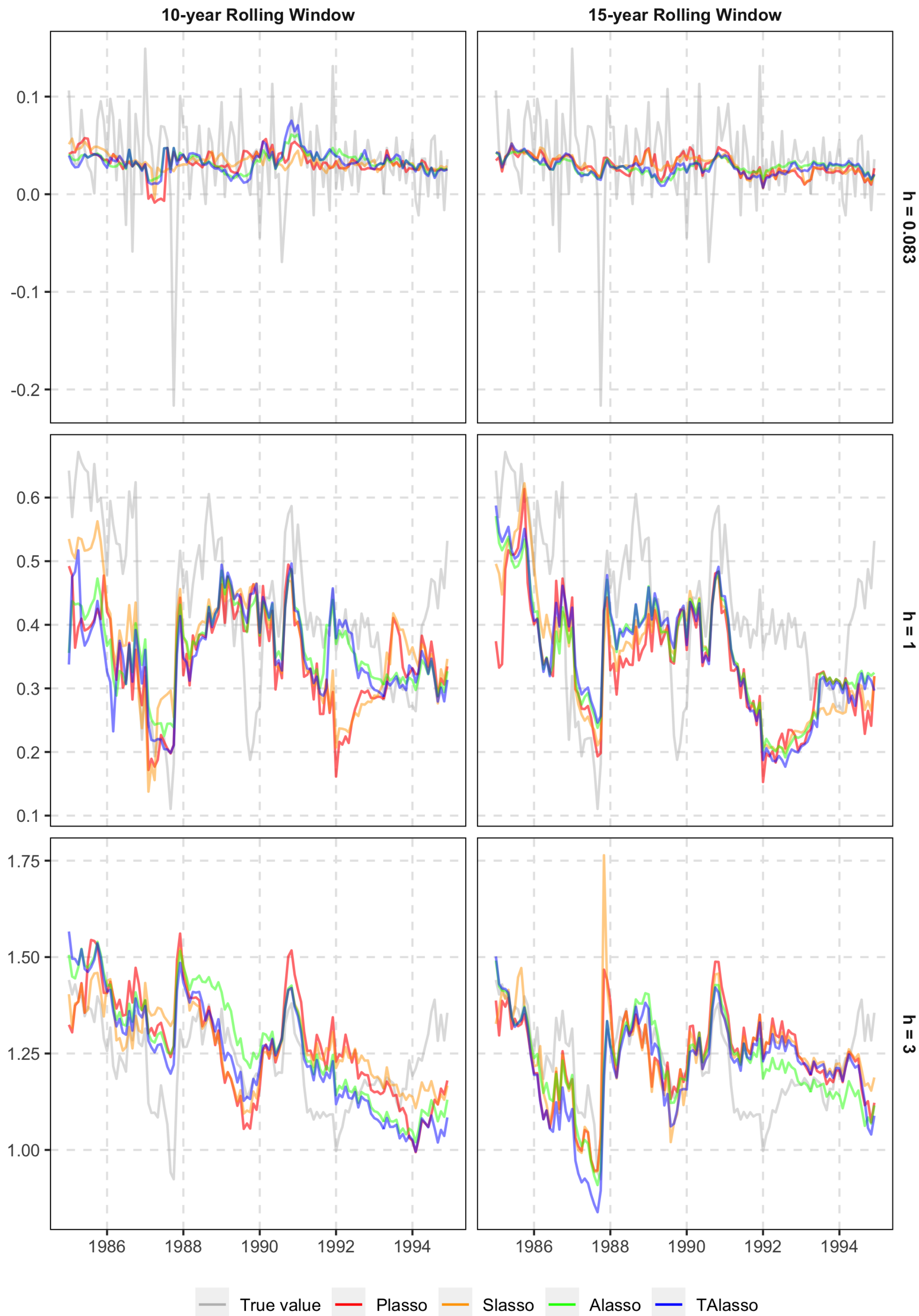


Figure S4: Realized return versus predicted returns: 1985–95

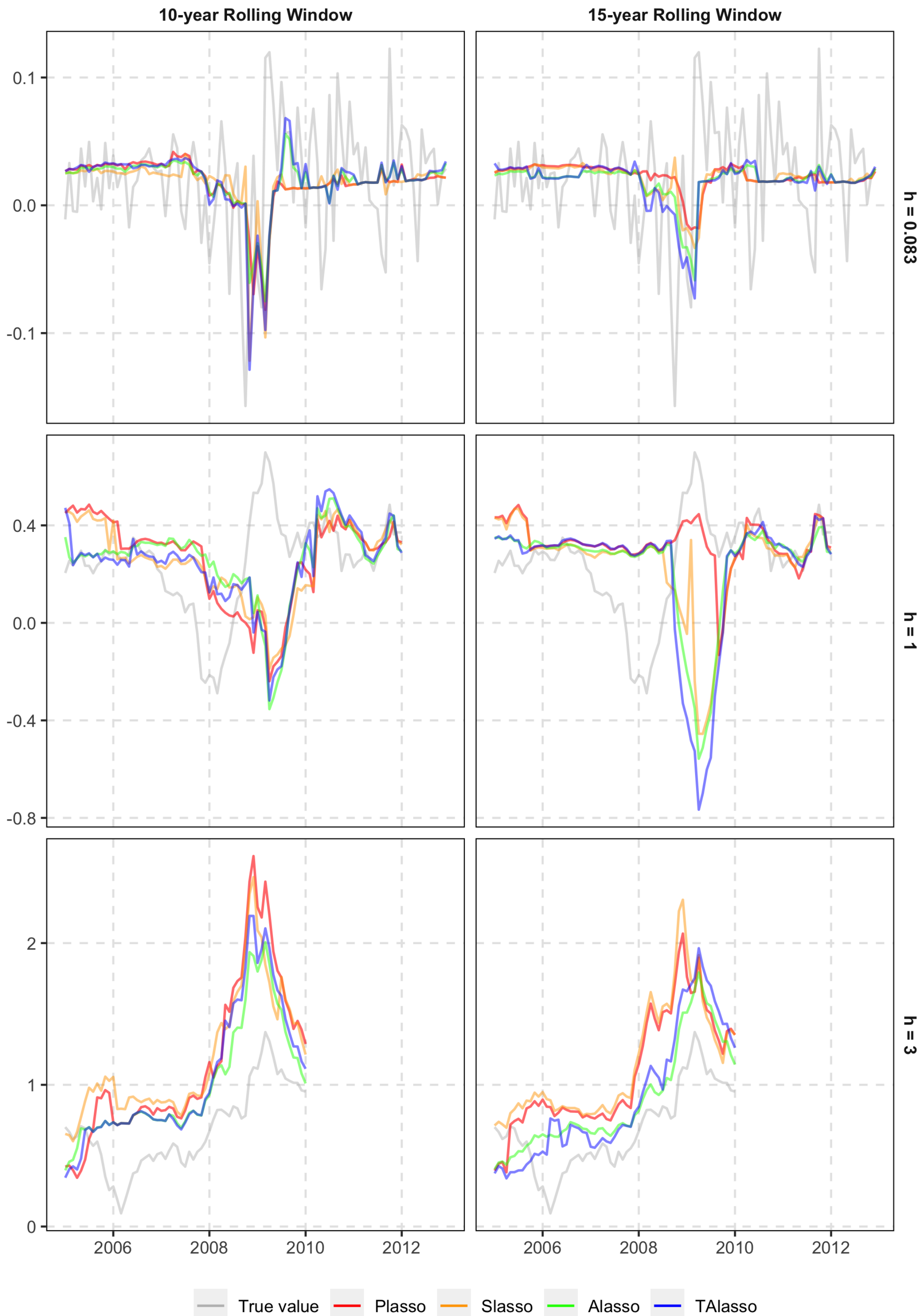


Figure S5: Realized return versus predicted returns: 1995–05

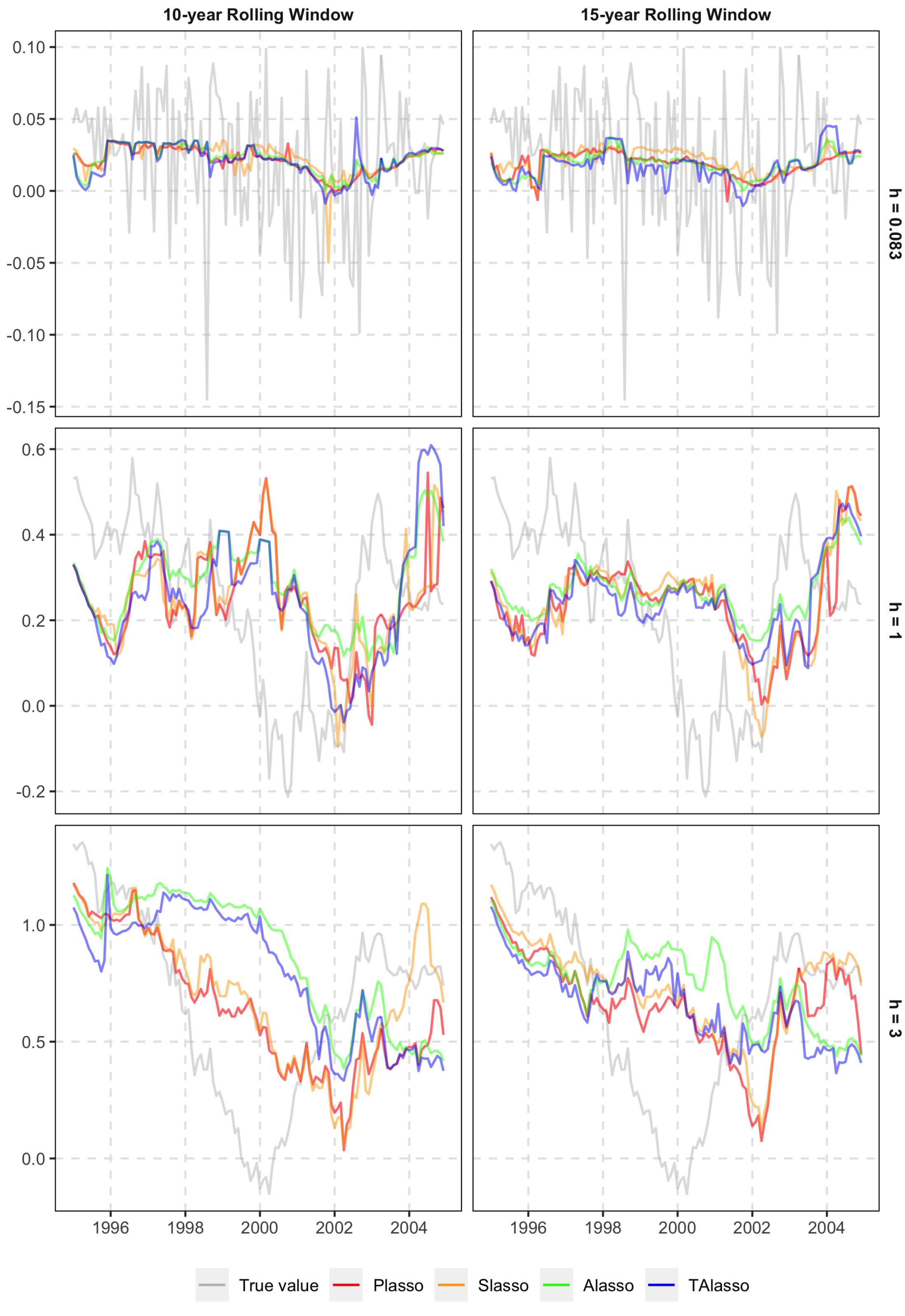


Figure S6: Realized return versus predicted returns: 2005–12