Electronic Appendix to "Nonlinear Forecasting With Many Predictors Using Kernel Ridge Regression"

Peter Exterkate [a, b, *] Patrick J.F. Groenen [c, d] Christiaan Heij [c] Dick van Dijk [c, d, e]

[a] *School of Economics, University of Sydney, Australia*
[b] *CREATES, Aarhus University, Denmark*
[c] *Econometric Institute, Erasmus University Rotterdam, The Netherlands*
[d] *Erasmus Research Institute of Management, Erasmus University Rotterdam, The Netherlands*
[e] *Tinbergen Institute, Erasmus University Rotterdam, The Netherlands*

November 10, 2015

Abstract

This Electronic Appendix includes extended versions of all tables in our article titled "Nonlinear Forecasting With Many Predictors Using Kernel Ridge Regression". The extensions entail the addition of several simulation DGPs, two further benchmark methods, as well as different forms of cross-validation for tuning parameter selection.

In the simulation experiments documented in the article, only DGPs with the factors explaining equal fractions of the variance of the predictors (R_x^2) and of the dependent variable (R_y^2) are considered. The analogous DGPs with $R_x^2 \neq R_y^2$ are included in the tables in this Appendix, and we observe that varying R_y^2 has a much larger effect than varying R_x^2.

The benchmarks considered in the article include three forecast combination schemes using all possible one-regressor models, labeled "Comb EW" for equal weights, "Comb iMSE" for inverse mean squared error weights, and "Comb JMA" for Jackknife Model Averaging weights. In this Appendix we also explore forecast combination using all possible two-regressor models. Using the EW and iMSE weighting schemes, the results turn out to be only marginally different from the averages over one-regressor models. Performing JMA over all two-regressor models appears to be computationally infeasible.

The tuning parameters in kernel ridge regression are selected using five-fold cross-validation in the article. As a robustness check, we also report results based on ten-fold as well as leave-one-out cross-validation in this Appendix; observe that leave-one-out corresponds to 120-fold cross-validation in our applications. We find that the results are not very sensitive to the choice of the number of folds.

*Corresponding author. School of Economics, Merewether Building H04, The University of Sydney, NSW 2006, Australia; email: peter.exterkate@sydney.edu.au; phone: +61 2 9351 8532.

Table E.1: Relative mean squared prediction errors for the DGPs (6)-(9), with observable factors.

DGP	Linear		Squared		Cross-product		Threshold	
$R^2 =$	0.4	0.8	0.4	0.8	0.4	0.8	0.4	0.8
Benchmark methods								
Mean	1.03	1.02	1.02	1.05	1.05	1.08	1.03	1.03
Comb EW	0.72	*0.42*	0.99	*0.99*	1.02	*1.02*	0.82	0.60
Comb iMSE	*0.72*	0.42	*0.99*	0.99	*1.02*	1.02	0.81	0.59
Comb JMA	0.73	0.42	0.99	1.00	1.02	1.02	*0.81*	*0.59*
Principal-components-based methods								
PC	***0.62***	*0.20*	1.00	0.99	1.02	1.02	0.74	0.45
PC^2	0.62	0.21	***0.63***	0.28	***0.63***	*0.26*	*0.74*	*0.44*
SPC	0.83	0.63	0.84	0.68	1.06	1.08	0.90	0.76
PC-Sieve	0.62	0.21	0.65	*0.28*	0.64	0.26	0.75	0.45
PC-NW	0.79	0.55	0.92	0.81	0.92	0.77	0.85	0.68
Kernel ridge regression, five-fold cross-validation								
Poly(1)	*0.62*	*0.20*	1.00	0.99	1.02	1.02	0.74	0.45
Poly(2)	0.62	0.21	*0.65*	***0.27***	*0.65*	***0.26***	0.74	0.44
Gauss	0.63	0.21	0.72	0.33	0.78	0.42	*0.73*	***0.34***
Kernel ridge regression, ten-fold cross-validation								
Poly(1)	*0.62*	*0.20*	0.99	0.99	1.02	1.03	0.74	0.45
Poly(2)	0.62	0.21	*0.65*	*0.27*	*0.65*	*0.26*	0.74	0.44
Gauss	0.63	0.21	0.72	0.33	0.78	0.42	*0.73*	*0.35*
Kernel ridge regression, leave-one-out cross-validation								
Poly(1)	*0.62*	***0.20***	0.99	0.99	1.02	1.02	0.74	0.45
Poly(2)	0.62	0.21	*0.65*	*0.27*	*0.65*	*0.26*	0.74	0.44
Gauss	0.63	0.21	0.72	0.33	0.78	0.42	***0.73***	*0.35*
Forecast combinations								
Linear	0.72	0.41	0.99	0.99	1.02	1.02	0.81	0.59
No KRR	0.69	0.35	0.81	0.63	0.85	0.70	0.79	0.53
All 5-f	*0.66*	*0.29*	*0.78*	*0.56*	0.83	*0.62*	0.76	0.47
All 10-f	0.66	0.29	0.78	0.56	0.83	0.62	0.76	0.47
All LOO	0.66	0.29	0.78	0.56	*0.83*	0.62	*0.76*	*0.47*
Diebold-Mariano tests								
Nonlin.	**19.04**	**31.99**	**23.09**	**28.65**	**22.60**	**26.84**	**16.65**	**29.45**
Kernel 5-f	**18.65**	**34.47**	**19.56**	**24.39**	**20.92**	**24.53**	**19.34**	**38.96**
Kernel 10-f	**18.65**	**34.50**	**19.44**	**24.23**	**20.99**	**24.86**	**19.42**	**39.16**
Kernel LOO	**18.70**	**34.47**	**19.40**	**24.23**	**20.87**	**24.80**	**19.60**	**39.10**

Notes: This table extends Table 1 in the article. It reports mean squared prediction errors (MSPEs) for models (6)-(9), averaged over 5000 forecasts, and relative to the variance of the series being predicted. It is assumed that $x_t = f_t$; that is, the factors are observed. These DGPs have no dynamic structure, so that x_{T+1} is used to forecast y_{T+1}. Three sets of kernel ridge regression forecasts are shown, with tuning parameters selected using five-fold, ten-fold, and leave-one-out cross-validation, respectively. The combination forecasts are averages of the Mean, Comb, and PC forecasts (“Linear”), all benchmark and PC-based methods forecasts (“No KRR”), all benchmark, PC-based, and five-fold kernel-based forecasts (“All 5-f”), and similarly for “All 10-f” and “All LOO”. The smallest relative MSPE for each DGP (column) within each group of methods (benchmarks, PC-based, five-fold kernel-based, ten-fold kernel-based, leave-one-out kernel-based, or combinations) is printed in italics, with the overall smallest in boldface italics. The last four rows report the t statistics of Diebold-Mariano tests for equal predictive ability. “Nonlin.” compares “Linear” to “No KRR”; a positive statistic indicates better performance of the latter. Similarly, “Kernel 5-f” compares “No KRR” to “All 5-f”, “Kernel 10-f” compares “No KRR” to “All 10-f”, and “Kernel LOO” compares “No KRR” to “All LOO”. The statistic is printed in boldface if it is significant at the 5% level.

Table E.2: Relative mean squared prediction errors for the DGPs (6)-(9), with i.i.d. latent factors.

DGP	Linear				Squared				Cross-product				Threshold			
$R_y^2 =$	0.4		0.8		0.4		0.8		0.4		0.8		0.4		0.8	
$R_x^2 =$	0.4	0.8	0.4	0.8	0.4	0.8	0.4	0.8	0.4	0.8	0.4	0.8	0.4	0.8	0.4	0.8
Benchmark methods																
Mean	1.02	1.02	1.03	1.03	1.00	1.00	1.02	1.02	1.02	1.02	1.04	1.04	1.00	1.00	1.00	1.00
C1 EW	0.87	0.76	0.74	0.50	0.99	0.98	0.99	0.97	1.01	1.00	1.01	0.99	0.90	0.82	0.80	0.64
C1 iMSE	0.87	0.74	0.71	0.41	0.99	0.98	0.99	0.97	1.01	1.00	1.01	0.98	0.90	0.81	0.79	0.60
C1 JMA	*0.77*	*0.65*	*0.54*	*0.26*	0.99	0.98	0.97	0.96	1.01	1.00	1.00	0.99	*0.83*	*0.75*	*0.66*	*0.48*
C2 EW	0.79	0.66	0.56	0.30	0.98	0.98	0.97	*0.96*	1.00	*1.00*	0.99	*0.97*	0.84	0.75	0.68	0.50
C2 iMSE	0.78	0.66	0.54	0.27	*0.98*	*0.98*	*0.97*	0.96	*1.00*	1.00	*0.99*	0.97	0.84	0.75	0.67	0.48
Principal-components-based methods																
PC	***0.63***	0.62	0.23	0.21	1.00	1.00	0.98	0.98	1.01	1.02	0.99	0.99	0.74	*0.73*	0.45	*0.44*
PC^2	0.64	0.63	0.23	0.21	***0.66***	***0.64***	***0.29***	0.27	0.85	0.85	0.66	0.64	0.75	0.74	0.45	0.44
SPC	0.63	0.64	0.23	0.22	0.69	0.65	0.34	0.27	0.77	0.66	0.48	0.28	***0.74***	0.75	*0.45*	0.45
PC-Sieve	0.63	***0.62***	***0.23***	***0.21***	0.68	0.65	0.30	*0.27*	***0.67***	***0.65***	***0.31***	***0.27***	0.74	0.73	0.46	0.44
PC-NW	0.79	0.78	0.56	0.55	0.92	0.91	0.82	0.80	0.92	0.91	0.77	0.75	0.83	0.82	0.67	0.65
Kernel ridge regression, five-fold cross-validation																
Poly(1)	*0.64*	*0.62*	*0.24*	*0.21*	0.99	0.99	0.97	0.97	1.01	1.00	1.00	0.99	0.75	0.74	0.47	0.45
Poly(2)	0.65	0.63	0.24	0.21	*0.71*	*0.65*	*0.35*	***0.26***	*0.72*	*0.66*	*0.35*	*0.27*	0.76	0.74	0.47	0.44
Gauss	0.65	0.63	0.25	0.22	0.79	0.71	0.50	0.34	0.86	0.77	0.62	0.43	*0.75*	*0.72*	*0.44*	*0.37*
Kernel ridge regression, ten-fold cross-validation																
Poly(1)	*0.64*	*0.62*	*0.24*	*0.21*	0.99	0.99	0.97	0.97	1.01	1.01	0.99	0.99	0.75	0.73	0.47	0.44
Poly(2)	0.65	0.63	0.24	0.21	*0.71*	*0.65*	*0.35*	*0.26*	*0.72*	*0.66*	*0.35*	*0.27*	0.76	0.74	0.47	0.44
Gauss	0.65	0.63	0.24	0.22	0.79	0.71	0.50	0.34	0.86	0.77	0.62	0.43	*0.75*	*0.72*	*0.44*	*0.37*
Kernel ridge regression, leave-one-out cross-validation																
Poly(1)	*0.64*	*0.62*	*0.24*	*0.21*	0.99	0.99	0.97	0.97	1.01	1.01	1.00	0.99	0.75	0.73	0.47	0.44
Poly(2)	0.65	0.63	0.24	0.21	*0.71*	*0.65*	*0.35*	*0.26*	*0.72*	*0.66*	*0.35*	*0.27*	0.76	0.74	0.47	0.45
Gauss	0.65	0.63	0.24	0.22	0.79	0.71	0.50	0.34	0.86	0.77	0.62	0.43	*0.75*	***0.72***	***0.44***	***0.37***
Forecast combinations																
Linear	0.78	0.69	0.55	0.35	0.98	0.98	0.97	0.96	1.00	1.00	0.99	0.98	0.84	0.77	0.67	0.54
No KRR	0.71	0.66	0.41	0.29	0.83	0.81	0.66	0.62	0.88	0.85	0.74	0.68	0.79	0.75	0.56	0.48
All 5-f	*0.69*	*0.65*	*0.35*	*0.26*	*0.82*	*0.79*	*0.62*	*0.56*	*0.87*	*0.82*	*0.70*	0.62	*0.77*	*0.74*	*0.52*	0.45
All 10-f	0.69	0.65	0.35	0.27	0.82	0.79	0.62	0.56	0.87	0.82	0.70	*0.62*	0.77	0.74	0.52	*0.45*
All LOO	0.69	0.65	0.35	0.27	0.82	0.79	0.62	0.56	0.87	0.83	0.70	0.62	0.77	0.74	0.52	0.45
Diebold-Mariano tests																
Nonlin.	**24.65**	**17.27**	**39.44**	**33.06**	**22.49**	**22.73**	**26.97**	**28.14**	**21.33**	**21.74**	**25.97**	**25.91**	**21.08**	**14.24**	**37.67**	**30.73**
Kernel 5-f	**18.45**	**14.32**	**35.67**	**30.82**	**10.49**	**21.68**	**20.82**	**25.25**	**14.05**	**21.77**	**23.79**	**29.74**	**14.74**	**14.36**	**31.36**	**31.84**
Kernel 10-f	**18.51**	**14.47**	**35.81**	**30.88**	**10.20**	**21.43**	**20.61**	**25.22**	**14.05**	**21.82**	**23.73**	**28.74**	**14.84**	**14.30**	**31.92**	**32.42**
Kernel LOO	**18.65**	**14.56**	**35.89**	**30.83**	**10.05**	**21.30**	**20.43**	**25.19**	**13.78**	**21.73**	**23.71**	**28.60**	**14.88**	**14.30**	**32.15**	**32.64**

Notes: This table extends Table 2 in the article. It has the same structure as Table E.1. The f_t are now treated as latent factors and only $x_t = \Theta f_t + \eta_t$ are observed. These DGPs have no dynamic structure, so that x_{T+1} is used to forecast y_{T+1}. The benchmarks methods labeled "Comb" in the article are now labeled "C1". Two additional benchmarks labeled "C2" have been added, which use the same schemes as the corresponding "C1" benchmarks to average over all two-regressor models. Note that results for the DGPs with $R_x^2 \neq R_y^2$ are omitted in the article.

Table E.3: Relative mean squared prediction errors for the DGPs (6)-(9), with AR(1) latent factors.

DGP	Linear				Squared				Cross-product				Threshold			
$R_y^2 =$	0.4		0.8		0.4		0.8		0.4		0.8		0.4		0.8	
$R_x^2 =$	0.4	0.8	0.4	0.8	0.4	0.8	0.4	0.8	0.4	0.8	0.4	0.8	0.4	0.8	0.4	0.8
Benchmark methods																
Mean	1.05	1.05	1.10	1.10	1.08	1.08	1.16	1.16	1.08	1.08	1.17	1.17	1.04	1.04	1.06	1.06
RW	1.54	1.54	0.97	0.97	1.73	1.73	1.39	1.39	1.73	1.73	1.38	1.38	1.65	1.65	1.20	1.20
AR	0.99	0.99	*0.78*	0.78	*1.05*	*1.05*	***1.01***	***1.01***	*1.05*	*1.05*	***1.01***	***1.01***	1.01	1.01	*0.88*	0.88
SETAR	1.05	1.05	0.81	0.81	1.12	1.12	1.09	1.09	1.11	1.11	1.10	1.10	1.07	1.07	0.94	0.94
C1 EW	0.98	0.93	0.96	0.85	1.07	1.07	1.15	1.14	1.07	1.07	1.16	1.15	0.99	0.96	0.96	0.89
C1 iMSE	0.98	0.93	0.95	0.83	1.07	1.07	1.15	1.14	1.07	1.07	1.16	1.15	0.99	0.95	0.96	0.88
C1 JMA	*0.94*	0.89	0.87	*0.74*	1.09	1.09	1.16	1.16	1.09	1.09	1.17	1.17	0.97	0.93	0.90	*0.81*
C2 EW	0.95	0.89	0.88	0.75	1.07	1.07	1.14	1.14	1.07	1.07	1.15	1.15	0.97	0.93	0.91	0.82
C2 iMSE	0.94	*0.89*	0.87	0.74	1.07	1.07	1.14	1.14	1.07	1.07	1.15	1.15	*0.97*	***0.93***	0.91	0.82
Principal-components-based methods																
PC	***0.88***	***0.88***	***0.72***	***0.71***	1.09	1.09	1.16	1.16	1.09	1.09	1.17	1.17	***0.93***	*0.93*	***0.80***	***0.79***
PC^2	0.91	0.90	0.74	0.72	1.06	1.05	1.08	1.06	1.07	1.08	1.12	1.11	0.96	0.95	0.82	0.81
SPC	0.89	0.90	0.73	0.74	1.06	1.05	*1.07*	*1.05*	1.08	1.08	1.12	1.10	0.94	0.95	0.81	0.82
PC-Sieve	0.88	0.88	0.72	0.71	1.07	1.07	1.12	1.11	1.08	1.08	1.10	1.10	0.93	0.93	0.81	0.80
PC-NW	0.94	0.94	0.87	0.86	*1.05*	*1.05*	1.11	1.10	*1.05*	*1.04*	*1.09*	*1.09*	0.96	0.96	0.91	0.90
Kernel ridge regression, five-fold cross-validation																
Poly(1)	0.90	*0.88*	*0.75*	*0.72*	1.08	1.08	1.16	1.15	1.08	1.07	1.16	1.15	0.94	*0.93*	0.83	0.80
Poly(2)	0.90	0.88	0.75	0.72	*1.04*	1.03	*1.07*	*1.04*	*1.04*	*1.03*	*1.06*	*1.04*	0.95	0.93	0.84	0.81
Gauss	*0.90*	0.88	0.75	0.72	1.05	*1.03*	1.08	1.04	1.06	1.04	1.11	1.07	*0.94*	0.93	*0.83*	*0.80*
Kernel ridge regression, ten-fold cross-validation																
Poly(1)	*0.90*	*0.88*	*0.75*	*0.71*	1.08	1.08	1.15	1.15	1.08	1.07	1.16	1.15	0.94	*0.93*	*0.83*	0.80
Poly(2)	0.90	0.88	0.75	0.72	*1.04*	1.03	*1.07*	*1.03*	*1.04*	***1.03***	*1.06*	*1.04*	0.94	0.93	0.83	0.81
Gauss	0.90	0.89	0.75	0.72	1.04	*1.03*	1.07	1.04	1.06	1.05	1.11	1.07	*0.94*	0.93	0.83	*0.80*
Kernel ridge regression, leave-one-out cross-validation																
Poly(1)	*0.90*	*0.88*	*0.74*	*0.71*	1.08	1.08	1.16	1.16	1.08	1.08	1.16	1.16	0.94	*0.93*	*0.83*	*0.80*
Poly(2)	0.90	0.89	0.75	0.72	*1.04*	1.04	*1.07*	*1.04*	*1.04*	*1.03*	*1.06*	*1.04*	0.94	0.93	0.83	0.80
Gauss	0.90	0.89	0.75	0.72	1.04	*1.03*	1.07	1.04	1.06	1.04	1.10	1.06	*0.94*	0.93	0.83	0.81
Forecast combinations																
Linear	0.93	0.90	0.80	0.75	1.04	1.04	1.05	1.05	1.04	1.04	1.06	1.06	0.96	0.94	0.85	0.81
No KRR	0.91	0.90	0.77	0.73	1.03	1.03	*1.04*	1.03	***1.04***	1.04	*1.05*	*1.04*	0.95	0.94	0.83	0.81
All 5-f	*0.90*	*0.89*	*0.76*	*0.73*	1.03	***1.03***	1.05	1.03	1.04	*1.04*	1.05	1.05	*0.94*	*0.93*	*0.83*	*0.80*
All 10-f	0.91	0.89	0.76	0.73	***1.03***	1.03	1.05	*1.03*	1.04	1.04	1.06	1.05	0.94	0.93	0.83	0.80
All LOO	0.91	0.89	0.76	0.73	1.04	1.03	1.05	1.03	1.04	1.04	1.06	1.05	0.94	0.93	0.83	0.80
Diebold-Mariano tests																
Nonlin.	**11.44**	**5.95**	**19.29**	**11.32**	**2.70**	**4.03**	**4.52**	**6.62**	**1.98**	**2.74**	**2.64**	**4.15**	**8.37**	**3.82**	**14.98**	**6.21**
Kernel 5-f	**8.92**	**9.32**	**11.16**	**10.83**	-0.29	**2.81**	**-6.02**	-0.40	0.06	**3.46**	**-5.30**	-0.62	**6.72**	**7.12**	**7.84**	**8.24**
Kernel 10-f	**8.80**	**9.19**	**11.01**	**10.86**	-0.08	**2.76**	**-5.44**	0.20	-0.34	**3.49**	**-5.92**	-0.69	**6.67**	**6.90**	**7.77**	**8.21**
Kernel LOO	**8.83**	**8.94**	**11.09**	**10.73**	-0.64	1.93	**-6.12**	-1.24	-0.46	**3.14**	**-6.48**	-1.86	**6.56**	**6.64**	**7.74**	**8.18**

Notes: This table extends Table 3 in the article. It has the same structure as Table E.2, except that the "Linear" combination forecast additionally includes the RW and AR forecasts. The f_t are assumed to be latent factors following AR(1) processes, so x_T and y_T are used to forecast y_{T+1}.

Table E.4: Relative mean squared prediction errors for the threshold autoregressive DGPs (10).

DGP	Self-exciting	Observed	Weak factor	Strong factor
Benchmark methods				
Mean	1.47	1.10	1.08	1.08
RW	0.65	0.54	0.50	0.50
AR	*0.56*	0.48	***0.46***	0.46
SETAR	0.59	0.53	0.50	0.50
C1 EW	–	0.47	0.88	0.77
C1 iMSPE	–	0.45	0.87	0.76
C1 JMA	–	0.44	0.48	*0.45*
C2 EW	–	0.38	0.81	0.75
C2 iMSPE	–	*0.38*	0.79	0.74
Principal-components-based methods				
PC	–	0.39	*0.71*	0.43
PC^2	–	*0.38*	0.75	0.44
SPC	–	0.50	0.76	0.68
PC-Sieve	*0.62*	0.39	0.72	***0.43***
PC-NW	0.75	0.65	0.87	0.69
Kernel ridge regression, five-fold cross-validation				
Poly(1)	0.57	0.38	0.58	0.48
Poly(2)	0.57	0.37	0.58	0.48
Gauss	*0.55*	***0.37***	*0.55*	*0.45*
Kernel ridge regression, ten-fold cross-validation				
Poly(1)	0.57	0.38	0.58	0.48
Poly(2)	0.58	0.37	0.58	0.48
Gauss	***0.55***	*0.37*	*0.55*	*0.45*
Kernel ridge regression, leave-one-out cross-validation				
Poly(1)	0.57	0.38	0.58	0.48
Poly(2)	0.58	*0.37*	0.58	0.48
Gauss	*0.56*	0.37	*0.55*	*0.45*
Forecast combinations				
Linear	0.63	0.42	0.59	0.52
No KRR	0.59	0.41	0.58	0.49
All 5-f	0.57	0.39	0.56	*0.47*
All 10-f	0.57	0.39	*0.56*	0.47
All LOO	*0.57*	*0.39*	0.56	0.47
Diebold-Mariano tests				
Nonlin.	**10.19**	**8.67**	**2.26**	**21.19**
Kernel 5-f	**12.61**	**23.30**	**16.85**	**18.60**
Kernel 10-f	**12.99**	**23.56**	**16.95**	**18.62**
Kernel LOO	**13.91**	**24.39**	**16.92**	**18.64**

Notes: This table extends Table 4 in the article. It has the same structure as Table E.3. y_T and x_T are used to forecast y_{T+1}, except in the self-exciting DGP, where only y_T is available.

Table E.5: Application frequency of the insanity filter for the macroeconomic series.

Forecast		Industrial Production				Personal Income				Manuf. & Trade Sales				Employment			
method	$h =$	1	3	6	12	1	3	6	12	1	3	6	12	1	3	6	12
Benchmark methods																	
Mean		–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
RW		0.2	0.2	–	–	0.4	–	–	–	–	–	–	–	0.2	–	–	–
AR		–	–	–	–	0.2	–	–	–	–	–	–	–	–	–	–	–
SETAR		0.4	0.2	0.2	–	–	–	–	–	0.6	–	–	–	–	–	–	–
C1 EW		–	–	–	–	0.2	–	–	–	–	–	–	–	–	–	–	–
C1 iMSE		–	–	–	–	0.2	–	–	–	–	–	–	–	–	–	–	–
C1 JMA		–	–	–	–	0.2	–	–	–	–	–	–	–	–	–	–	–
C2 EW		–	–	–	–	0.2	–	–	–	–	–	–	–	–	–	–	–
C2 iMSE		–	–	–	–	0.2	–	–	–	–	–	–	–	–	–	–	–
Principal-components-based methods																	
PC		–	–	–	–	0.2	–	–	–	–	–	–	–	–	–	–	–
PC^2		0.2	0.4	0.4	–	0.2	–	0.2	–	–	–	–	–	–	0.2	0.4	0.7
SPC		0.2	0.6	1.5	0.7	0.2	0.4	0.2	0.2	–	0.4	1.7	0.7	–	0.8	0.6	–
PC-Sieve		0.8	0.2	–	–	0.2	–	0.2	–	–	–	–	–	–	0.2	–	–
PC-NW		–	–	–	–	0.2	–	–	–	–	–	–	–	–	–	–	–
Kernel ridge regression, five-fold cross-validation																	
Poly(1)		–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Poly(2)		0.4	0.2	–	–	–	–	–	–	–	–	0.4	–	–	–	–	–
Gauss		–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Kernel ridge regression, ten-fold cross-validation																	
Poly(1)		–	–	–	–	–	–	–	–	–	–	0.2	–	–	–	–	–
Poly(2)		0.4	–	–	–	–	–	–	–	–	–	0.4	–	–	–	0.2	–
Gauss		–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Kernel ridge regression, leave-one-out cross-validation																	
Poly(1)		–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Poly(2)		0.2	0.4	0.6	–	–	–	–	–	–	–	0.4	–	–	–	0.4	–
Gauss		–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

Notes: This table extends Table 5 in the article. It lists the percentage of forecasts to which an insanity filter was applied, for each forecast method and for each target series. If the filter was never applied, this is indicated by a dash (−). The same naming convention as in Table E.2 applies to the "C1" and "C2" benchmarks.

Table E.6: Relative mean squared prediction errors for the macroeconomic series.

Forecast	Industrial Production				Personal Income				Manuf. & Trade Sales				Employment			
method $h =$	1	3	6	12	1	3	6	12	1	3	6	12	1	3	6	12
Benchmark methods																
Mean	1.01	1.04	1.05	1.07	1.02	1.05	1.09	1.16	1.01	1.02	1.04	1.08	0.98	0.96	0.96	0.97
RW	1.18	1.09	1.36	1.61	1.20	1.36	1.14	1.34	2.18	1.49	1.45	1.50	1.58	0.95	0.99	1.19
AR	0.92	0.87	1.01	1.02	1.04	1.04	1.09	1.14	1.01	1.01	1.09	1.08	0.96	0.85	0.89	0.95
SETAR	1.01	1.27	1.17	1.06	*0.94*	1.00	1.17	1.13	1.03	1.02	1.05	1.08	0.98	0.89	0.90	0.98
C1 EW	0.85	0.81	0.90	0.91	0.97	0.94	0.93	0.98	0.98	0.96	0.98	0.98	0.89	0.75	0.76	0.81
C1 iMSE	0.85	0.81	0.89	0.87	0.97	0.94	0.92	0.95	0.98	0.96	0.96	0.94	0.89	0.74	0.74	0.78
C1 JMA	*0.75*	*0.74*	0.92	*0.71*	0.95	*0.83*	*0.81*	*0.83*	*0.94*	0.94	0.91	*0.71*	*0.81*	*0.63*	*0.66*	0.74
C2 EW	0.83	0.78	0.86	0.85	0.95	0.90	0.88	0.92	0.95	0.93	0.93	0.92	0.86	0.70	0.69	0.74
C2 iMSE	0.83	0.77	*0.84*	0.79	0.95	0.89	0.86	0.89	0.95	*0.92*	*0.91*	0.85	0.85	0.69	0.67	*0.70*
Principal-components-based methods																
PC	0.80	*0.73*	*0.76*	0.68	*0.89*	***0.78***	*0.86*	*0.87*	0.88	***0.80***	0.80	*0.68*	0.76	***0.56***	***0.49***	*0.54*
PC^2	*0.78*	0.79	0.85	0.75	0.90	0.85	0.91	0.95	0.91	0.83	*0.77*	0.77	***0.75***	0.60	0.53	0.57
SPC	0.85	0.86	0.80	0.97	0.96	0.90	0.95	1.11	0.97	1.03	0.90	1.04	0.78	0.69	0.61	0.80
PC-Sieve	0.94	0.77	0.82	*0.67*	0.90	0.79	0.92	0.98	*0.88*	0.81	0.79	0.68	0.76	0.57	0.53	0.56
PC-NW	0.90	0.82	0.86	0.83	0.94	0.95	0.90	0.94	0.95	0.93	0.92	0.91	0.86	0.70	0.68	0.73
Kernel ridge regression, five-fold cross-validation																
Poly(1)	*0.77*	*0.69*	0.82	0.72	0.89	0.83	0.80	0.80	*0.88*	0.86	0.84	*0.65*	*0.80*	0.58	0.55	0.50
Poly(2)	0.79	0.70	0.80	0.68	0.89	*0.83*	0.81	0.77	0.90	0.86	0.81	0.65	0.80	0.59	0.55	***0.48***
Gauss	0.79	0.71	*0.74*	*0.64*	***0.87***	0.83	***0.76***	*0.74*	0.89	*0.84*	*0.80*	0.66	0.80	*0.57*	*0.52*	0.51
Kernel ridge regression, ten-fold cross-validation																
Poly(1)	*0.76*	0.69	0.80	*0.61*	0.91	0.83	0.86	0.85	0.88	0.84	0.80	*0.63*	0.81	*0.59*	0.54	0.53
Poly(2)	0.78	*0.66*	0.75	0.64	0.92	0.83	0.83	0.78	0.88	0.86	*0.76*	0.64	*0.80*	0.63	0.56	*0.48*
Gauss	0.77	0.71	***0.73***	0.62	*0.89*	*0.82*	*0.82*	***0.73***	***0.87***	*0.83*	0.79	0.66	0.80	0.60	*0.53*	0.49
Kernel ridge regression, leave-one-out cross-validation																
Poly(1)	***0.73***	***0.64***	0.74	***0.60***	0.92	0.81	0.83	0.87	0.89	0.85	***0.75***	***0.59***	0.84	*0.59*	*0.54*	*0.51*
Poly(2)	0.78	0.69	*0.73*	0.64	0.90	*0.79*	0.93	0.81	0.90	0.91	0.80	0.66	0.83	0.62	0.59	0.58
Gauss	0.75	0.71	0.76	0.63	*0.88*	0.80	*0.79*	*0.77*	*0.89*	*0.85*	0.80	0.77	*0.82*	0.62	0.59	0.58
Forecast combinations																
Linear	0.80	0.76	0.83	0.80	0.92	0.87	0.83	0.88	0.95	0.89	0.88	0.83	0.85	0.67	0.66	0.73
No KRR	0.80	0.74	0.78	0.75	0.90	0.83	0.80	0.86	0.92	0.86	0.83	0.81	0.82	0.63	0.61	0.68
All 5-f	0.77	0.71	0.75	0.70	*0.89*	0.81	*0.78*	0.81	0.90	0.84	0.80	0.74	0.80	*0.61*	0.57	0.62
All 10-f	0.77	*0.71*	0.74	*0.68*	0.89	0.81	0.78	*0.81*	*0.89*	*0.84*	*0.79*	*0.74*	*0.80*	0.61	*0.57*	*0.62*
All LOO	*0.77*	0.71	*0.74*	0.69	0.89	*0.81*	0.78	0.81	0.89	0.84	0.79	0.75	0.80	0.62	0.58	0.63
Diebold-Mariano tests																
Nonlin.	1.06	1.10	**2.25**	1.77	1.89	**2.32**	1.17	1.04	**3.62**	**2.73**	**3.25**	1.30	**4.65**	**2.48**	**2.58**	**3.32**
Kernel 5-f	**3.63**	**2.70**	1.53	1.52	**2.89**	**2.49**	**2.23**	**2.35**	**3.98**	**2.35**	**2.08**	**2.02**	**3.90**	**4.01**	**2.95**	**1.98**
Kernel 10-f	**3.69**	**3.31**	1.94	1.91	**2.54**	**2.51**	1.80	**2.67**	**4.29**	**2.72**	**2.45**	**2.10**	**3.96**	**3.41**	**3.20**	**2.05**
Kernel LOO	**4.27**	**3.30**	**2.00**	1.92	**2.52**	**3.63**	**1.99**	**2.90**	**4.08**	**2.91**	**2.47**	**2.49**	**2.96**	**3.15**	**3.11**	**2.37**

Notes: This table extends Table 6 in the article. It reports mean squared prediction errors (MSPEs) for four macroeconomic series, over the period 1970-2010, relative to the variance of the series being predicted. Three sets of kernel ridge regression forecasts are shown, with tuning parameters selected using five-fold, ten-fold, and leave-one-out cross-validation, respectively. The same naming convention as in Table E.2 applies to the "C1" and "C2" benchmarks. The combination forecasts are averages of the Mean, RW, AR, C1, C2, and PC forecasts ("Linear"), all benchmark and PC-based methods forecasts ("No KRR"), all benchmark, PC-based, and five-fold kernel-based forecasts ("All 5-f"), and similarly for "All 10-f" and "All LOO". The smallest relative MSPE for each DGP (column) within each group of methods (benchmarks, PC-based, five-fold kernel-based, ten-fold kernel-based, leave-one-out kernel-based, or combinations) is printed in italics, with the overall smallest in boldface italics. The last four rows report the t statistics of Diebold-Mariano tests for equal predictive ability. "Nonlin." compares "Linear" to "No KRR"; a positive statistic indicates better performance of the latter. Similarly, "Kernel 5-f" compares "No KRR" to "All 5-f", "Kernel 10-f" compares "No KRR" to "All 10-f", and "Kernel LOO" compares "No KRR" to "All LOO". The statistic is printed in boldface if it is significant at the 5% level.

Table E.7: Estimated coefficients $\hat{\alpha}$ from the forecast combining regression (11).

Forecast	Industrial Production				Personal Income			
method	$h=1$	$h=3$	$h=6$	$h=12$	$h=1$	$h=3$	$h=6$	$h=12$
Principal-components-based methods								
PC^2	0.61 (0.33)	$0.13^{\dagger}$ (0.16)	$0.09^{\dagger}$ (0.22)	$0.12^{\dagger}$ (0.19)	$0.44^{*\dagger}$(0.22)	$0.18^{\dagger}$ (0.15)	$0.38^{*\dagger}$(0.19)	$0.23^{\dagger}$ (0.21)
SPC	$0.29^{*\dagger}$(0.11)	$0.23^{\dagger}$ (0.23)	$0.44^{*\dagger}$(0.16)	$0.12^{\dagger}$ (0.10)	$0.02^{\dagger}$ (0.13)	$0.21^{\dagger}$ (0.19)	$0.34^{\dagger}$ (0.28)	$0.15^{\dagger}$ (0.13)
PC-Sieve	$-0.24^{\dagger}$ (0.19)	$-0.15^{\dagger}$ (0.27)	$-0.21^{\dagger}$ (0.36)	0.67 (0.59)	$-0.37^{*\dagger}$(0.18)	$-0.28^{\dagger}$ (0.49)	$0.29^{*\dagger}$(0.12)	$0.01^{\dagger}$ (0.19)
PC-NW	$0.11^{\dagger}$ (0.20)	$0.28^{\dagger}$ (0.16)	$0.34^{*\dagger}$(0.16)	$0.22^{\dagger}$ (0.16)	$0.17^{\dagger}$ (0.38)	$-0.03^{\dagger}$ (0.19)	$0.40^{*\dagger}$(0.13)	$0.36^{*\dagger}$(0.12)
Kernel ridge regression, five-fold cross-validation								
Poly(1)	$0.57^{*\dagger}$(0.11)	$0.62^{*\dagger}$(0.12)	$0.33^{*\dagger}$(0.13)	$0.42^{*\dagger}$(0.19)	$0.56^{*\dagger}$(0.18)	$0.22^{\dagger}$ (0.27)	0.70^{*} (0.22)	$0.62^{*\dagger}$(0.19)
Poly(2)	$0.52^{*\dagger}$(0.14)	$0.59^{*\dagger}$(0.14)	$0.37^{*\dagger}$(0.18)	$0.50^{*\dagger}$(0.21)	$0.54^{*\dagger}$(0.15)	$0.35^{\dagger}$ (0.21)	$0.61^{*\dagger}$(0.16)	0.81^{*} (0.22)
Gauss	$0.54^{*\dagger}$(0.14)	$0.56^{*\dagger}$(0.18)	$0.58^{*\dagger}$(0.15)	0.62^{*} (0.21)	$0.65^{*\dagger}$(0.17)	$0.28^{\dagger}$ (0.22)	0.75^{*} (0.21)	0.88^{*} (0.21)
Kernel ridge regression, ten-fold cross-validation								
Poly(1)	$0.59^{*\dagger}$(0.10)	$0.59^{*\dagger}$(0.12)	$0.41^{*\dagger}$(0.10)	$0.62^{*\dagger}$(0.17)	$0.43^{*\dagger}$(0.20)	$0.26^{\dagger}$ (0.28)	$0.49^{*\dagger}$(0.22)	$0.53^{*\dagger}$(0.18)
Poly(2)	$0.53^{*\dagger}$(0.12)	$0.69^{*\dagger}$(0.11)	$0.55^{*\dagger}$(0.17)	$0.61^{*\dagger}$(0.18)	$0.37^{*\dagger}$(0.16)	$0.33^{\dagger}$ (0.22)	$0.55^{*\dagger}$(0.19)	0.74^{*} (0.18)
Gauss	$0.62^{*\dagger}$(0.13)	$0.56^{*\dagger}$(0.17)	$0.62^{*\dagger}$(0.14)	0.70^{*} (0.20)	$0.56^{*\dagger}$(0.14)	$0.31^{\dagger}$ (0.24)	0.60^{*} (0.21)	0.92^{*} (0.19)
Kernel ridge regression, leave-one-out cross-validation								
Poly(1)	$0.71^{*\dagger}$(0.11)	0.84^{*} (0.17)	$0.55^{*\dagger}$(0.13)	0.68^{*} (0.19)	$0.32^{\dagger}$ (0.17)	$0.32^{\dagger}$ (0.30)	0.59^{*} (0.24)	$0.49^{*\dagger}$(0.17)
Poly(2)	$0.56^{*\dagger}$(0.16)	$0.61^{*\dagger}$(0.18)	$0.56^{*\dagger}$(0.14)	0.62^{*} (0.20)	$0.47^{*\dagger}$(0.16)	$0.46^{*\dagger}$(0.20)	$0.38^{*\dagger}$(0.19)	$0.63^{*\dagger}$(0.16)
Gauss	$0.71^{*\dagger}$(0.13)	$0.55^{*\dagger}$(0.23)	$0.51^{*\dagger}$(0.14)	0.68^{*} (0.18)	$0.64^{*\dagger}$(0.15)	$0.42^{\dagger}$ (0.24)	0.66^{*} (0.22)	0.75^{*} (0.17)

Forecast	Manufacturing & Trade Sales				Employment			
method	$h=1$	$h=3$	$h=6$	$h=12$	$h=1$	$h=3$	$h=6$	$h=12$
Principal-components-based methods								
PC^2	$-0.04^{\dagger}$ (0.15)	$0.27^{\dagger}$ (0.17)	$0.61^{*\dagger}$(0.19)	$-0.06^{\dagger}$ (0.27)	0.67^{*} (0.18)	$0.15^{\dagger}$ (0.21)	$0.26^{\dagger}$ (0.15)	$0.38^{*\dagger}$(0.15)
SPC	$-0.04^{\dagger}$ (0.15)	$-0.08^{\dagger}$ (0.16)	$0.31^{\dagger}$ (0.19)	$-0.04^{\dagger}$ (0.09)	$0.36^{*\dagger}$(0.12)	$0.09^{\dagger}$ (0.10)	$0.12^{\dagger}$ (0.12)	$-0.17^{\dagger}$ (0.17)
PC-Sieve	–	$0.36^{*\dagger}$(0.03)	$0.53^{*\dagger}$(0.14)	0.32 (1.65)	–	$0.29^{\dagger}$ (0.27)	$-0.37^{\dagger}$ (0.21)	$0.29^{\dagger}$ (0.23)
PC-NW	$0.08^{\dagger}$ (0.21)	$-0.06^{\dagger}$ (0.22)	$0.20^{\dagger}$ (0.22)	$-0.11^{\dagger}$ (0.19)	$-0.11^{\dagger}$ (0.13)	$-0.11^{\dagger}$ (0.15)	$-0.03^{\dagger}$ (0.16)	$-0.06^{\dagger}$ (0.18)
Kernel ridge regression, five-fold cross-validation								
Poly(1)	$0.48^{*\dagger}$(0.12)	$0.29^{*\dagger}$(0.13)	$0.38^{*\dagger}$(0.14)	0.58^{*} (0.22)	$0.32^{*\dagger}$(0.11)	$0.37^{*\dagger}$(0.13)	$0.20^{\dagger}$ (0.13)	0.63^{*} (0.25)
Poly(2)	$0.41^{*\dagger}$(0.15)	$0.25^{\dagger}$ (0.20)	$0.47^{*\dagger}$(0.12)	0.59^{*} (0.23)	$0.30^{*\dagger}$(0.12)	$0.33^{*\dagger}$(0.16)	$0.24^{\dagger}$ (0.14)	0.76^{*} (0.25)
Gauss	$0.40^{*\dagger}$(0.17)	$0.30^{\dagger}$ (0.19)	$0.50^{*\dagger}$(0.13)	0.57^{*} (0.26)	$0.25^{\dagger}$ (0.13)	$0.39^{*\dagger}$(0.16)	$0.34^{*\dagger}$(0.16)	0.62^{*} (0.22)
Kernel ridge regression, ten-fold cross-validation								
Poly(1)	$0.47^{*\dagger}$(0.13)	$0.35^{*\dagger}$(0.13)	$0.50^{*\dagger}$(0.12)	0.63^{*} (0.22)	$0.29^{*\dagger}$(0.11)	$0.27^{\dagger}$ (0.15)	$0.25^{*\dagger}$(0.12)	$0.53^{*\dagger}$(0.23)
Poly(2)	$0.50^{*\dagger}$(0.14)	$0.25^{\dagger}$ (0.15)	$0.59^{*\dagger}$(0.13)	0.62^{*} (0.24)	$0.33^{*\dagger}$(0.12)	$0.13^{\dagger}$ (0.18)	$0.22^{\dagger}$ (0.12)	0.72^{*} (0.24)
Gauss	$0.51^{*\dagger}$(0.17)	$0.35^{\dagger}$ (0.18)	$0.51^{*\dagger}$(0.13)	0.60^{*} (0.27)	$0.26^{\dagger}$ (0.14)	$0.20^{\dagger}$ (0.16)	$0.31^{*\dagger}$(0.15)	0.67^{*} (0.21)
Kernel ridge regression, leave-one-out cross-validation								
Poly(1)	$0.44^{*\dagger}$(0.14)	$0.27^{\dagger}$ (0.14)	$0.66^{*\dagger}$(0.16)	0.82^{*} (0.28)	$0.20^{\dagger}$ (0.11)	$0.22^{\dagger}$ (0.18)	$0.18^{\dagger}$ (0.16)	0.62^{*} (0.27)
Poly(2)	$0.40^{*\dagger}$(0.14)	$0.12^{\dagger}$ (0.15)	$0.49^{*\dagger}$(0.16)	0.60^{*} (0.26)	$0.19^{\dagger}$ (0.12)	$0.07^{\dagger}$ (0.19)	$0.16^{\dagger}$ (0.11)	$0.38^{\dagger}$ (0.19)
Gauss	$0.44^{*\dagger}$(0.17)	$0.29^{\dagger}$ (0.18)	$0.51^{*\dagger}$(0.14)	$0.14^{\dagger}$ (0.23)	$0.18^{\dagger}$ (0.13)	$0.12^{\dagger}$ (0.18)	$0.16^{\dagger}$ (0.19)	$0.37^{\dagger}$ (0.22)

Notes: This table extends Table 7 in the article. It reports $\hat{\alpha}$, the weight placed on the candidate forecast in the forecast combining regression (11). HAC standard errors follow in parentheses. An asterisk ($*$) indicates rejection of the hypothesis $\alpha = 0$ and a dagger (†) indicates rejection of $\alpha = 1$, at 5% significance. The PC and PC-Sieve forecasts of the one-month growth rates of Manufacturing & Trade Sales and Employment are equal in every estimation window; hence, no $\hat{\alpha}$ can be computed in these two cases.