Online Appendix for:

How far can we forecast?
Statistical tests of the predictive content

by *Jörg Breitung* and *Malte Knüppel*

OA.1 Monte Carlo experiments with multiple forecast horizons

In order to gain insights into the small sample behaviour of the tests in typical time-series environments with multiple forecast horizons, we conduct Monte Carlo experiments based on the cases displayed in Table OA.1.[1] In the first four cases, univariate DGPs are considered, whereas the last case refers to the simple multivariate DGP used in Section 5 of the main text. The first three DGPs are moving-average processes considered in Stock and Watson (2007) to describe the behaviour of first differences of quarterly US inflation. The first process is based on their MA(1)-model estimated for the post-1984 period, whereas the second process refers to their pre-1984 estimation results. The third process is based on the quarterly version of Nelson and Schwert's (1977) model reported in Stock and Watson (2007). The fourth process uses the estimation result for an AR(1)-process of quarterly growth of US real GDP from 1996 to 2016 which broadly corresponds to the sample employed in Section 6 of the main text. For all four univariate cases, the forecast models assume an AR(1) process. While these models are misspecified in the first three cases, the parameters θ_h will be determined such that the respective null hypotheses $h^* = 1$ and $h^* = 2$ are nevertheless correct. For the case $AR(1)$-$AR(1)$, a finite h^* does not exist, because the forecasts are informative for all $h < \infty$. The last process mimics a forecasting equation for financial returns and implies an R^2 of about 0.04 for $h = 1$. This would be considered a "large" value in forecasting stock price returns given the usual empirical results as reported, for example, in Fama and French (1988). In this case, the maximum forecast horizon is $h^* = 1$. All processes are specified such that the variance of Y_t equals 1.

Forecasts are assumed to be made in a direct manner. That is, the parameters of the forecast models depend on the forecast horizon h. For the scenario with forecasts being equal to their conditional means (scenario 1) and the corresponding scenario with noise (scenario 2), the parameters θ_h are set to their pseudo-true values implied by the DGP and the forecast model with direct forecasts. This leads, for instance, to $\theta_1 = -0.28/(1 + 0.28^2)$ in the case $MA(1)_a$-$AR(1)$. Of course, direct forecasts imply $\theta_h = 0$ for $h > h^*$ in all cases considered. Correspondingly, for the scenario of model predictions (scenario 3), the parameters of the forecast models result from regressions of Y_t on the explanatory variables known at time $t - h$.[2] We consider 1- to 4-step-ahead forecasts.

---

[1] We do not consider the test of Giacomini and White (2006) in these experiments, because it is difficult to implement their null hypothesis into the DGPs and because of the lower power of the test.

[2] That is, Y_t is regressed on a constant and X_{t-h} in the case *multivar*, and on a constant and Y_{t-h} in the other cases.

Table OA.1: Cases considered in our Monte Carlo study

case	DGP	h^*	forecast model	$R^2(h^*)$
$MA(1)_a$-$AR(1)$	$Y_t = \varepsilon_t - 0.28\varepsilon_{t-1}$	1	$\widehat{Y}^*_{t+h} = \theta_h Y_t$	0.07
$MA(1)_b$-$AR(1)$	$Y_t = \varepsilon_t - 0.66\varepsilon_{t-1}$	1	$\widehat{Y}^*_{t+h} = \theta_h Y_t$	0.21
$MA(2)$-$AR(1)$	$Y_t = \varepsilon_t - 0.49\varepsilon_{t-1} - 0.16\varepsilon_{t-2}$	2	$\widehat{Y}^*_{t+h} = \theta_h Y_t$	0.02
$AR(1)$-$AR(1)$	$Y_t = 0.42 Y_{t-1} + \varepsilon_t$	$-$	$\widehat{Y}^*_{t+h} = \theta_h Y_t$	$-$
$multivar$	$Y_t = 0.2 X_{t-1} + \varepsilon_t$	1	$\widehat{Y}^*_{t+h} = \theta_h X_t$	0.04

Note: ε_t and X_t are $\overset{iid}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2)$. h^* is the maximum forecast horizon. $R^2(h^*)$ is the asymptotic R^2 of the forecast model at horizon h^*. In each case, σ_ε^2 is chosen such that the variance of Y_t is equal to 1.

The forecast models in Table OA.1 coincide with the conditional mean only when the test statistics $dm_{0,h}$ and $dm_{T,h}$ are applied. For all other tests, the forecasts are modified as described in Table OA.2. For tests in the scenario with noise, the magnitude of the variance of the noise term added to the conditional mean forecast is comparable to the largest magnitude considered in Section 5.[3] Note that for the tests of $\beta_h = 0.5$ and $\gamma_h = 0.5$, the null hypotheses do not hold for $h > h^*$, because the MSPE of the noisy forecasts is larger than (i.e. not equal to) the MSPE of the benchmark. Therefore, the tests will reject less frequently than suggested by the size. Actually, in the case $AR(1)$-$AR(1)$, h^* for the tests of $\beta_h = 0.5$ and $\gamma_h = 0.5$ differs from h^* for the tests of $\beta_h = 0$ and $\gamma_h = 0$. While a finite h^* does not exist in the latter situation, h^* equals 2 in the former setting. The reason is that even without noise, the MSPE would only be marginally lower than the variance of Y_t at $h = 3$. With noise, the MSPE exceeds the variance at this horizon. In the scenario with model predictions, for tests of the null hypotheses $\beta_h = 0$ and $\gamma_h = 0$, the forecast models are estimated including a constant.[4] The standard errors of the test statistics are calculated as in Newey and West (1987) using the automatic lag length selection procedure proposed by Andrews (1991). We employ a significance level of 0.05.

Table OA.3 displays the results for the $MA_a(1)$ model, hence $h^* = 1$. The evaluation sample includes $n = 50$ or $n = 100$ forecasts. In scenario 3, the initial estimation samples are based on $T = 100$ observations, and a recursive estimation scheme is employed. The tests are conducted sequentially for the forecast horizons $h = 1, 2, 3, 4$. The last forecast horizon where the test rejects is identified as horizon $\widehat{h}^*$. If the test does not reject for any horizon, we conclude that $\widehat{h}^* \geq 4$.

---

[3] The largest value in Section 5 of the main text is $\sigma_\eta = 0.1$ which is identical to the value used here, but the variance of Y_t in Section 5 is $1 + b^2$, whereas it equals 1 here. However, the differences are very small, since $|b|$ does not exceed 0.2.

[4] As mentioned above, since we are considering direct forecasts, the model parameters are estimated separately for each h.

Table OA.2: Modifications of forecasts depending on scenario

scenario	test stat./hypothesis	forecasts
1: without noise	$dm_{0,h},\ dm_{T,h}$	$\widehat{Y}_{t+h} = \widehat{Y}^*_{t+h}$
2: with noise	$\beta_h = 0;\ \beta_h = 0.5$ $\gamma_h = 0;\ \gamma_h = 0.5$	$\widehat{Y}_{t+h} = \widehat{Y}^*_{t+h} + \eta_t$
3: model prediction	$\beta_h = 0$ $\gamma_h = 0$	$\widehat{Y}_{t+h} = \begin{cases} \widehat{\theta}_{1,h} + \widehat{\theta}_{2,h} X_t & \text{case } \textit{multivar} \\ \widehat{\theta}_{1,h} + \widehat{\theta}_{2,h} Y_t & \text{otherwise} \end{cases}$

Note: η_t is $\overset{iid}{\sim} \mathcal{N}(0, \sigma_\eta^2)$ with $\sigma_\eta = 0.1$. $\widehat{\theta}_{1,h}$ and $\widehat{\theta}_{2,h}$ are estimated by regressing Y_t on a constant and on X_{t-h} in the case *multivar*, and on a constant and Y_{t-h} in all other cases considered.

With $n = 50$ most tests have a power of about 0.6 to 0.7 at $h = 1$. In the scenario without noise and with the test $dm_{0,h}$, the power almost reaches 0.8. In the scenario with noise, the null hypotheses $\beta_h = 0.5$ and $\gamma_h = 0.5$ are rejected in 25% of the simulations only. For $n = 100$, the power increases to about 0.3 for these tests, while all other tests attain a power of about 0.8 to 0.9. The power of the tests at $h = 1$ broadly corresponds to the percentage of simulations in which h^* is identified correctly. The latter percentage is marginally lower than the power, which is due to those cases where the tests also reject at $h > 1$. The percentage of these cases is mostly close to the significance level of the tests, i.e. to 5%.

For the remaining Monte Carlo experiments, we do not report results for the tests $dm_{0,h}$ and $dm_{T,h}$ because the underlying scenario without noise appears unrealistic in empirical applications, and violations of this assumption can easily cause pronounced size distortions as found in Section 5. Moreover, in what follows, we will not consider the tests based on γ_h in the scenario with noise, because with the magnitude of noise used here, the results are very similar to those based on β_h.

Table OA.4 contains the results for the cases $MA(1)_b$-$AR(1)$ and $MA(2)$-$AR(1)$. Concerning the case $MA_b(1)$-$AR(1)$, the absolute value of the MA-coefficient is much larger than in the case $MA_a(1)$-$AR(1)$, making it considerably easier for the tests to detect $h^* = 1$. The MSPE-variance ratio equals about 0.8 at $h = 1$, and the tests for $\beta_h = 0$ and $\gamma_h = 0$ virtually always reject at this horizon even with $n = 50$. $\widehat{h}^*$ is equal to h^* in 90% to 95% of the simulations, and it is mostly larger than h^* otherwise. Only the test for $\beta_h = 0.5$ has a considerably lower power of 0.5 with $n = 50$ and of 0.7 with $n = 100$ at $h = 1$. The share of correct identifications of h^* is very similar to the power at $h = 1$.

In the case $MA(2)$-$AR(1)$, the second MA-coefficient is very close to zero, making it difficult to identify $h^* = 2$. Note that the MSPE-variance ratio is virtually equal to 1 at this horizon. While for $h = 1$, the tests for $\beta_h = 0$ and $\gamma_h = 0$ reject in about 85% of the replications with $n = 50$ and in almost all of the replications with $n = 100$, these numbers are considerably lower for $h = 2$. The most frequent value of $\widehat{h}^*$ equals 1 for these tests. Given the large MSPE-variance

Table OA.3: Results for case $MA(1)_a$-$AR(1)$

h	0	1	2	3	4	0	1	2	3	4
			$n=50$					$n=100$		
			without noise (scenario 1)							
$\frac{\text{MSPE}}{\text{variance}}$		0.95	1.01	1.01	1.01		0.94	1.01	1.01	1.01
				rejections						
$dm_{h,0}$		0.77	0.05	0.05	0.05		0.89	0.05	0.05	0.05
$dm_{h,T}$		0.66	0.04	0.04	0.04		0.81	0.04	0.04	0.04
				$\widehat{h}^*$						
$dm_{h,0}$	0.23	**0.73**	0.00	0.00	0.04	0.11	**0.84**	0.00	0.00	0.05
$dm_{h,T}$	0.34	**0.62**	0.00	0.00	0.03	0.19	**0.77**	0.00	0.00	0.03
			with noise (scenario 2)							
$\frac{\text{MSPE}}{\text{variance}}$		0.96	1.02	1.02	1.02		0.95	1.02	1.02	1.01
				rejections						
$\beta_h = 0$		0.64	0.07	0.07	0.07		0.85	0.06	0.06	0.07
$\gamma_h = 0$		0.66	0.10	0.11	0.10		0.88	0.09	0.09	0.10
$\beta_h = 0.5$		0.25	0.04	0.04	0.04		0.32	0.02	0.02	0.02
$\gamma_h = 0.5$		0.25	0.04	0.04	0.04		0.33	0.03	0.03	0.03
				$\widehat{h}^*$						
$\beta_h = 0$	0.36	**0.59**	0.05	0.00	0.00	0.15	**0.79**	0.05	0.00	0.00
$\gamma_h = 0$	0.34	**0.59**	0.06	0.01	0.01	0.12	**0.80**	0.07	0.01	0.00
$\beta_h = 0.5$	0.75	**0.24**	0.01	0.00	0.00	0.68	**0.31**	0.01	0.00	0.00
$\gamma_h = 0.5$	0.75	**0.24**	0.01	0.00	0.00	0.67	**0.32**	0.01	0.00	0.00
			model prediction (scenario 3)							
$\frac{\text{MSPE}}{\text{variance}}$		0.96	1.02	1.02	1.02		0.95	1.02	1.02	1.02
				rejections						
$\beta_h = 0$		0.60	0.04	0.05	0.04		0.82	0.03	0.03	0.03
$\gamma_h = 0$		0.63	0.05	0.05	0.05		0.85	0.04	0.04	0.04
				$\widehat{h}^*$						
$\beta_h = 0$	0.40	**0.58**	0.02	0.00	0.00	0.18	**0.80**	0.02	0.00	0.00
$\gamma_h = 0$	0.37	**0.60**	0.02	0.00	0.00	0.15	**0.82**	0.02	0.00	0.00

Note: The values displayed in the category '*rejections*' denote the share of rejections for each horizon h. The values displayed in the category '$\widehat{h}^*$' denote the share of cases in which h is identified as the maximum forecast horizon. The significance level is set to 0.05. 'MSPE' is the mean-squared prediction error, 'variance' is the variance of the target variable in the evaluation sample. Bold entries refer to the true h^*. If a test rejects for all horizons, $\widehat{h}^*$ is set equal to the largest horizon $h = 4$. The forecasts differ between the 3 scenarios according to the modifications described in Table OA.2. For the tests corresponding to the model-based predictions, the estimation is carried out recursively with $T = 100$ observations used for the first parameter estimation. n is the number of observations for evaluation. The share of rejections for $h \leq \widehat{h}^*$ corresponds to the power of the tests. The share of rejections for $h > \widehat{h}^*$ corresponds to the size of the tests (except for $\beta_h = 0.5$ and $\gamma_h = 0.5$). The variance in the MSPE-variance-ratio is calculated dividing by n. All tests are one-sided. Results are based on 10,000 simulations.

ratio at $h^* = 2$, the test for $\beta_h = 0.5$ almost never identifies $\widehat{h}^* = 2$. Thus, with the MA(2)-specification chosen here, larger evaluation samples are needed in order to reliably determine h^*.

The results for the cases $AR(1)$-$AR(1)$ and *multivar* can be found in Table OA.5. Due to the persistence of the $AR(1)$-process, the sample variance is downward biased in small samples. Apparently, this bias, together with the effects of noise or estimation error, leads to MSPE-variance ratios often considerably larger than 1 for $h > 1$. Nevertheless, the tests for $\beta_h = 0$ and $\gamma_h = 0$ reject at $h = 2$ in about 15% to 40% of the replications. Yet, in about 60% to 70% of the simulations, $\widehat{h}^* = 1$ is found, and the correct identification of $h^* \geq 4$ almost never occurs. This result, however, is not surprising given the sample sizes under study, and the fact that the population correlation between Y_{t+4} and Y_t equals 0.03 only. The test for $\gamma_h = 0$ has a moderately larger power than the test for $\beta_h = 0$, which is more conservative. Therefore, the test for $\gamma_h = 0$ tends to obtain larger values of $\widehat{h}^*$ here. The test for $\beta_h = 0.5$ finds $\widehat{h}^* = 0$ in about 70% of the simulations with $n = 50$, and still in about 50% of the simulations with $n = 100$.

Concerning the multivariate case, the results are strongly related to those displayed in Tables 1, 2 and 3 in the main text. For $h = 1$, the power of the tests ranges from about 0.3 with model predictions, $\beta_h = 0$ and $n = 50$ to 0.55 with model predictions, $\gamma_h = 0$ and $n = 100$. For $h > 1$, the tests have an actual size close to their nominal size of 5%, with the tests for $\beta_h = 0$ being moderately liberal in the scenario with noise, and conservative in the scenario with model predictions. The test for $\beta_h = 0.5$ practically never rejects for $h > 1$, and also fails to reject at $h = 1$ in more than 80% of the simulations.

Finally, it might be interesting to see how many observations are needed in the cases $MA(2)$-$AR(1)$ and $AR(1)$-$AR(1)$ to find $\widehat{h}^* > 1$ with a large probability. Table OA.6 contains results for sample sizes of $n = 250$ and $n = 500$. For both sample sizes and in both cases, the MSPE-variance ratio equals 0.99 at $h = 2$, but exceeds 1 for $h > 2$. While in the case $MA(2)$-$AR(1)$ with $n = 250$, $\widehat{h}^* = 1$ continues to be the most frequent result, $\widehat{h}^* = 2$ is the most likely result of the tests for $\beta_h = 0$ and $\gamma_h = 0$ when $n = 500$. In the case $AR(1)$-$AR(1)$ with $n = 500$, the test for $\gamma_h = 0$ yields $\widehat{h}^* = 3$ in about 20% of the simulations, but still hardly delivers $\widehat{h}^* \geq 4$. The test for $\beta_h = 0.5$ mostly gives $\widehat{h}^* = 1$.[5]

To summarize, with the small sample sizes considered in the simulations, the forecasts $\widehat{Y}_{t+h|t}$ need to be at least moderately correlated with the target variable Y_{t+h} at h^* for the tests of $\beta_h = 0$ and $\gamma_h = 0$ to correctly identify h^*. For model predictions and a known value of T, the test for $\gamma_h = 0$ appears to be preferable due to its marginally larger power. For other situations, the test for $\beta_h = 0$ seems more appealing, because it does not require knowledge of or decisions about T but gives similar results.

---

[5]In additional simulations not reported here, we find that with $n = 10,000$ the tests for $\beta_h = 0.5$ and $\gamma_h = 0.5$ virtually always correctly identify $h^* = 2$. The tests for $\beta_h = 0$ and $\gamma_h = 0$ yield $\widehat{h}^* \geq 4$ in about 30% and 55% of the simulations, respectively.

Table OA.4: Results for cases $MA(1)_b$-$AR(1)$ and $MA(2)$-$AR(1)$

h	0	1	2	3	4	0	1	2	3	4
			$n=50$					$n=100$		
					$MA(1)_b - AR(1)$					
					with noise (scenario 2)					
$\frac{\text{MSPE}}{\text{variance}}$		0.82	1.01	1.01	1.01		0.81	1.01	1.01	1.01
					rejections					
$\beta_h = 0$		0.99	0.09	0.08	0.09		1.00	0.07	0.07	0.07
$\beta_h = 0.5$		0.51	0.04	0.04	0.04		0.72	0.03	0.03	0.02
					$\widehat{h}^*$					
$\beta_h = 0$	0.01	**0.90**	0.08	0.01	0.00	0.00	**0.93**	0.06	0.00	0.00
$\beta_h = 0.5$	0.49	**0.48**	0.02	0.00	0.00	0.28	**0.70**	0.02	0.00	0.00
					model prediction (scenario 3)					
$\frac{\text{MSPE}}{\text{variance}}$		0.81	1.01	1.01	1.01		0.80	1.01	1.01	1.01
					rejections					
$\beta_h = 0$		0.99	0.07	0.06	0.06		1.00	0.05	0.05	0.05
$\gamma_h = 0$		0.99	0.06	0.06	0.06		1.00	0.05	0.05	0.04
					$\widehat{h}^*$					
$\beta_h = 0$	0.01	**0.92**	0.05	0.01	0.00	0.00	**0.95**	0.04	0.01	0.00
$\gamma_h = 0$	0.01	**0.93**	0.05	0.01	0.00	0.00	**0.95**	0.04	0.01	0.00
					$MA(2) - AR(1)$					
					with noise (scenario 2)					
$\frac{\text{MSPE}}{\text{variance}}$		0.91	1.00	1.01	1.01		0.91	1.00	1.01	1.01
					rejections					
$\beta_h = 0$		0.86	0.25	0.08	0.08		0.98	0.31	0.06	0.06
$\beta_h = 0.5$		0.32	0.12	0.04	0.04		0.46	0.10	0.02	0.02
					$\widehat{h}^*$					
$\beta_h = 0$	0.14	0.67	**0.18**	0.01	0.00	0.02	0.68	**0.28**	0.02	0.00
$\beta_h = 0.5$	0.68	0.31	**0.01**	0.00	0.00	0.54	0.44	**0.02**	0.00	0.00
					model prediction (scenario 3)					
$\frac{\text{MSPE}}{\text{variance}}$		0.92	1.00	1.01	1.01		0.91	1.00	1.01	1.01
					rejections					
$\beta_h = 0$		0.84	0.21	0.06	0.06		0.98	0.25	0.04	0.04
$\gamma_h = 0$		0.88	0.21	0.05	0.05		0.99	0.26	0.04	0.04
					$\widehat{h}^*$					
$\beta_h = 0$	0.16	0.71	**0.12**	0.01	0.00	0.02	0.74	**0.23**	0.01	0.00
$\gamma_h = 0$	0.12	0.73	**0.14**	0.01	0.00	0.01	0.73	**0.24**	0.01	0.00

Note: See Table OA.3 for explanations.

Table OA.5: Results for cases $AR(1)$-$AR(1)$ and *multivar*

h	0	1	2	3	4	0	1	2	3	4
			$n=50$					$n=100$		
					$AR(1)-AR(1)$					
					with noise (scenario 2)					
$\frac{\text{MSPE}}{\text{variance}}$		0.89	1.04	1.06	1.06		0.86	1.01	1.03	1.03
					rejections					
$\beta_h = 0$		0.88	0.26	0.10	0.07		0.99	0.41	0.12	0.07
$\beta_h = 0.5$		0.34	0.09	0.05	0.04		0.54	0.11	0.04	0.03
					$\widehat{h}^*$					
$\beta_h = 0$	0.12	0.63	0.20	0.05	**0.01**	0.01	0.58	0.33	0.07	**0.01**
$\beta_h = 0.5$	0.66	0.27	**0.06**	0.01	0.00	0.46	0.44	**0.08**	0.01	0.00
					model prediction (scenario 3)					
$\frac{\text{MSPE}}{\text{variance}}$		0.89	1.05	1.08	1.09		0.86	1.02	1.05	1.05
					rejections					
$\beta_h = 0$		0.87	0.16	0.04	0.03		0.99	0.27	0.04	0.02
$\gamma_h = 0$		0.92	0.27	0.09	0.06		0.99	0.41	0.09	0.05
					$\widehat{h}^*$					
$\beta_h = 0$	0.13	0.71	0.14	0.01	**0.00**	0.01	0.72	0.24	0.02	**0.00**
$\gamma_h = 0$	0.08	0.66	0.21	0.04	**0.01**	0.01	0.59	0.33	0.06	**0.02**
					multivar					
					with noise (scenario 2)					
$\frac{\text{MSPE}}{\text{variance}}$		0.99	1.03	1.03	1.03		0.98	1.02	1.02	1.02
					rejections					
$\beta_h = 0$		0.39	0.07	0.07	0.07		0.57	0.06	0.07	0.06
$\beta_h = 0.5$		0.15	0.04	0.04	0.03		0.18	0.02	0.02	0.02
					$\widehat{h}^*$					
$\beta_h = 0$	0.61	**0.36**	0.03	0.00	0.00	0.43	**0.54**	0.03	0.00	0.00
$\beta_h = 0.5$	0.85	**0.15**	0.01	0.00	0.00	0.82	**0.18**	0.00	0.00	0.00
					model prediction (scenario 3)					
$\frac{\text{MSPE}}{\text{variance}}$		1.00	1.04	1.04	1.04		0.99	1.02	1.02	1.02
					rejections					
$\beta_h = 0$		0.33	0.03	0.03	0.03		0.50	0.02	0.02	0.02
$\gamma_h = 0$		0.39	0.05	0.05	0.05		0.57	0.03	0.04	0.04
					$\widehat{h}^*$					
$\beta_h = 0$	0.67	**0.32**	0.01	0.00	0.00	0.50	**0.49**	0.01	0.00	0.00
$\gamma_h = 0$	0.61	**0.37**	0.02	0.00	0.00	0.43	**0.55**	0.02	0.00	0.00

Note: See Table OA.3 for explanations. For the case '$AR(1)$-$AR(1)$', bold entries refer to the largest h tested except for the test of $\beta_h = 0.5$. In the latter case $h^* = 2$, because the MSPE exceeds the variance of Y_t at $h = 3$.

7

Table OA.6: Results for cases $MA(2)$-$AR(1)$ and $AR(1)$-$AR(1)$ with large n

h	0	1	2	3	4	0	1	2	3	4
			$n=250$					$n=500$		
					$MA(2)-AR(1)$					
					with noise (scenario 2)					
$\frac{\text{MSPE}}{\text{variance}}$		0.91	0.99	1.01	1.01		0.91	0.99	1.01	1.01
					rejections					
$\beta_h = 0$		1.00	0.49	0.06	0.06		1.00	0.70	0.05	0.05
$\beta_h = 0.5$		0.76	0.10	0.01	0.01		0.95	0.11	0.00	0.00
					$\widehat{h}^*$					
$\beta_h = 0$	0.00	0.51	**0.46**	0.02	0.00	0.00	0.30	**0.67**	0.03	0.00
$\beta_h = 0.5$	0.24	0.72	**0.05**	0.00	0.00	0.05	0.86	**0.09**	0.00	0.00
					model prediction (scenario 3)					
$\frac{\text{MSPE}}{\text{variance}}$		0.90	0.99	1.01	1.01		0.90	0.99	1.01	1.01
					rejections					
$\beta_h = 0$		1.00	0.43	0.03	0.03		1.00	0.66	0.03	0.03
$\gamma_h = 0$		1.00	0.45	0.03	0.03		1.00	0.69	0.03	0.03
					$\widehat{h}^*$					
$\beta_h = 0$	0.00	0.57	**0.41**	0.01	0.00	0.00	0.34	**0.64**	0.01	0.00
$\gamma_h = 0$	0.00	0.55	**0.44**	0.01	0.00	0.00	0.31	**0.67**	0.01	0.00
					$AR(1)-AR(1)$					
					with noise (scenario 2)					
$\frac{\text{MSPE}}{\text{variance}}$		0.85	0.99	1.02	1.02		0.84	0.99	1.01	1.01
					rejections					
$\beta_h = 0$		1.00	0.72	0.17	0.08		1.00	0.93	0.24	0.08
$\beta_h = 0.5$		0.87	0.16	0.03	0.01		0.99	0.23	0.02	0.01
					$\widehat{h}^*$					
$\beta_h = 0$	0.00	0.28	0.57	0.13	**0.02**	0.00	0.07	0.69	0.21	**0.03**
$\beta_h = 0.5$	0.13	0.71	**0.14**	0.01	0.00	0.01	0.76	**0.21**	0.01	0.00
					model prediction (scenario 3)					
$\frac{\text{MSPE}}{\text{variance}}$		0.84	0.99	1.02	1.03		0.84	0.99	1.01	1.02
					rejections					
$\beta_h = 0$		1.00	0.56	0.06	0.01		1.00	0.84	0.10	0.02
$\gamma_h = 0$		1.00	0.71	0.14	0.05		1.00	0.93	0.23	0.06
					$\widehat{h}^*$					
$\beta_h = 0$	0.00	0.44	0.50	0.05	**0.01**	0.00	0.16	0.74	0.09	**0.01**
$\gamma_h = 0$	0.00	0.29	0.58	0.11	**0.03**	0.00	0.07	0.70	0.19	**0.04**

Note: See Table OA.3 for explanations. For the case '$AR(1)$-$AR(1)$', bold entries refer to the largest h tested except for the test of $\beta_h = 0.5$. In the latter case $h^* = 2$, because the MSPE exceeds the variance of Y_t at $h = 3$.
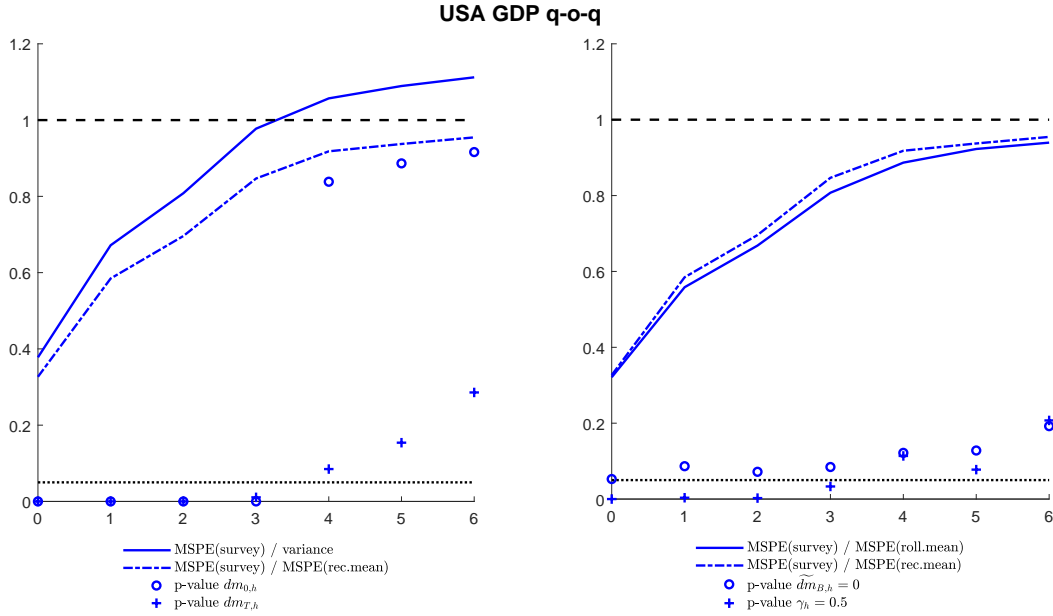
8

USA GDP q-o-q

Figure OA.1: Additional test results for quarter-on-quarter growth rates of US real GDP. The tests in the left panel assume that forecasts are conditional means without noise. The test based on $dm_{0,h}$ uses the in-sample mean as the benchmark. The solid line indicates the corresponding ratio of the Consensus forecasts' MSPE to the variance. The test based on $dm_{T,h}$ uses the recursive mean with $T = 20$ initial observations as the benchmark. The dash-dotted line indicates the corresponding ratio of the Consensus forecasts' MSPE to the MSPE of the recursive mean forecasts. The Giacomini-White test for a loss differential of zero ($\widetilde{dm}_{B,h} = 0$) in the right panel uses the mean from a rolling estimation window with $B = 20$ observations as the benchmark. The solid line indicates the corresponding ratio of the Consensus forecasts' MSPE to the MSPE of the rolling mean forecasts. The test for $\gamma_h = 0.5$ is identical to the test in Figure 2 in the main text and is displayed for comparison only. For further explanations, see Figure 1 in the main text.

OA.2 Further empirical results

The sequential p-values of the tests $dm_{0,h}$, $dm_{T,h}$, and the test of Giacomini and White (2006) when applied to US real GDP growth are shown in Figure OA.1. The test results for the scenario without noise are displayed in the left panel, and both tests arrive at $\widehat{h}^* = 3$. The same result was obtained with the test for $\beta_h = 0$ (see Table 5). The test of Giacomini and White (2006) is shown in the right panel together with the test for $\gamma_h = 0.5$ for comparison. While the ratios of the survey forecasts' MSPEs to the MSPEs of the respective benchmark forecasts (rolling and recursive means) are very similar, the p-values of the test by Giacomini and White (2006) are larger than those of the test for $\gamma_h = 0.5$ for most horizons. Since the p-value of the former test is larger than 5% even for the nowcast, the survey forecasts are found to be uninformative at all horizons.

9

OA.3 Proofs

OA.3.1 Proof of Theorem A.2:

Under the null hypothesis (2), Assumptions 1 – 2 and with the recursive mean as the uninformative benchmark, we obtain

$$\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}\left[(Y_{t+h}-\widehat{Y}_{t+h|t})^2-(Y_{t+h}-\overline{Y}_t)^2\right]$$

$$=\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}\left[(u_{h,t}-\eta_t)^2-\left(\mu_{h,t}^*+u_{h,t}\right)^2\right]\ =\ 0$$

where

$$\mu_{h,t}^* = \mu_{h,t}-\mu-\zeta_{T,t}$$
$$\zeta_{T,t} = \overline{Y}_t-\mu\ =\ O_p\left((T+t)^{-1/2}\right).$$

and $\mathbb{E}(\mu_{h,t}^* u_{h,t})=0$. It follows that the null hypothesis implies:

$$\sigma_\eta^2 = \mathbb{E}\left[\frac{1}{n}\sum_{t=1}^{n}(\mu_{h,t}^*)^2\right].$$

Consider the numerator of the statistic τ_a for the hypothesis $\gamma_h=0.5$ with

$$a_t = (Y_{t+h}-\overline{Y}_t)(\widehat{Y}_{t+h|t}-\overline{Y}_t)-\frac{1}{2}(\widehat{Y}_{t+h|t}-\overline{Y}_t)^2$$

$$= (\mu_{h,t}^*+u_{h,t})(\mu_{h,t}^*+\eta_t)-\frac{1}{2}(\mu_{h,t}^*+\eta_t)^2$$

$$= u_{h,t}(\mu_{h,t}^*+\eta_t)+\frac{1}{2}(\mu_{h,t}^*)^2-\frac{1}{2}\eta_t^2.$$

Let $S_t = \sum_{s=-T+1}^{t}(Y_s-\mu)$. According to Assumption 1 a functional central limit theorem applies (see Phillips and Solo 1992) such that

$$\frac{1}{\omega_y\sqrt{T+n}}S_{[a(T+n)]} \Rightarrow W(a)$$

where ω_y^2 denotes the long-run variance of Y_t. Then $u_{h,t}\zeta_{T,t} = (T+t)^{-1}u_{h,t}S_t^h$ and

$$\sum_{t=1}^{n}u_{h,t}\zeta_{T,t} = \sum_{t=1}^{n}\frac{T+n}{T+t}\frac{S_t}{\sqrt{T+n}}\frac{u_{h,t}}{\sqrt{T+n}}$$

$$\Rightarrow \omega_y\omega_h\int_\pi^1\frac{1}{a}W(a)\mathrm{d}W_h(a)$$

10

where $(T+n)^{-1/2} \sum_{s=-T+1}^{[a(T+n)]} u_{h,s} \Rightarrow \omega_h W_h(a)$ and $\pi = T/(T+n)$. Therefore, $\sum_{t=1}^{n} u_{h,t} \zeta_{T,t}$ is $O_p(1)$. It follows that

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} a_t = \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} u_{h,t}(\mu_{h,t} - \mu + \eta_t) + \frac{1}{2}(\mu_{h,t}^*)^2 - \frac{1}{2}\eta_t^2 \right) + O_p(n^{-1/2})$$

By Assumptions 1 and 2 u_t is uncorrelated with $\mu_{h,t}$ and η_t yielding

$$\mathbb{E}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} u_{h,t}(\mu_{h,t} - \mu + \eta_t) \right) = 0.$$

Furthermore, the null hypothesis implies

$$\mathbb{E}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} (\mu_{h,t}^*)^2 - \eta_t^2 \right) = 0.$$

It follows that $\lim_{n \to \infty} \mathbb{E}(n^{-1/2} \sum_{t=1}^{n} a_t) = 0$.

Under the conditions of Assumptions 1 and 2 it can be shown that $n^{-1} \sum_{t=j+1}^{n} a_t a_{t-j}$ converges in probability to its expectation for any finite j and $n \to \infty$. Therefore the test statistic for $\gamma_h = 0.5$ has a standard normal limiting distribution.

The asymptotic distribution for the test of the hypothesis $\gamma_h = 0$ follows in a similar manner. The test statistic is based on

$$
\begin{aligned}
a_t &= (Y_{t+h} - \overline{Y}_t)(\widehat{Y}_{t+h|t} - \overline{Y}_t) \\
&= (\mu_{h,t}^* + u_{h,t})(\mu_{h,t}^* + \eta_t)
\end{aligned}
$$

Since $\mathbb{E}(\mu_{h,t}^* \eta_t) = 0$ and $\mathbb{E}(\mu_{h,t}^* u_{h,t}) = 0$ we have

$$
\begin{aligned}
\mathbb{E}\left(\sum_{t=1}^{n} (\mu_{h,t}^*)^2 \right) &= \mathbb{E}\left(\sum_{t=1}^{n} (\mu_{h,t} - \mu)^2 \right) + \mathbb{E}\left(\sum_{t=1}^{n} \zeta_{T,t}^2 \right) \\
&= \mathbb{E}\left(\sum_{t=1}^{n} (\mu_{h,t} - \mu)^2 \right) + \sum_{t=1}^{n} \frac{(T+n)^2}{(T+t)^2} \left(\frac{S_t}{\sqrt{T+n}} \right)^2 \frac{1}{T+n} \\
\mathbb{E}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} (\mu_{h,t}^*)^2 \right) &= \mathbb{E}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} (\mu_{h,t} - \mu)^2 \right) + O\left(n^{-1/2} \right)
\end{aligned}
$$

If $\mu_{h,t} = \mu$ it follows that

$$\lim_{n \to \infty} \mathbb{E}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} a_t \right) = 0$$

and the sample covariances of a_t converge in probability to the respective covariances as $n \to \infty$ and the HAC t-statistic possesses a standard normal limiting distribution.

OA.3.2 Proof of Theorem 3:

Consider some $h > h^*$. We first analyze

$$\sum_{t=1}^{n} (\widehat{Y}_{t+h|t} - \overline{\widehat{Y}}_h)(Y_{t+h} - \overline{Y}^h) = \sum_{t=1}^{n} \widehat{Y}_{t+h|t}(u_{h,t} - \overline{u}_h).$$

An important problem with analysing this expression is that the estimation error in $\widehat{Y}_{t+h|t} = \mu_{h,t}(\widehat{\theta}_t)$ is correlated with $\overline{u}_h$. To sidestep this difficulty we decompose the forecast into one component, where the estimator is computed at $t = 0$ such that $\mu_{h,t}(\widehat{\theta}_0)$ is uncorrelated with $\{u_{1+h}, \ldots, u_{n+h}\}$ and show that the remaining component is asymptotically negligible. Applying a mean value expansion yields

$$\widehat{Y}_{t+h|t} = \mu_{h,t}(\widehat{\theta}_0) + D_{h,t}(\bar{\theta}_t)'(\widehat{\theta}_t - \widehat{\theta}_0)$$

where $D_{h,t}(\theta) = \partial \mu_{h,t}(\theta)/\partial\theta$ and $\bar{\theta}_t$ denotes some value between $\widehat{\theta}_0$ and $\widehat{\theta}_t$. Note that $\mu_{h,t}(\widehat{\theta}_0)$ and $u_{1+h}, u_{2+h}, \ldots, u_{n+h}$ are uncorrelated as the forecast errors are uncorrelated with the past. Next we analyze

$$\sum_{t=1}^{n} \left[\mu_{h,t}(\widehat{\theta}_0) + D_{h,t}(\bar{\theta}_t)'(\widehat{\theta}_t - \widehat{\theta}_0) \right] (u_{h,t} - \overline{u}_h) = A_{T,n} + B_{T,n}^1 + B_{T,n}^2$$

where

$$A_{T,n} = \sum_{t=1}^{n} \mu_{h,t}(\widehat{\theta}_0)(u_{h,t} - \overline{u}_h)$$

$$B_{T,n}^1 = \sum_{t=1}^{n} D_{h,t}(\bar{\theta}_t)'(\widehat{\theta}_t - \widehat{\theta}_0)u_{h,t}$$

$$B_{T,n}^2 = \overline{u}_h \sum_{t=1}^{n} D_{h,t}(\bar{\theta}_t)'(\widehat{\theta}_t - \widehat{\theta}_0).$$

Another mean value expansion around the true value θ with $\mu_{h,t}(\theta) = \mu$ (the null hypothesis) yields

$$A_{T,n} = (\widehat{\theta}_0 - \theta)' \sum_{t=1}^{n} D_{h,t}(\bar{\theta}_0)u_{h,t} - \overline{u}_h(\widehat{\theta}_0 - \theta)' \sum_{t=1}^{n} D_{h,t}(\bar{\theta}_0)$$

$$= A_{T,n}^1 + A_{T,n}^2$$

where $\bar{\theta}_0$ is again some value between $\widehat{\theta}_0$ and θ. Since $\widehat{\theta}_0$ and $D_{h,t}(\bar{\theta}_0)$ are uncorrelated with $u_{h,t}$ it follows that $A_{T,n}^1 = O_p(T^{-1/2})O_p(n^{1/2})$ and $A_{T,n}^2 = O_p(T^{-1/2})O_p(n^{-1/2})O_p(n)$. Thus, $A_{T,n}$ is $O_p(\sqrt{n/T})$. Under the null hypothesis $\widehat{\theta}_t - \widehat{\theta}_0$ and $D_{h,t}(\bar{\theta}_t)$ are uncorrelated with $u_{h,t}$.

Furthermore, Assumptions 3 (ii) and (iii) imply

$$\sum_{t=1}^{n} ||\widehat{\theta}_t - \widehat{\theta}_0||^2 ||D_{h,t}(\bar{\theta}_t)||^2 u_{h,t}^2 = O_p\left(\frac{n}{T^2}\right) \sum_{t=1}^{n} ||D_{h,t}(\bar{\theta}_t)||^2 u_{h,t}^2 = O_p\left(\frac{n^2}{T^2}\right).$$

and, therefore,

$$B_{T,n}^1 = \sum_{t=1}^{n} (\widehat{\theta}_t - \widehat{\theta}_0)' D_{h,t}(\bar{\theta}_t) u_{h,t} = O_p\left(\frac{n}{T}\right).$$

Since

$$\sum_{t=1}^{n} (\widehat{\theta}_t - \widehat{\theta}_0)' D_{h,t}(\bar{\theta}_t) = O_p\left(\frac{\sqrt{n}}{T}\right) \sum_{t=1}^{n} D_{h,t}(\bar{\theta}_t)$$

$$= O_p\left(\frac{n^{3/2}}{T}\right)$$

it follows that $B_{T,n}^2 = O_p(n^{-1/2}) O_p(n^{3/2}/T) = O_p(n/T)$. As $n/T \to 0$ It follows that

$$\sqrt{\frac{T}{n}} \sum_{t=1}^{n} (\widehat{Y}_{t+h|t} - \overline{Y}_h)(Y_{t+h} - \overline{Y}_h) = \sqrt{\frac{T}{n}} (A_{T,n} + B_{T,n}^1 + B_{T,n}^2)$$

$$= \sqrt{\frac{T}{n}} A_{T,n} + O_p\left(\sqrt{\frac{n}{T}}\right)$$

$$= \sqrt{T}(\widehat{\theta}_0 - \theta)' \frac{1}{\sqrt{n}} \sum_{t=1}^{n} D_{h,t}(\bar{\theta}_0)(u_{h,t} - \overline{u}_h) + O_p\left(\sqrt{\frac{n}{T}}\right).$$

Next we analyze

$$\sum_{t=1}^{n} (\widehat{Y}_{t+h|t} - \overline{\widehat{Y}}_h)^2 (Y_{t+h} - \overline{Y}_h)^2 = \sum_{t=1}^{n} (\widehat{Y}_{t+h|t} - \overline{\widehat{Y}}_h)^2 (u_{h,t} - \overline{u}_h)^2.$$

Using the above mean value expansions we obtain

$$\widehat{Y}_{t+h|t} = \mu_{h,t}(\widehat{\theta}_0) + D_{h,t}(\bar{\theta}_t)'(\widehat{\theta}_t - \widehat{\theta}_0)$$

$$= \mu + D_{h,t}(\bar{\theta}_0)'(\widehat{\theta}_0 - \theta) + D_{h,t}(\bar{\theta}_t)'(\widehat{\theta}_t - \widehat{\theta}_0)$$

$$\widehat{Y}_{t+h|t} - \overline{\widehat{Y}}_h = \widetilde{D}_{h,t}(\bar{\theta}_0)'(\widehat{\theta}_0 - \theta) + \widetilde{\Psi}_{h,t}(\widehat{\theta}_t, \widehat{\theta}_0)$$

13

where

$$\widetilde{D}_{h,t}(\bar{\theta}_0) = D_{h,t}(\bar{\theta}_0) - n^{-1}\sum_{s=1}^{n} D_{h,s}(\bar{\theta}_0)$$

$$\widetilde{\Psi}_{h,t}(\widehat{\theta}_t, \widehat{\theta}_0) = D_{h,t}(\bar{\theta}_t)'(\widehat{\theta}_t - \widehat{\theta}_0) - \frac{1}{n}\sum_{s=1}^{n} D_{h,s}(\bar{\theta}_s)'(\widehat{\theta}_s - \widehat{\theta}_0).$$

It follows that

$$\sum_{t=1}^{n}(\widehat{Y}_{t+h|t} - \overline{\widehat{Y}}_h)^2(Y_{t+h} - \overline{Y}_h)^2 = \sum_{t=1}^{n}(\widehat{\theta}_0 - \theta)'\widetilde{D}_{h,t}(\bar{\theta}_0)\widetilde{D}_{h,t}(\bar{\theta}_0)'(\widehat{\theta}_0 - \theta)(u_{h,t} - \bar{u}_h)^2$$
$$+ \sum_{t=1}^{n}\widetilde{\Psi}_{h,t}(\widehat{\theta}_t, \widehat{\theta}_0)^2(u_{h,t} - \bar{u}_h)^2$$
$$+ 2\sum_{t=1}^{n}(\widehat{\theta}_0 - \theta)'\widetilde{D}_{h,t}(\bar{\theta}_0)\widetilde{\Psi}_{h,t}(\widehat{\theta}_t, \widehat{\theta}_0)(u_{t+j} - \bar{u}_h)^2$$
$$= C_{T,n}^0 + C_{T,n}^1 + 2C_{T,n}^2 \ .$$

For the leading term we obtain

$$C_{T,n}^0 = O_p(T^{-1})O_p(n) \ = \ O_p(n/T)$$

For the second term $C_{T,n}^1$ we note that

$$\sum_{t=1}^{n}||\widehat{\theta}_t - \widehat{\theta}_0||^2||D_{h,t}(\bar{\theta}_t)||^2(u_{h,t} - \bar{u}_h)^2 = O_p\left(\frac{n}{T^2}\right)\sum_{t=1}^{n}||D_{h,t}(\bar{\theta}_t)||^2(u_{h,t} - \bar{u}_h)^2 \ = \ O_p(n^2/T^2).$$

Since the mean adjustment does not affect the order of magnitude we conclude that

$$C_{T,n}^1 = O_p(n^2/T^2).$$

For the last term $C_{T,n}^2$ we have

$$\sum_{t=1}^{n}(\widehat{\theta}_0 - \theta)'D_{h,t}(\bar{\theta}_0)D_{h,t}(\bar{\theta}_t)'(\widehat{\theta}_t - \widehat{\theta}_0) = O_p\left(\frac{1}{\sqrt{T}}\right)\left(\sum_{t=1}^{n}D_{h,t}(\bar{\theta}_0)D_{h,t}(\bar{\theta}_t)\right)O_p\left(\frac{\sqrt{n}}{T}\right) \ = \ O_p(n^{3/2}/T)$$

and, since the mean-adjustments do not affect the order of magnitude,

$$C_{T,n}^2 = O_p(n^{3/2}/T^{3/2}).$$

Combining these results yields

$$\frac{T}{n} \sum_{t=1}^{n} (\widehat{Y}_{t+h|t} - \overline{\widehat{Y}}_h)^2 (Y_{t+h} - \overline{Y}_h)^2$$

$$= \frac{1}{n} \sum_{t=1}^{n} \sqrt{T}(\widehat{\theta}_0 - \theta)' \widetilde{D}_{h,t}(\bar{\theta}_0) \widetilde{D}_{h,t}(\bar{\theta}_0)' \sqrt{T}(\widehat{\theta}_0 - \theta)(u_t - \bar{u}_h)^2 + O_p\left(\sqrt{\frac{n}{T}}\right).$$

In the same manner we can analyze the estimator for the long run variance. Let $\widehat{V}_{n,T} = \widehat{\Gamma}_{0,n,T} + 2 \sum_{j=1}^{h-1} \widehat{\Gamma}_{j,n,T}$ with

$$\widehat{\Gamma}_{j,n,T} = \frac{1}{n} \sum_{t=1+j}^{n} (\widehat{Y}_{t+h|t} - \overline{\widehat{Y}}_h)(\widehat{Y}_{t+h-j|t} - \overline{\widehat{Y}}_h)(Y_{t+h} - \overline{Y}_h)(Y_{t+h-j} - \overline{Y}_h)$$

$$\frac{T}{n}\widehat{\Gamma}_{j,n,T} = \frac{1}{n} \sum_{t=1}^{n} \sqrt{T}(\widehat{\theta}_0 - \theta)' \widetilde{D}_{h,t}(\bar{\theta}_0) \widetilde{D}_{t+h-j}(\bar{\theta}_0)' \sqrt{T}(\widehat{\theta}_0 - \theta)'(u_{h,t} - \bar{u}_h)(u_{t+h-j} - \bar{u}_h) + O_p\left(\sqrt{\frac{n}{T}}\right).$$

It follows that

$$\frac{T}{n}\widehat{V}_{n,T} = \mathbb{E}\left[\frac{1}{n}\left(\sum_{t=1}^{n} \sqrt{T}(\widehat{\theta}_0 - \theta)' \widetilde{D}_{t+h}(\theta)(u_{h,t} - \bar{u}_h)\right)^2\right] + o_p(1).$$

Finally we obtain

$$\frac{1}{\sqrt{nV_{n,T}}} \sum_{t=1}^{n} (\widehat{Y}_{t+h|t} - \overline{\widehat{Y}}_h)(u_{h,t} - \bar{u}_n) = \frac{\frac{1}{\sqrt{n}} \sum_{t=1}^{n} (\widehat{\theta}_0 - \theta)' \widetilde{D}_{h,t}(\theta)(u_{h,t} - \bar{u}_h)}{\sqrt{\mathbb{E}\left[\frac{1}{n}\left(\sum_{t=1}^{n} (\widehat{\theta}_0 - \theta)' \widetilde{D}_{h,t}(\theta)(u_{h,t} - \bar{u}_h)\right)^2\right]}} + o_p(1)$$

$$\xrightarrow{d} \mathcal{N}(0,1)$$

OA.4 Critical values of $dm_{T,h}$

Table OA.7: Critical values of $dm_{T,h}$ depending on π

π	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$
0.05	-8.84	-6.11	-5.01
0.10	-7.98	-5.33	-4.26
0.15	-7.42	-4.83	-3.80
0.20	-6.96	-4.48	-3.46
0.25	-6.60	-4.16	-3.18
0.30	-6.16	-3.87	-2.93
0.35	-5.84	-3.62	-2.72
0.40	-5.54	-3.40	-2.53
0.45	-5.22	-3.16	-2.33
0.50	-4.93	-2.96	-2.16
0.55	-4.60	-2.73	-1.99
0.60	-4.29	-2.52	-1.82
0.65	-3.97	-2.32	-1.65
0.70	-3.65	-2.11	-1.49
0.75	-3.27	-1.89	-1.32
0.80	-2.91	-1.65	-1.15
0.85	-2.46	-1.40	-0.96
0.90	-2.00	-1.12	-0.76
0.95	-1.38	-0.76	-0.51
0.96	-1.23	-0.68	-0.46
0.97	-1.06	-0.58	-0.39
0.98	-0.86	-0.47	-0.31
0.99	-0.61	-0.33	-0.22

Note: π denotes the ratio $T/(T+n)$ where n is the size of the evaluation sample and $T + n$ is the full sample size. H_0 is rejected at the significance level α if $dm_{T,h}$ is smaller than the critical value. Critical values are based on 1 million simulations.

References

ANDREWS, D. W. K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59(3), 817–858.

FAMA, E. F., AND K. R. FRENCH (1988): "Dividend yields and expected stock returns," *Journal of Financial Economics*, 22(1), 3 – 25.

GIACOMINI, R., AND H. WHITE (2006): "Tests of Conditional Predictive Ability," *Econometrica*, 74(6), 1545–1578.

NELSON, C. R., AND G. W. SCHWERT (1977): "Short-Term Interest Rates as Predictors of Inflation: On Testing the Hypothesis That the Real Rate of Interest is Constant," *American Economic Review*, 67(3), 478–486.

NEWEY, W. K., AND K. D. WEST (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3), 703–708.

PHILLIPS, P. C. B., AND V. SOLO (1992): "Asymptotics for Linear Processes," *The Annals of Statistics*, 20(2), 971–1001.

STOCK, J. H., AND M. W. WATSON (2007): "Why Has U.S. Inflation Become Harder to Forecast?," *Journal of Money, Credit and Banking*, 39(s1), 3–33.