Online Appendix to

Reassessing the Predictive Power of the Yield Spread for Recessions in the United States

Patrick J. Coe
Carleton University and CAMA

Shaun P. Vahey
University of Warwick and CAMA

October 9, 2024

A1 Data

In this section, we discuss our data sources.

A1.1 Recession Target Variable

In the main paper, we discuss results using the "first-final" release of real GDP (GNP) obtained from the monthly data vintages for $ROUTPUT$, contained in the *Real-Time Data Set for Macroeconomists*, https://www.philadelphiafed.org/. The first-final data for each observation (quarter) is taken from the vintage published in the 3^{rd} month of the following quarter. For example, the first-final release for the observation referring to 2022:1 is from the June 2022 vintage. We use first-final data to estimate the parameters of the probit models and to evaluate the forecasts for the results in both tables contained in the main paper.

In this appendix, we provide further tables based on two alternative measures of real economic activity constructed from $ROUTPUT$. Our first alternative uses the "advance" (or initial) release for each quarter. For each observation, this is taken from the vintage published in the 2^{nd} month of the following quarter. For example, for the observation referring to 2022:1, this initial measurement comes from the vintage dated May 2022. Our second alternative measure uses our "final" data vintage, dated April 22^{nd} 2024.

We follow Rudebusch and Williams (2009, RW) and define a recession as one quarter of negative economic growth, constructing $R1$ as:

$$R1_t = \begin{cases} 0 \text{ if } y_t \geq y_{t-1}. \\ 1 \text{ if } y_t < y_{t-1}. \end{cases}$$

where y_t is real GDP (GNP) in quarter t.

A1.2 Yield Spread and National Financial Conditions Index

To construct the yield spread, we again follow the approach used by RW. We subtract the 3-Month Treasury Bill Secondary Market Rate (FRED series TB3MS) from the Market Yield on U.S. Treasury Securities at 10-Year Constant Maturity (FRED series GS10). Both series were downloaded from https://fred.stlouisfed.org/ on April 22nd, 2024.

The National Financial Conditions Index (NFCI) is produced by FRB Chicago. In the broad replication section of the main paper, we use Amburgey and McCracken's (2023) real-time NFCI. They provide vintages of weekly NFCI data from January 6^{th} 1988 to July 28^{th} 2021. Within each vintage, we aggregate the NFCI to the quarterly frequency by averaging the weekly data.

For the results reported in the main paper, the "first-final" real output measurement for each observation comes from the vintage for the third month of the following quarter; see the discussion in Section 1.1 of this appendix. Therefore, we use the NFCI vintage from the 11^{th} week of each quarter to construct the appropriate predictor variable for financial conditions.

In this appendix, we also report additional results based on two further measures of financial conditions. The first uses the real-time vintages but matches the timing of the advance release $R1$ variable, using the NFCI vintage from the 7^{th} week of each quarter. The second uses the single vintage FRED series NFCI, downloaded from https://fred.stlouisfed.org/ on April 22nd, 2024 (ie not real-time).

A1.3 Survey of Professional Forecasters

Following RW, we use the average recession probability forecasts of the experts in the Survey of Professional Forecasters (SPF) as the benchmark. These forecasts are available from the Federal Reserve Bank of Philadelphia's website, https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/survey-of-professional-forecasters. See Federal Reserve Bank of Philadelphia (2024) for more information.

A2 Forecast Performance Metrics

Following RW, we use the end of evaluation sample Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Log Probability Score (LPS) to compare forecast performance. These are defined as:

$$
\begin{aligned}
MAE(h) &= \tfrac{1}{T^E}\textstyle\sum_{t=1}^{T^E}\left|F_{t+h|t} - R1_{t+h}\right| \\
RMSE(h) &= \sqrt{\tfrac{1}{T^E}\textstyle\sum_{t=1}^{T^E}\left(F_{t+h|t} - R1_{t+h}\right)^2} \\
LPS(h) &= -\tfrac{1}{T^E}\textstyle\sum_{t=1}^{T^E}\left[(1 - R1_{t+h})\ln\left(1 - F_{t+h|t}\right) + R1_{t+h}\ln\left(F_{t+h|t}\right)\right]
\end{aligned}
$$

where h is the forecast horizon, T^E is the length of the evaluation sample, $R1_{t+h}$ is the realisation for the target variable and $F_{t+h|t}$ is the forecast probability of a recession made at forecast origin t.

A3 Additional Results

In this section, we report some additional results. For the tables in this section of the appendix, we use various measures of real economic activity to define the $R1$ variable for probit parameter estimation and for forecast evaluation. We also use a final vintage version of the NFCI from April 22nd, 2024, a dummy variable for yield curve inversions and consider robust estimation methods.

A3.1 Narrow Replication

In figure A1, the counterpart to Figure 1 of the main paper, we define the $R1$ recessions using the advance release data, rather than the first-final data. Similarly, Table A1 is the advance release counterpart to Table 1 of the main paper. That is, we use advance release $R1$ to estimate the parameters of the probit models and to construct the realisations for forecast evaluation.

In Table A2, we display additional results reverting to first-final $R1$ for probit parameter estimation. However, in contrast to Table 1 in the main paper, we evaluate the forecasts using realisations constructed using the final vintage data for $ROUTPUT$ (April 22^{nd} 2024).

The results displayed in Tables A1 and A2 are consistent with the findings displayed in Table 1 in the main paper. The probit model with the yield spread predictor variable performs well for the longer horizons and the SPF provides a more competitive benchmark for the shorter horizons.

In Table A3, we provide additional results for Croushore and Marsten's (2016, CM) evaluation sample with forecast origins from 1968:4 to 2013:1. The rows labelled "CM SPF" and "CM Yield Spread" are from the "Full sample" columns of Table 2 of CM. The rows labelled "SPF" and "Yield Spread" refer to our replication. We use the first-final data to define $R1$, matching Table 1 in our main paper and Table 2 in CM. That is, we use first-final $R1$ to estimate the parameters of the probit models and to construct the realisations for forecast evaluation.

CM argue that the relative predictive content of the yield spread at longer horizons owes much to the earlier part of their evaluation samples. We explore this in Figure A2 which contains four panels and focuses on our longest horizon, $h = 4$. Panels (a) and (c) display the three forecast metrics we employ in the paper, but calculated over rolling m-quarter windows rather than the full evaluation sample, as reported in Table 1 of our paper. For panel (a), we set $m = 40$. And, for panel (c), we set $m = 60$. For each metric, the figure displays the ratio of the yield spread to the SPF. Therefore, values below one indicate the yield spread performs better than the benchmark.[1] We plot each metric against the midpoint of the rolling window.[2] Consistent with CM, we find that the relative performance of the yield spread is strongest before 2000.

In panels (b) and (d) of Figure A2 we display the results from the fluctuation test of Giacomini and Rossi (2010). Rossi and Sekhposyan (2010) use this test to study the ability of models to predict output growth and inflation over time. In our application, negative values of the test statistic indicate that the probit with the (lagged) yield spread outperforms the benchmark. The null hypothesis for the fluctuation test is that the two models have equal forecasting performance for all observations in the evaluation. The alternative hypothesis is that one of the models performs better for at least one observation. Under the null, the test statistic is always between the two

[1] Figure 2 in CM uses the inverse of this ratio. Hence, values above one in their figure indicate superior performance for the yield spread.

[2] Figure 2 in CM uses the end of the rolling window.

critical values indicated by the dashed lines in panels (b) and (d). We perform this test using rolling windows of lengths $m = 40$ and $m = 60$. The null is rejected in favour of the alternative in which the yield spread has superior performance in windows centered from around 1980 to the mid-1990s.

A3.2 Broad Replication

Table A4 is the counterpart to Table 2 in the main paper, where we use advance release data to define $R1$ for probit parameter estimation and for forecast evaluation. Whereas, in Table 2, we use first-final $R1$. The real-time NFCI data refer to the 7^{th} week of each quarter.

In Table A5, we provide results for our broad replication reverting to first-final data to estimate the probit parameters, matching Table 2. However, in this case, we evaluate the forecasts using $R1$ realisations constructed with the final vintage data (April 2024). The real-time NFCI data refer to the 11^{th} week of the quarter.

In Table A6, we display results for forecast origins from 1982:2 to 2023:4 and "final vintage" NFCI measurements from April 2024. For Table 2 in the main paper, we use forecast origins from 1988:1 to 2021:2 to match the availability of the real-time NFCI data. This single vintage NFCI extends back to 1971:1 and so using this series rather than real-time data allows us to begin our forecasting exercise in 1982:2 (ensuring a minimum of 40 observations for estimation after accounting for lags in the probit equation). Using this final vintage also allows us to extend the sample to 2023:4. As in Table 2, we use first-final data to define $R1$ for probit parameter estimation and forecast evaluation.

Our additional results for the broad replication provided in this appendix are broadly consistent with those provided in the main paper. The probit models with the NFCI predictor are competitive with the SPF benchmark for some horizons and some forecast performance metrics. But, generally speaking, the SPF is hard to beat at shorter horizons and the probit models perform better for longer horizons. One difference is that the yield spread now gives the lowest score for each metric at horizon $h = 3$ when we use the final vintage for the NFCI. Whereas, with the real-time NFCI (over a shorter sample), the NFCI typically gives the lowest score for each metric at this horizon, as reported in the main text.

Figure A3 is similar to A2, but focuses on the relative performance of the real-time NFCI over time, rather than the yield spread. It reports rolling forecast metrics and fluctuation tests for the probit model with the NFCI predictor relative to the SPF for $h = 4$. As for the yield spread, the NFCI is relatively effective prior to the Global Financial Crisis.

In Table A7, we display results for a probit model in which we use a (lagged) yield curve inversion predictor variable for forecast origins from 1974:1 to 2023:4. This binary variable takes on the value one when the yield spread is negative and the value zero otherwise. As in the main paper, we use first-final data to define $R1$ for probit parameter estimation and evaluation. We find that the yield curve inversion variable provides superior forecasts to the SPF at longer horizons, but not shorter horizons. However, a consistent pattern is that the yield curve inversion variable does not predict as well as the yield spread.

A3.3 Alternative Estimation Methods

Table A8 reports the first four moments of the yield spread and the NFCI along with $p-$values from a Shapiro-Wilk test for normality. Figure A4 reports plots of the empirical PDFs for the yield spread and the NFCI for 1971:1 to 2023:4. In each case, these are overlaid with the PDF from a normal distribution with the same mean and standard deviation. Estimates of skewness and kurtosis suggest departures from normality, especially for the NFCI. These are confirmed by the Shapiro-Wilk tests.

Given the non-Gaussian nature of the NFCI, Cheris, Coe and Vahey (2023) use copula-based methods to estimate the dependence between the NFCI and real activity. In this appendix, we use related methods to estimate the probit models and to forecast recessions. Let x_t be the data for the yield spread (NFCI). We construct the (ascending) ranked data as $x_t^r = R_{x,t}/(T+1)$, where $R_{x,t}$ denotes the rank of each observation for x_t relative to its own history and T is the length of the time series. The ranked data variable is then transformed to be Gaussian using an inverse Gaussian Cumulative Distribution Function (CDF). The approach generates standard Normal "pseudo" data, x_t^p for the yield spread (NFCI). We then estimate the probit model based on the pseudo variable and generate forecasts of the recession target variable.

In Table A9, we provide results using robust regression based on ranks for our broad replication. This table is analogous to Table 2 of our paper, with the difference being that the probit is estimated using pseudo data for the predictor variable and the out-of-sample forecast is generated using the observation on x_{t-1}^p. Overall, the results are similar to Table 2 of the main paper. The SPF continues to dominate the yield spread at shorter horizons, but the yield spread performs better at longer horizons. Using pseudo data, does give a slight improvement in the performance of the NFCI at longer horizons.

In Table A10, we display results using a logit model rather than a probit model. These results are almost identical to those we report in Table 2 of the paper for the probit model.

References

Amburgey, Aaron J. and Michael W. McCracken (2023) “On the Real-Time Predictive Content of Financial Condition Indices for Growth” *Journal of Applied Econometrics* 38(2) pp137-163.

Chernis, Tony, Patrick J. Coe and Shaun P. Vahey (2023) “Reassessing the Dependence Between Economic Growth and Financial Conditions Since 1973” *Journal of Applied Econometrics* 38(2) pp260-267.

Dean Croushore and Katherine Marsten (2016) “Reassessing the Relative Power of the Yield Spread in Forecasting Recessions” *Journal of Applied Econometrics* 31(6) pp1183-1191.

Federal Reserve Bank of Philadelphia (2024) *Survey of Professional Forecasters: Documentation* Available from philadelphiafed.org

Federal Reserve Bank of Philadelphia (2023) *Real-Time Dataset for Macroeconomists.* Available from philadelphiafed.org

Giacomini, Raffaella and Barbara Rossi (2010) “Forecast Comparisons in Unstable Environments” *Journal of Applied Econometrics* 25(4) pp595-620.

Rossi, Barbara and Tatevik Sekhposyan (2010) “Have Economic Models’ Forecasting Performance for US Output Growth and Inflation Changed Over Time, and When?” *International Journal of Forecasting* 26(4) pp808–835.

Rudebusch, Glenn D. and John C. Williams (2009) “Forecasting Recessions: The Puzzle of the Enduring Power of the Yield Curve” *Journal of Business and Economic Statistics* 27(4) pp492-503.

Table A1: Evaluation of Real-time Recession Probability Forecasts, Advance Data

	Forecast Origins for Evaluation Sample					
	1968:4 - 2007:1			1968:4 - 2023:4		
	MAE	RMSE	LPS	MAE	RMSE	LPS
$h = 0$						
RW SPF	**0.166**	**0.257**	**0.236**			
RW Yield Spread	0.205	0.313	0.350			
SPF	**0.166**	**0.257**	**0.236**	**0.174**	**0.263**	**0.248**
Yield Spread	0.206	0.313†	0.345†	0.214†	0.327†	0.378†
$h = 1$						
RW SPF	0.216	0.298	0.299			
RW Yield Spread	**0.182**	**0.289**	**0.281**			
SPF	0.216	0.298	0.299	0.223	**0.305**	**0.315**
Yield Spread	**0.180**	**0.285**	**0.278**	**0.195**	0.311	0.340
$h = 2$						
RW SPF	0.239	0.321	0.346			
RW Yield Spread	**0.190**†	**0.294**	**0.320**			
SPF	0.239	0.321	0.346	0.244	0.325	0.355
Yield Spread	**0.188**†	**0.292**	**0.316**	**0.198**†	**0.312**	**0.350**
$h = 3$						
RW SPF	0.254	0.339	0.387			
RW Yield Spread	**0.203**†	**0.310**	**0.328**†			
SPF	0.254	0.339	0.387	0.257	0.342	0.391
Yield Spread	**0.201**†	**0.309**	**0.328**	**0.204**†	**0.321**	**0.350**
$h = 4$						
RW SPF	0.265	0.351	0.418			
RW Yield Spread	**0.209**†	**0.317**†	**0.330**†			
SPF	0.265	0.351	0.418	0.264	0.349	0.412
Yield Spread	**0.214**†	**0.318**†	**0.333**†	**0.211**†	**0.323**†	**0.342**†

Notes: Yield Spread refers to the forecast from a probit model using the lagged yield spread as the single predictor variable. MAE, RMSE, and LPS refer to mean absolute error, root mean squared error and (-1 times) the log probability score, respectively. The lowest value is indicated in bold for each measure of forecast accuracy at each horizon (h). As a rough guide to statistical significance, differences from the SPF forecast at the 5% level, according to a two-tailed Diebold–Mariano test with HAC standard errors, are labelled †. Rows labelled RW contain the values reported by Rudebusch and Williams (2009).

Table A2: Real-time Recession Probability Forecasts, Evaluated Using Final Vintage Data

	Forecast Origins for Evaluation Sample					
	1968:4 - 2007:1			1968:4 - 2023:4		
	MAE	RMSE	LPS	MAE	RMSE	LPS
$h = 0$						
SPF	**0.166**	**0.256**	**0.234**	**0.181**	**0.276**	**0.267**
Yield Spread	0.201	0.306	0.313	0.222†	0.340†	0.393†
$h = 1$						
SPF	0.213	0.293	0.297	0.231	**0.318**	**0.341**
Yield Spread	**0.173**†	**0.273**	**0.251**	**0.202**†	0.322	0.365
$h = 2$						
SPF	0.238	0.319	0.347	0.252	0.339	0.381
Yield Spread	**0.185**†	**0.287**†	**0.285**†	**0.206**†	**0.325**	**0.366**
$h = 3$						
SPF	0.247	0.329	0.373	0.262	0.349	0.407
Yield Spread	**0.188**†	**0.288**†	**0.273**†	**0.205**†	**0.323**†	**0.352**
$h = 4$						
SPF	0.257	0.340	0.398	0.268	0.356	0.423
Yield Spread	**0.203**†	**0.300**†	**0.303**†	**0.214**†	**0.327**†	**0.360**†

Notes: Yield Spread refers to the forecast from a probit model using the lagged yield spread as the single predictor variable. MAE, RMSE, and LPS refer to mean absolute error, root mean squared error and (-1 times) the log probability score, respectively. The lowest value is indicated in bold for each measure of forecast accuracy at each horizon (h). As a rough guide to statistical significance, differences from the SPF forecast at the 5% level, according to a two-tailed Diebold–Mariano test with HAC standard errors, are labelled †.

Table A3: Evaluation of Real-time Recession Probability Forecasts Forecast Origins 1968:4 - 2013:1

	MAE	RMSE	LPS
$h = 0$			
CM SPF	**0.165**	**0.247**	**0.218**
CM Yield Spread	0.211†	0.322†	0.364†
SPF	**0.170**	**0.257**	**0.237**
Yield Spread	0.211	0.327†	0.379†
$h = 1$			
CM SPF	0.215	**0.292**	**0.296**
CM Yield Spread	**0.186**	0.309	0.328
SPF	0.214	**0.289**	**0.289**
Yield Spread	**0.193**	0.314	0.345
$h = 2$			
CM SPF	0.252	0.337	0.381
CM Yield Spread	**0.190**†	**0.312**	**0.325**
SPF	0.237	**0.314**	**0.335**
Yield Spread	**0.199**†	0.316	0.359
$h = 3$			
CM SPF	0.261	0.349	0.414
CM Yield Spread	**0.202**†	**0.314**†	**0.323**†
SPF	0.251	0.334	0.377
Yield Spread	**0.202**†	**0.319**	**0.342**
$h = 4$			
CM SPF	0.270	0.357	0.436
CM Yield Spread	**0.203**†	**0.312**†	**0.314**†
SPF	0.260	0.345	0.403
Yield Spread	**0.211**†	**0.320**†	**0.333**†

Notes: Yield Spread refers to the forecast from a probit model using the lagged yield spread as the single predictor variable. MAE, RMSE, and LPS refer to mean absolute error, root mean squared error and (-1 times) the log probability score, respectively. The lowest value is indicated in bold for each measure of forecast accuracy at each horizon (h). As a rough guide to statistical significance, differences from the SPF forecast at the 5% level, according to a two-tailed Diebold–Mariano test with HAC standard errors, are labelled †. Rows labelled CM contain the values reported by Croushore and Marsten (2016)

Table A4: Evaluation of Real-time Recession Probability Forecasts, Advance Data

	Forecast Origins for Evaluation Sample					
	1988:1 - 2007:1			1988:1 - 2021:2		
	MAE	RMSE	LPS	MAE	RMSE	LPS
$h = 0$						
SPF	0.131	0.200	0.165	0.143	**0.215**	**0.189**
Yield Spread	0.163	0.233	0.222	0.184	0.283	0.304
Real-time NFCI	**0.086**†	**0.189**	**0.145**	**0.108**	0.233	0.207
$h = 1$						
SPF	0.156	0.203	0.187	0.179	**0.240**	0.231
Yield Spread	0.149	0.224	0.196	0.175	0.282	0.308
Real-time NFCI	**0.104**†	**0.193**	**0.158**	**0.129**†	0.248	**0.228**
$h = 2$						
SPF	0.174	0.217	0.212	0.197	**0.256**	0.258
Yield Spread	0.148	0.208	0.185	0.172	0.269	0.277
Real-time NFCI	**0.124**†	**0.198**	**0.174**	**0.149**†	0.258	**0.247**
$h = 3$						
SPF	0.189	0.233	0.238	0.212	0.274	0.288
Yield Spread	0.133†	0.199†	0.166†	0.159†	0.265	0.260
Real-time NFCI	**0.115**†	**0.193**†	**0.163**†	**0.146**†	**0.259**	**0.249**
$h = 4$						
SPF	0.197	0.242	0.253	0.218	0.281	0.300
Yield Spread	0.136†	0.204†	**0.172**†	**0.158**†	**0.263**	**0.249**†
Real-time NFCI	**0.134**†	**0.200**†	0.180†	0.162†	0.267	0.271

Notes: Yield Spread and Real-time NFCI refer to the forecasts from probit models, each using a single predictor variable. MAE, RMSE, and LPS refer to mean absolute error, root mean squared error and (-1 times) the log probability score, respectively. For each forecast performance metric, the lowest value is indicated in bold. As a rough guide to statistical significance, differences from the SPF forecast at the 5% level, according to a two-tailed Diebold–Mariano test with HAC standard errors, are labelled †.

Table A5: Real-time Recession Probability Forecasts, Evaluated with Final Vintage Data

	Forecast Origins for Evaluation Sample					
	1988:1 - 2007:1			1988:1 - 2021:2		
	MAE	RMSE	LPS	MAE	RMSE	LPS
$h = 0$						
SPF	0.135	**0.208**	**0.172**	0.157	**0.246**	**0.227**
Yield Spread	0.168	0.245	0.233	0.207†	0.320	0.370†
Real-time NFCI	**0.096**	0.214	0.172	**0.135**	0.285	0.296
$h = 1$						
SPF	0.167	0.227	0.215	0.201	**0.283**	**0.292**
Yield Spread	0.154	0.234	0.206	0.197	0.318	0.383
Real-time NFCI	**0.114**†	**0.218**	**0.186**	**0.155**†	0.297	0.310
$h = 2$						
SPF	0.184	0.238	0.237	0.219	**0.296**	**0.315**
Yield Spread	0.156	0.226	0.203	0.194	0.308	0.348
Real-time NFCI	**0.135**†	**0.222**	**0.200**	**0.175**†	0.304	0.325
$h = 3$						
SPF	0.198	0.252	0.261	0.233	0.311	0.343
Yield Spread	0.143†	0.223†	**0.191**†	0.181†	**0.304**	0.346
Real-time NFCI	**0.127**†	**0.220**†	0.194†	**0.172**†	0.306	**0.327**
$h = 4$						
SPF	0.206	0.260	0.274	0.240	0.317	0.354
Yield Spread	0.146†	0.228†	**0.199**†	**0.180**†	**0.303**	0.339
Real-time NFCI	**0.144**†	**0.225**†	0.208†	0.186†	0.308	**0.336**

Notes: Yield Spread and Real-time NFCI refer to the forecasts from probit models, each using a single predictor variable. MAE, RMSE, and LPS refer to mean absolute error, root mean squared error and (-1 times) the log probability score, respectively. For each forecast performance metric, the lowest value is indicated in bold. As a rough guide to statistical significance, differences from the SPF forecast at the 5% level, according to a two-tailed Diebold–Mariano test with HAC standard errors, are labelled †.

Table A6: Evaluation of Real-time Recession Probability Forecasts, Final Vintage National Financial Conditions Index

	Forecast Origins for Evaluation Sample					
	1982:2 - 2007:1			1982:2 - 2023:4		
	MAE	RMSE	LPS	MAE	RMSE	LPS
$h = 0$						
SPF	0.133	**0.198**	0.165	0.161	**0.244**	**0.225**
Yield Spread	0.164	0.250	0.262	0.198	0.307	0.356†
NFCI	**0.089**†	0.204	**0.162**	**0.121**†	0.265	0.263
$h = 1$						
SPF	0.159	**0.203**	0.189	0.197	**0.268**	**0.271**
Yield Spread	0.144	0.239	0.217	0.187	0.307	0.351
NFCI	**0.113**†	0.217	**0.184**	**0.144**†	0.279	0.284
$h = 2$						
SPF	0.178	**0.221**	0.220	0.214	**0.282**	**0.297**
Yield Spread	**0.152**	0.225	**0.203**	0.188	0.294	0.312
NFCI	0.153†	0.227	0.215	**0.178**†	0.290	0.307
$h = 3$						
SPF	0.193	0.233	0.241	0.227	0.293	0.319
Yield Spread	**0.134**†	**0.201**†	**0.168**†	**0.173**†	**0.280**	**0.284**
NFCI	0.149†	0.218	0.207	0.176†	0.287	0.305
$h = 4$						
SPF	0.209	0.252	0.267	0.236	0.302	0.335
Yield Spread	**0.140**†	**0.204**†	**0.175**†	**0.173**†	**0.275**†	**0.269**†
NFCI	0.169†	0.225†	0.223†	0.191†	0.291	0.315

Notes: Yield Spread and Real-time NFCI refer to the forecasts from probit models, each using a single predictor variable. MAE, RMSE, and LPS refer to mean absolute error, root mean squared error and (-1 times) the log probability score, respectively. For each forecast performance metric, the lowest value is indicated in bold. As a rough guide to statistical significance, differences from the SPF forecast at the 5% level, according to a two-tailed Diebold–Mariano test with HAC standard errors, are labelled †.

Table A7: Evaluation of Real-time Recession Probability Forecasts, Yield Spread and Yield Curve Inversions

	Forecast Origins for Evaluation Sample					
	1975:1 - 2007:1			1975:1 - 2023:4		
	MAE	RMSE	LPS	MAE	RMSE	LPS
$h = 0$						
SPF	**0.166**	**0.258**	**0.239**	**0.179**	**0.273**	**0.265**
Yield Spread	0.182	0.293	0.317	0.204	0.324	0.380
Yield Curve Inversion	0.211†	0.298	0.318	0.225†	0.327†	0.368†
$h = 1$						
SPF	0.201	0.273	0.269	0.219	**0.299**	**0.311**
Yield Spread	**0.159**†	**0.273**	**0.258**	**0.188**	0.315	0.356
Yield Curve Inversion	0.203	0.296	0.309	0.218	0.326	0.362
$h = 2$						
SPF	0.213	**0.278**	**0.287**	0.232	**0.306**	**0.330**
Yield Spread	**0.173**†	0.286	0.314	**0.194**†	0.318	0.369
Yield Curve Inversion	0.209	0.281	0.296	0.222	0.314	0.348
$h = 3$						
SPF	0.230	0.296	0.323	0.245	0.321	0.361
Yield Spread	**0.169**†	**0.276**	**0.279**	**0.188**†	**0.310**	**0.338**
Yield Curve Inversion	0.204†	0.279	0.289	0.220†	0.320	0.358
$h = 4$						
SPF	0.241	0.307	0.345	0.252	0.327	0.375
Yield Spread	**0.174**†	**0.268**†	**0.257**†	**0.189**†	**0.300**†	**0.308**†
Yield Curve Inversion	0.211†	0.282†	0.297†	0.223†	0.317	0.354

Notes: Yield Spread and Yield Curve Inversion refer to the forecasts from probit models, each using a single predictor variable. MAE, RMSE, and LPS refer to mean absolute error, root mean squared error and (-1 times) the log probability score, respectively. For each forecast performance metric, the lowest value is indicated in bold. As a rough guide to statistical significance, differences from the SPF forecast at the 5% level, according to a two-tailed Diebold–Mariano test with HAC standard errors, are labelled †.

Table A8: Summary Statistics and Shapiro-Wilk Tests

	1953:2 - 2023:4	1971:1 - 2023:4	
	Yield Spread	Yield Spread	NFCI
Mean	1.404	1.599	0.003
Standard Deviation	1.161	1.247	0.983
Skewness	-0.093	-0.455	1.936
Kurtosis	2.478	2.511	6.635
Shapiro-Wilk p-value	0.006	0.0001	< 0.0001

Yield Spread refers to the Market Yield on 10-Year Treasury Securities minus the 3-Month Treasury Bill Secondary Market Rate. NFCI refers to the final vintage release (as of April 22nd, 2024) of the Chicago Fed's National Financial Conditions Index.

Table A9: Real-time Recession Probability Forecasts, Probit Estimation With Pseudo Data

	Forecast Origins for Evaluation Sample					
	1988:1 - 2007:1			1988:1 - 2021:2		
	MAE	RMSE	LPS	MAE	RMSE	LPS
$h = 0$						
SPF	0.129	**0.193**	**0.160**	0.148	**0.227**	**0.205**
Yield Spread	0.172	0.252	0.242	0.196	0.303	0.336
Real-time NFCI	**0.103**	0.215	0.169	**0.126**	0.268	0.355
$h = 1$						
SPF	0.154	**0.196**	0.181	0.184	**0.251**	**0.248**
Yield Spread	0.159	0.240	0.214	0.189	0.302	0.345
Real-time NFCI	**0.121**	0.213	**0.175**	**0.145**†	0.276	0.336
$h = 2$						
SPF	0.173	**0.214**	0.210	0.203	**0.267**	**0.274**
Yield Spread	0.161	0.231	0.211	0.187	0.292	0.311
Real-time NFCI	**0.135**†	0.217	**0.188**	**0.159**†	0.282	0.322
$h = 3$						
SPF	0.194	0.244	0.250	0.221	0.290	0.311
Yield Spread	0.149†	0.225†	**0.195**†	0.178†	0.290	**0.303**
Real-time NFCI	**0.127**†	**0.215**†	0.180†	**0.156**†	**0.284**	0.310
$h = 4$						
SPF	0.205	0.259	0.272	0.229	0.300	0.328
Yield Spread	0.150†	0.226†	0.198†	0.175†	**0.285**	**0.285**†
Real-time NFCI	**0.138**†	**0.217**†	**0.190**†	**0.166**†	0.287	0.309

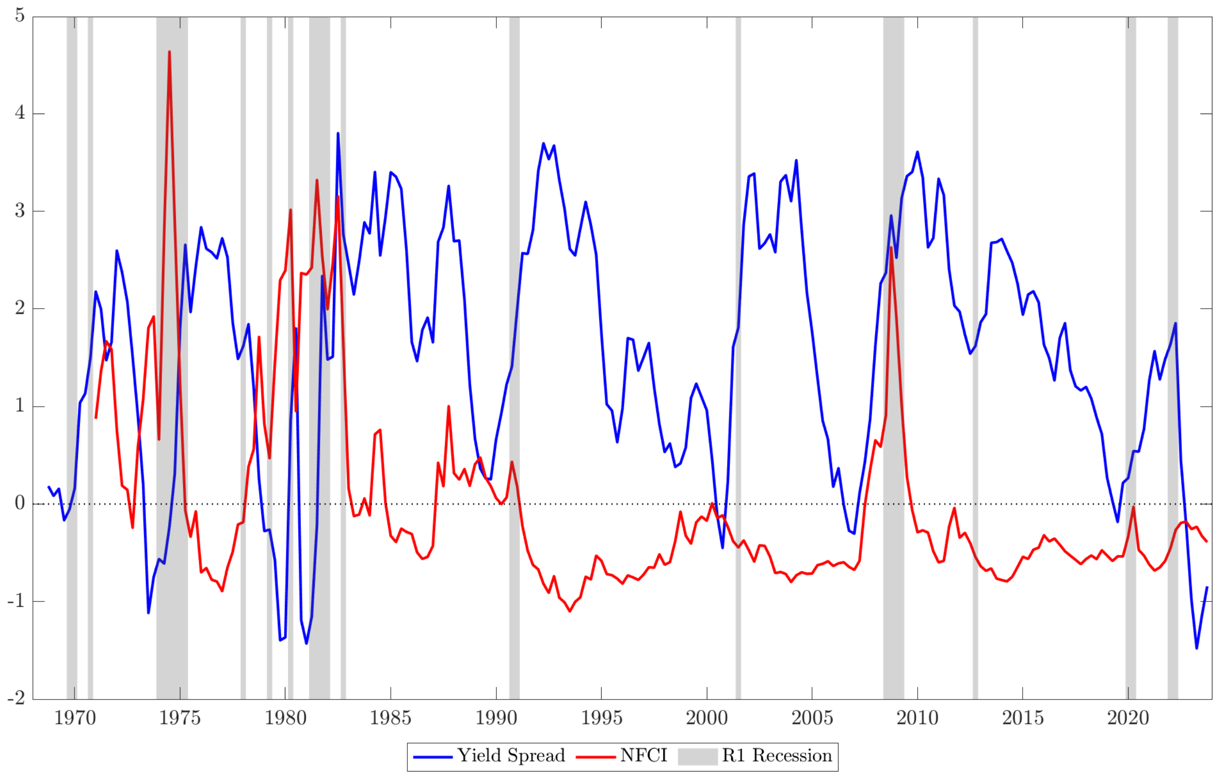
Notes: Yield Spread and Real-time NFCI refer to the forecasts from probit models, each using pseudo data for a single predictor variable. MAE, RMSE, and LPS refer to mean absolute error, root mean squared error and (-1 times) the log probability score, respectively. For each forecast performance metric, the lowest value is indicated in bold. As a rough guide to statistical significance, differences from the SPF forecast at the 5% level, according to a two-tailed Diebold–Mariano test with HAC standard errors, are labelled †.

Table A10: Real-time Recession Probability Forecasts, Logit Estimation

	Forecast Origins for Evaluation Sample					
	1988:1 - 2007:1			1988:1 - 2021:2		
	MAE	RMSE	LPS	MAE	RMSE	LPS
$h = 0$						
SPF	0.129	**0.193**	**0.160**	0.148	**0.227**	**0.205**
Yield Spread	0.172	0.252	0.248	0.196†	0.303	0.339†
Real-time NFCI	**0.095**	0.216	0.180	**0.122**	0.264	0.263
$h = 1$						
SPF	0.154	**0.196**	**0.181**	0.184	**0.251**	**0.248**
Yield Spread	0.154	0.241	0.216	0.184	0.302	0.345
Real-time NFCI	**0.114**†	0.217	0.185	**0.144**†	0.275	0.274
$h = 2$						
SPF	0.173	**0.214**	0.210	0.203	**0.267**	**0.274**
Yield Spread	0.152	0.224	**0.200**	0.180	0.289	0.309
Real-time NFCI	**0.137**†	0.222	0.201	**0.165**†	0.283	0.291
$h = 3$						
SPF	0.194	0.244	0.250	0.221	0.290	0.311
Yield Spread	0.137†	**0.216**†	**0.180**†	0.167†	**0.284**	0.293
Real-time NFCI	**0.128**†	0.219	0.193†	**0.162**†	0.285	**0.292**
$h = 4$						
SPF	0.205	0.259	0.272	0.229	0.300	0.328
Yield Spread	**0.139**†	**0.218**†	**0.184**†	**0.166**†	**0.282**	**0.279**†
Real-time NFCI	0.147†	0.223†	0.207†	0.177†	0.288	0.304

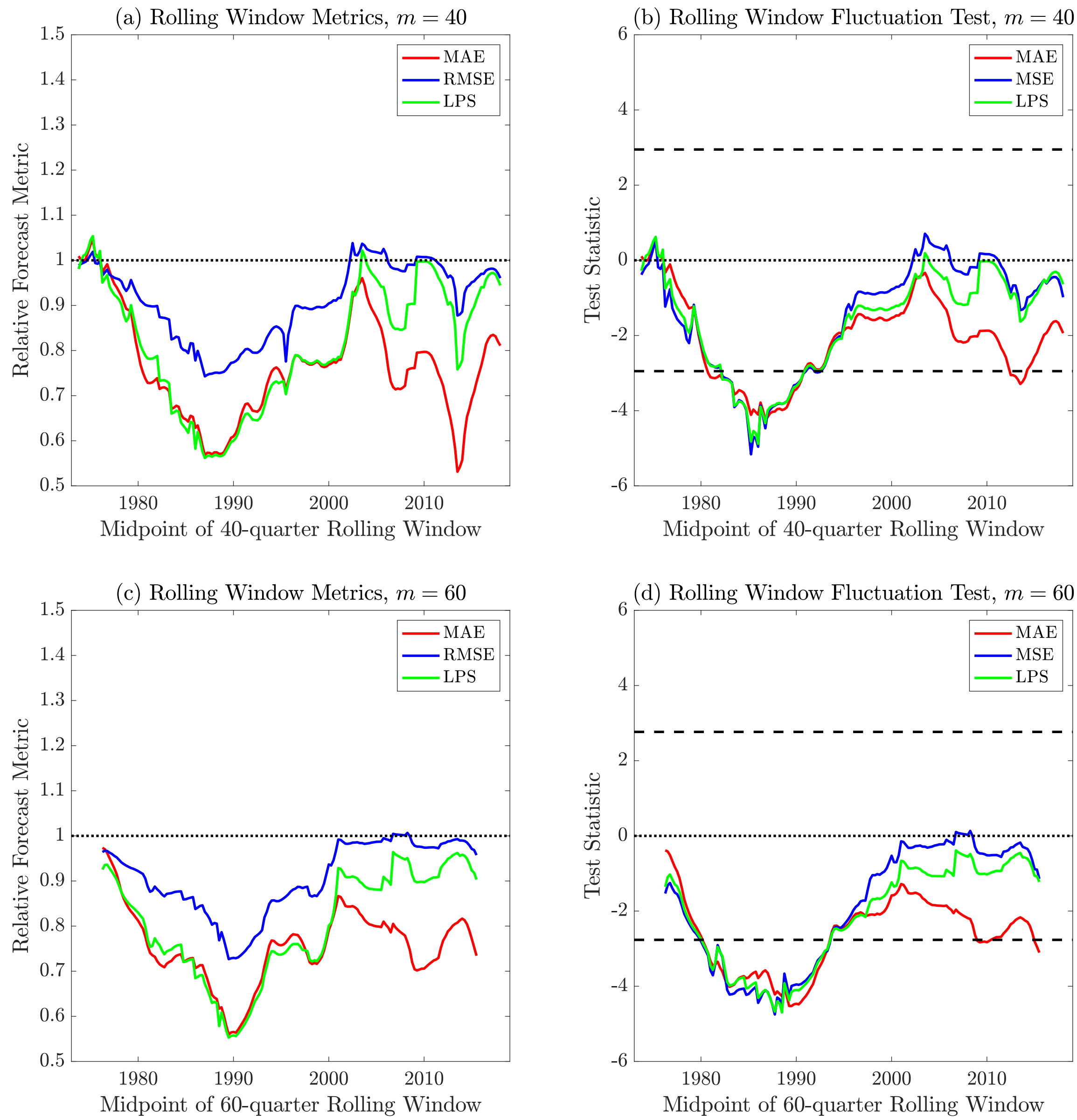
Notes: Yield Spread and Real-time NFCI refer to the forecasts from logit models, each using a single predictor variable. MAE, RMSE, and LPS refer to mean absolute error, root mean squared error and (-1 times) the log probability score, respectively. For each forecast performance metric, the lowest value is indicated in bold. As a rough guide to statistical significance, differences from the SPF forecast at the 5% level, according to a two-tailed Diebold–Mariano test with HAC standard errors, are labelled †.

Figure A1: $R1$ Recessions, the Yield Spread and the NFCI, 1968:4-2023:4



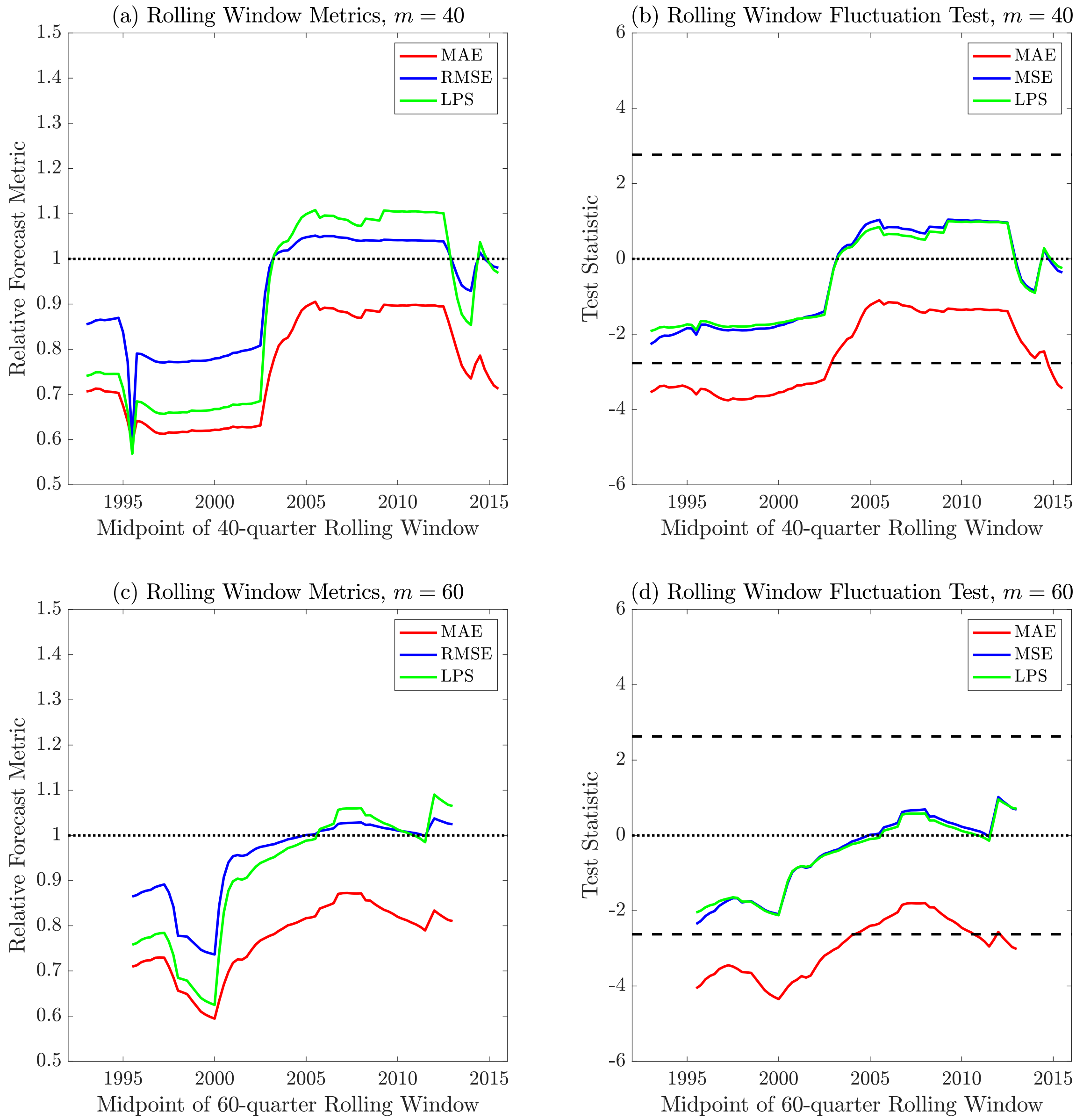
Notes: The $R1$ one-quarter recession timing is plotted using the advance releases of real-time data. The Yield Spread is the difference between the yields on 10-year Treasury bonds and 3-month Treasury bills. The NFCI is the National Financial Conditions Index produced by the Federal Reserve Bank of Chicago.

Figure A2: Rolling Forecast Metrics and Fluctuation Tests, $h = 4$



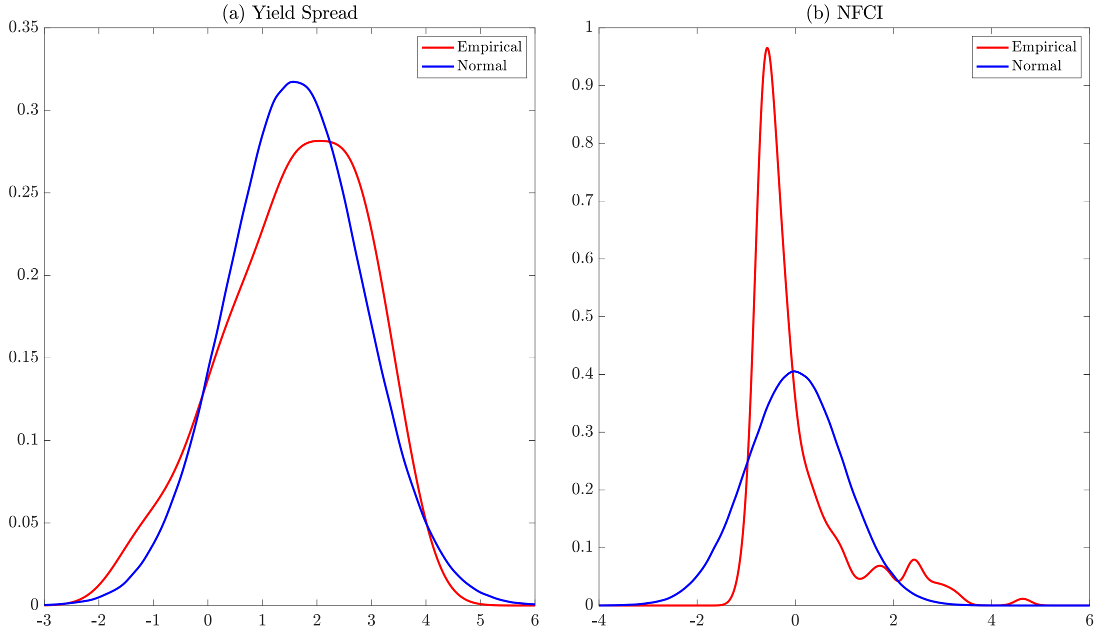
Notes: Relative forecast metrics are the metric for the yield spread probit forecast divided by the SPF forecast calculated over rolling $m-$quarter windows. A value less than one indicates the yield spread probit performs better over the window in question. Under the null hypothesis of the fluctuation test, the two models have equal forecasting performance at each point in time and the test statistic is always between the two critical values indicated by the dashed lines in panels (b) and (d). Under the alternative hypothesis, one of the models performs better for at least part of the evaluation sample. A test statistic below (above) the lower (upper) critical value indicates it is the yield spread probit (SPF) which performs better.

Figure A3: Rolling Forecast Metrics and Fluctuation Tests, $h = 4$



Notes: Relative forecast metrics are the metric for the NFCI probit forecast divided by the SPF forecast calculated over rolling $m-$quarter windows. A value less than one indicates the NFCI probit performs better over the window in question. Under the null hypothesis of the fluctuation test, the two models have equal forecasting performance at each point in time and the test statistic is always between the two critical values indicated by the dashed lines in panels (b) and (d). Under the alternative hypothesis, one of the models performs better for at least part of the evaluation sample. A test statistic below (above) the lower (upper) critical value indicates it is the NFCI probit (SPF) which performs better.

Figure A4: PDFs, 1971:1-2023:4



Notes: Red lines denote empirical Probability Density Functions (PDF). The blue lines show the PDF from a normal distribution with the same mean and standard deviation as the corresponding empirical PDF.