Forecasting the 10-Year Treasury Yield

A Supply-Side-Informed Error Correction Approach When the Phillips Curve Is Flat

Frank Schmid

frank.schmid@frankschmid.com

This version: 8 November 2026
First version: 19 March 2026

Abstract. This paper develops a forecasting model for the 10-year U.S. Treasury yield and evaluates it on data as they were published at the time. The model is an error correction equation. It ties the yield to its long-run relation with the federal funds rate and adds the latest changes in the yield, the funds rate, the yield spread, the unemployment rate, and a supply-side inflation signal, the producer price index for finished goods, which takes the place of consumer inflation on the proposition that, when the Phillips curve is flat, cost pressure upstream of the consumer carries the information about inflation that bond markets price. The target is the monthly average of the yield. A monthly average of a daily series is predictable from its own last change by construction, so every forecast is read against two benchmarks, the no-change forecast of the average and the yield at the month-end close. Extended with the month-end close, the model forecasts the next month's average more accurately than the no-change forecast; it does not improve on the close itself at any horizon from one month through twelve, and beyond one month neither form of the model improves on no change. The yield's own dynamics carry the model's fit in sample. No other coefficient is individually distinguishable from zero: not that on the change in the producer price inflation rate, in the full sample or in any segment set by the adoption of an inflation target, nor that on the change in the consumer price inflation rate in its place, the funds rate, the spread, or the unemployment rate, nor, read against its bootstrap distribution, that on the deviation from the long-run relation. Tests at overlapping horizons, and the test of the error correction term, are read against bootstrap distributions. On data from the Federal Open Market Committee's internal agreement on an inflation objective, the model's terms beyond the close jointly carry information beyond the regression on the close alone at horizons of one, three, and six months, though the equation with the close remains less accurate than the close; on data from January 2012 the test does not reject at any horizon.

JEL classification: C53, E43, E47, G12.

This paper was developed with assistance from generative AI tools across the design, analysis, coding, and drafting stages. The author supervised all stages, verified the results, and takes full responsibility for the content.

Contents

Abbreviations and Acronyms **iii**

1 Introduction **1**

2 Literature Review **2**

2.1 The Random Walk Benchmark 2
2.2 Factor Models of the Yield Curve 2
2.3 Macro-Finance and Survey-Based Approaches 3
2.4 Machine Learning Approaches 3
2.5 Time-Averaged Targets and Real-Time Data 3
2.6 Gap in the Literature 4

3 Data and Real-Time Methodology **4**

3.1 Target and Forecast Origin 4
3.2 What Is Known at the Origin 5

4 Unit Root and Random Walk Analysis **6**

4.1 Persistence 6
4.2 Time Averaging and the Two Benchmarks 7

5 Cointegration and Error Correction **9**

6 Supply-Side versus Demand-Side Inflation Signals **9**

7 Model Specification and Estimation **11**

7.1 Specification 11
7.2 Tests and Reference Distributions 11
7.3 Estimates 12
7.4 Diagnostics 14

8 Out-of-Sample Performance and Comparison with the Literature **15**

8.1 Scheme 16
8.2 Forecast Accuracy 16
8.3 Comparison with Published Benchmarks 19
8.4 The PPI Term 20
8.5 Robustness 20
8.6 Caveats and Limitations 24

9 The Inflation-Targeting Regimes **24**

10 Conditional Forecast **26**

11 Yield Decomposition **28**

12 Connection to the Taylor Rule and Observational Equivalence **29**

13 Conclusion **31**

Bibliography **32**

Data Sources **34**

Abbreviations and Acronyms

AI	Artificial intelligence
ARCH	Autoregressive conditional heteroskedasticity
BIC	Bayesian information criterion
FOMC	Federal Open Market Committee
FRED	Federal Reserve Economic Data, Federal Reserve Bank of St. Louis
GSCPI	Global Supply Chain Pressure Index, Federal Reserve Bank of New York
JEL	Journal of Economic Literature
LM	Lagrange multiplier
PCE	Personal consumption expenditures
PPI	Producer price index
RMSE	Root mean squared error

Greek Letters

α (alpha)	Intercept of the long-run relation between the yield and the federal funds rate
β (beta)	Slope of the long-run relation
γ (gamma)	Coefficient on the deviation from the long-run relation
δ (delta)	Coefficients on the short-run regressors
Δ (capital delta)	One-month change
ε (epsilon)	Disturbance of the forecasting equation
π (pi)	Year-over-year inflation rate
ρ (rho)	First-order autocorrelation
φ (phi)	Coefficient on the month-end close less the monthly average

1 Introduction

The 10-year Treasury yield is the reference rate for mortgages, for corporate borrowing, and for the discounting of long-dated liabilities. A large literature finds that forecasts of it seldom improve on the forecast that it will not change.[1] This paper develops a forecasting model that the literature has not examined: an error correction equation in which the inflation signal is taken from producer prices rather than from consumer prices.

The specification comes from two companion papers. The first examines which accounts of the policy rate real-time data can separate, and its estimate of the slope of inflation on the unemployment gap in the quarters since the adoption of the inflation target is not statistically distinguishable from zero.[2] The second evaluates supply-side leading indicators, among them the Global Supply Chain Pressure Index (GSCPI) and the producer price index (PPI) for finished goods, as forecasters of core personal consumption expenditures (PCE) inflation in the inflation-targeting era, on the premise that the Phillips curve is flat.[3] The slope of inflation on unemployment is small in U.S. data,[4] and one account attributes the flatness to monetary policy that targets inflation.[5] If slack carries little information about inflation, the hypothesis runs, cost pressure upstream of the consumer carries more, and a bond market that prices expected inflation responds to it. The model puts that hypothesis in a form that can fail: the change in the year-over-year rate of the PPI for finished goods enters a forecasting equation for the yield beside the yield's own dynamics, the policy rate, the yield spread, and the unemployment rate, and the equation is evaluated out of sample.

Three features of the evaluation determine what it can show. First, every regressor is read as it stood when the forecast would have been made; revised data enter only as a variant. Second, the target is the monthly average of the yield, which is predictable from its own last change by construction, so the paper reads each forecast against two benchmarks, the no-change forecast of the literature and the yield at the month-end close, and reports the model with and without the close. Third, the hypothesis is conditional on a monetary regime, so the paper splits the sample at two dates for the start of inflation targeting, each taken from the record of the Federal Open Market Committee (FOMC).

The route is as follows. Section 2 reviews the literature. Section 3 states the target, the forecast origin, and the real-time information set. Section 4 examines persistence and the random walk property and derives the two benchmarks. Section 5 estimates the long-run relation. Section 6 compares the supply-side and the demand-side inflation signal. Section 7 states the model, estimates it, and reports its diagnostics. Section 8, the principal section, evaluates it out of sample and sets the results beside published ones: with the close, the model's root mean squared error (RMSE) at one month is 0.831 times that of the no-change forecast and 1.105 times that of the close, and

[1] Gregory R. Duffee, "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance* 57, no. 1 (2002): 405–443; Francis X. Diebold and Canlin Li, "Forecasting the Term Structure of Government Bond Yields," *Journal of Econometrics* 130, no. 2 (2006): 337–364.

[2] Frank Schmid, *Observational Equivalence in Taylor Rule Estimation: What Is Identified Depends on the Output Gap and the Sample Period* (working paper, 26 July 2026).

[3] Frank Schmid, *Supply-Side Inflation Forecasting When the Phillips Curve Is Flat: A Real-Time Data Approach for the U.S. Inflation-Targeting Era, 2012:Q1–2025:Q4* (working paper, 19 March 2026); Gianluca Benigno, Julian di Giovanni, Jan J. J. Groen, and Adam I. Noble, "The GSCPI: A New Barometer of Global Supply Chain Pressures," Staff Report no. 1017 (Federal Reserve Bank of New York, May 2022).

[4] Jonathon Hazell, Juan Herreño, Emi Nakamura, and Jón Steinsson, "The Slope of the Phillips Curve: Evidence from U.S. States," *Quarterly Journal of Economics* 137, no. 3 (2022): 1299–1344.

[5] Michael McLeay and Silvana Tenreyro, "Optimal Inflation and the Identification of the Phillips Curve," *NBER Macroeconomics Annual* 34 (2020): 199–255.

the test of the PPI term's contribution to the forecasts rejects at none of the four horizons evaluated. Section 9 repeats the analysis inside each inflation-targeting regime. Section 10 reports the conditional forecast from the last origin, Section 11 divides the yield into a real and an inflation component, Section 12 relates the findings to the companion papers, and Section 13 concludes.

2 Literature Review

This section reviews five strands of work and states the gap the model addresses.

2.1 The Random Walk Benchmark

The no-change forecast, the forecast of a random walk, is the benchmark of the literature on interest rate forecasting. Duffee finds that affine term structure models of the standard form forecast Treasury yields less accurately than the no-change forecast, and Diebold and Li, reporting that finding, describe the model of that form as one that "appears largely useless for forecasting."[6] Professional forecasts meet the same benchmark. Kolb and Stekler examine analysts' forecasts of interest rates six months ahead and conclude that they were not more accurate than random walk forecasts.[7]

2.2 Factor Models of the Yield Curve

Diebold and Li model the yield curve by three factors that evolve over time: level, slope, and curvature. Their model improves on no change at the twelve-month horizon; at one month its accuracy is not statistically distinguishable from that of no change. Christensen, Diebold, and Rudebusch derive the arbitrage-free counterpart of that model and report that imposing its structure can improve predictive performance.[8] Later work adds macroeconomic information. Diebold, Rudebusch, and Aruoba join the three factors to measures of real activity, inflation, and the policy rate, and find that the macroeconomic variables affect subsequent movements of the yield curve, with evidence of an influence in the reverse direction as well.[9] Ang and Piazzesi find that models with macroeconomic factors forecast better than models with unobservable factors alone, while the unobservable factors still account for most of the movement at the long end of the curve.[10] A factor-augmented vector autoregression lowers forecast error at horizons of six to twelve months.[11]

[6] Duffee, "Term Premia," 405–443; Diebold and Li, "Forecasting the Term Structure," 337–364.

[7] R. A. Kolb and H. O. Stekler, "How Well Do Analysts Forecast Interest Rates?" *Journal of Forecasting* 15, no. 5 (1996): 385–394.

[8] Jens H. E. Christensen, Francis X. Diebold, and Glenn D. Rudebusch, "The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models," *Journal of Econometrics* 164, no. 1 (2011): 4–20.

[9] Francis X. Diebold, Glenn D. Rudebusch, and S. Boragan Aruoba, "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach," *Journal of Econometrics* 131, nos. 1–2 (2006): 309–338.

[10] Andrew Ang and Monika Piazzesi, "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," *Journal of Monetary Economics* 50, no. 4 (2003): 745–787.

[11] Emanuel Moench, "Forecasting the Yield Curve in a Data-Rich Environment: A No-Arbitrage Factor-Augmented VAR Approach," *Journal of Econometrics* 146, no. 1 (2008): 26–43.

2.3 Macro-Finance and Survey-Based Approaches

Allowance for slow-moving trends in inflation and the real rate improves long-horizon forecasts.[12] For the 10-year Treasury rate itself, Baghestani compares the random walk with an autoregression augmented by the expected inflation rate and finds that the augmented model forecasts more accurately.[13] Cieslak documents persistent errors in investors' expectations of the short rate over the business cycle, errors that make excess returns on Treasury bonds predictable after the fact without a time-varying risk premium.[14] The model of this paper belongs to the error correction tradition, in which two cointegrated series, here the long rate and the short rate, share a long-run relation and deviations from it forecast subsequent changes.[15] Under the expectations hypothesis the spread between the two rates forecasts their changes; Campbell and Shiller find that it forecasts changes in short rates in the direction the hypothesis implies and changes in the long rate in the opposite direction.[16]

2.4 Machine Learning Approaches

A recent strand applies machine learning to the same task; Wang compares several such models for short-term forecasts of U.S. Treasury yields.[17] The model of this paper is linear and holds seven regressors, so that each term can be tested on its own.

2.5 Time-Averaged Targets and Real-Time Data

Two further strands govern how a forecast of the yield is evaluated. The first concerns the target. Working showed that the first differences of averages of a random walk are serially correlated, with a first-order autocorrelation that approaches one quarter as the number of observations averaged grows.[18] A forecast of a period average can therefore improve on the no-change forecast of the average without any information beyond the path of the series inside the last period. For the real price of oil, Ellwanger and Snudden show that the no-change forecast taken from the final observation of the month is more accurate than the no-change forecast of the monthly average at horizons up to one year, and that at the one-month horizon the forecasts that improve on the monthly-average benchmark do not improve on the end-of-month one.[19] Section 4.2 applies the same reasoning to the yield.

[12] Michael D. Bauer and Glenn D. Rudebusch, "Interest Rates under Falling Stars," *American Economic Review* 110, no. 5 (2020): 1316–1354.

[13] Hamid Baghestani, "Forecasting the 10-Year US Treasury Rate," *Journal of Forecasting* 29, no. 8 (2010): 673–688.

[14] Anna Cieslak, "Short-Rate Expectations and Unexpected Returns in Treasury Bonds," *Review of Financial Studies* 31, no. 9 (2018): 3265–3306.

[15] Robert F. Engle and C. W. J. Granger, "Co-Integration and Error Correction: Representation, Estimation, and Testing," *Econometrica* 55, no. 2 (1987): 251–276.

[16] John Y. Campbell and Robert J. Shiller, "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies* 58, no. 3 (1991): 495–514.

[17] Y.-F. Wang, "Comparing Machine Learning Models for Short-Term U.S. Treasury Yield Forecasting," *Concurrency and Computation: Practice and Experience* (2025): e70265.

[18] Holbrook Working, "Note on the Correlation of First Differences of Averages in a Random Chain," *Econometrica* 28, no. 4 (1960): 916–918.

[19] Reinhard Ellwanger and Stephen Snudden, "Forecasts of the Real Price of Oil Revisited: Do They Beat the Random Walk?" *Journal of Banking & Finance* 154 (2023): 106962.

The second is real-time data. Estimates on revised data use values that no forecaster had, and forecasts made from different vintages of one series can differ in particular periods.[20] The evaluation below is conducted in real time throughout.

2.6 Gap in the Literature

The work reviewed here takes its inflation information from consumer prices, from expected inflation, or from the factors of the yield curve itself. It has not, to the author's knowledge, placed a supply-side inflation signal in a forecasting equation for the yield. The hypothesis this paper tests is that, if the Phillips curve is flat, the channel from slack to inflation carries little and the inflation signal a bond market would price moves upstream to producer costs. This paper builds the equation that hypothesis implies and evaluates it.

3 Data and Real-Time Methodology

This section states what is forecast, when the forecast is made, and what is known at that moment.

3.1 Target and Forecast Origin

The target is the monthly average of the 10-year constant-maturity Treasury yield, the average of business days published by the Board of Governors of the Federal Reserve System in its H.15 release. The forecast origin is the close of the final business day of a month. From an origin in month t, the paper forecasts the monthly average of each of the twelve following months.

Two quantities are known at the origin and are kept apart throughout: the average of the origin month, y_t, and the yield at the origin's close, c_t. On the last origin, 30 September 2026, the average of the month was 4.99 percent and the close was 5.29 percent. Figure 1 shows the monthly average with the effective federal funds rate over the sample.

[20] Dean Croushore and Tom Stark, "A Real-Time Data Set for Macroeconomists," *Journal of Econometrics* 105, no. 1 (2001): 111–130.

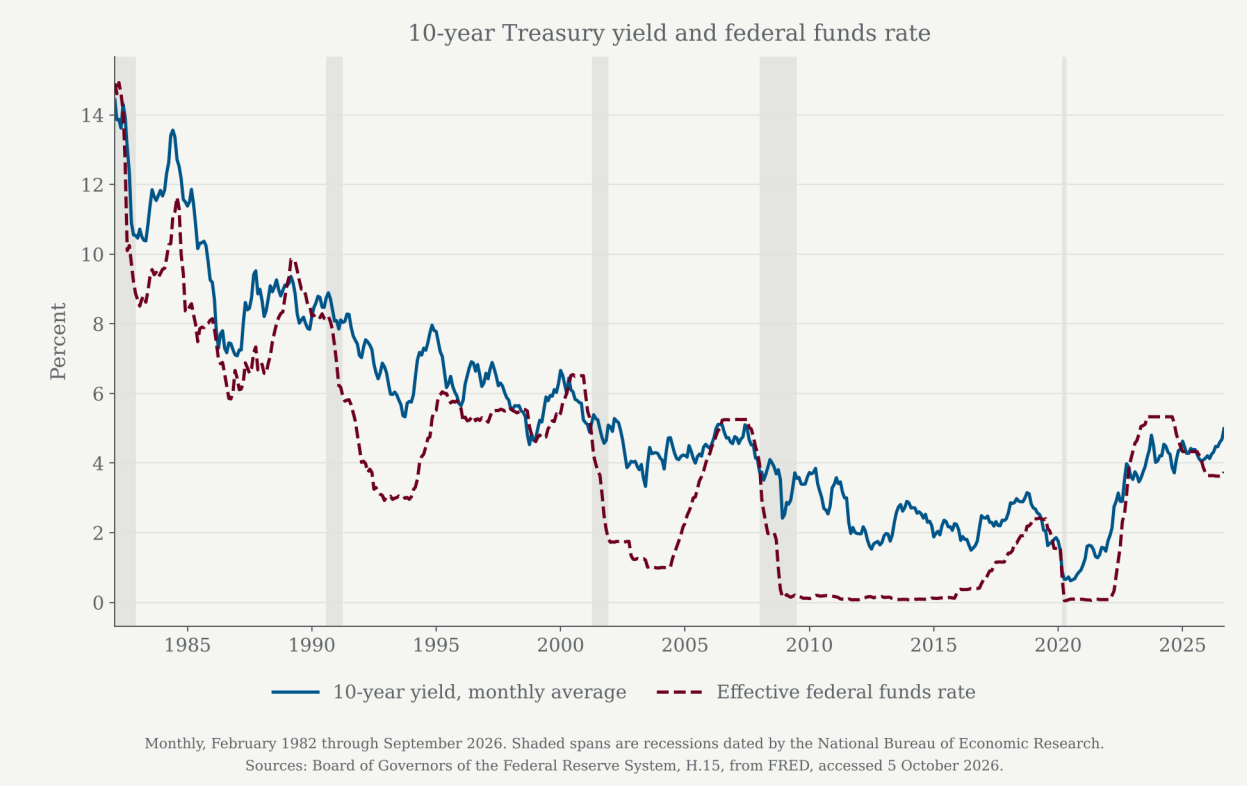


Figure 1. Monthly average of the 10-year constant-maturity Treasury yield (solid deep blue line) and the effective federal funds rate (dashed maroon line), in percent on the vertical axis, monthly, February 1982 through September 2026. Shaded spans are recessions dated by the National Bureau of Economic Research. Sources: Board of Governors of the Federal Reserve System, H.15, from FRED, Federal Reserve Bank of St. Louis, accessed 5 October 2026.

3.2 What Is Known at the Origin

The series fall into two classes. Market rates are not revised, and their values through the origin's close are known at the close. The effective federal funds rate enters as the mean of the daily rate over the origin month through the day before the origin, because the rate for a given day is published on the next.

The PPI for finished goods, the unemployment rate, and the PCE price index are revised after first publication. For each of them the paper reads the vintage in force on the final calendar day of the origin month from the real-time archive of FRED, and takes every transformation inside that one vintage. The year-over-year PPI rate at an origin is the latest rate that vintage contains, and its monthly change is the difference between that rate and the rate one month earlier in the same vintage. In 531 of the 536 origins the latest PPI observation is for the month before the origin; in 4 it is two months old and in 1 it is three months old. The unemployment rate is one month old at 534 origins and two months old at 2.

The archive's first PPI vintage is dated 12 February 1982, which sets the first origin. The series is the index for finished goods throughout; the archive carries it under one series code until March 2015 and under its successor code from then. The first PCE vintage is dated 1 August 2000, so

in-sample comparisons that use the PCE rate begin with the origin of August 2000, and out-of-sample ones with that of August 2010, the first whose rolling window of 120 months lies within the archived vintages.

A second real-time source covers part of the same ground. The Real-Time Data Set for Macroeconomists of the Federal Reserve Bank of Philadelphia carries monthly vintages of the finished-goods index from November 1998. Over the 335 month-end vintages the two sources share, the latest observation is for the same month in 334, and of the 670 year-over-year rates compared, 659 agree to within 0.05 percentage point, with a largest difference of 0.052.

Revision is small relative to the signal. Over the sample, the monthly change in the PPI rate as first published has a standard deviation of 1.01 percentage points; its difference from the same change in the latest vintage has a standard deviation of 0.18, and the two are correlated at 0.984.

Two further series are descriptive and enter no forecast: the GSCPI, which is republished whole at each release and is read in the vintage at the access date, and the yield on 10-year inflation-protected Treasury securities with the breakeven inflation rate it implies, which are market rates.

4 Unit Root and Random Walk Analysis

This section examines the persistence of the series, shows what time averaging does to the target, and states the two benchmarks.

4.1 Persistence

Table 1 reports augmented Dickey-Fuller statistics.[21] For the level of the yield the statistic is -2.898 ($p = 0.046$), at the edge of the 5 percent level over a sample that begins near the highest yields of the record, and for the funds rate it is -2.596 ($p = 0.094$). The monthly change in the yield has a statistic of -11.790. The PPI rate and the GSCPI reject a unit root ($p = 0.002$ and 0.002). The paper treats the two interest rates as persistent series and models the yield in changes.

[21] Said E. Said and David A. Dickey, "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order," *Biometrika* 71, no. 3 (1984): 599–607.

Table 1. Augmented Dickey-Fuller statistics (constant; lag length by the Akaike criterion, maximum 12) and Engle-Granger statistics for the 10-year yield and the federal funds rate, monthly, February 1982 through September 2026 unless a longer sample is stated. The 5 percent critical value of the Engle-Granger statistic is in the final column. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, each from FRED and its real-time archive, Federal Reserve Bank of St. Louis, accessed 5 October 2026; author's calculations.

Series or pair	Statistic	p-value	5% critical value
10-year yield, monthly average	−2.898	0.046	—
Federal funds rate	−2.596	0.094	—
PPI finished goods, year-over-year rate	−3.852	0.002	—
Unemployment rate	−3.017	0.033	—
Supply chain pressure index, January 1998 through August 2026	−3.881	0.002	—
10-year yield, monthly change	−11.790	<0.001	—
Yield and funds rate, Engle-Granger	−3.273	0.059	−3.348
Yield and funds rate, from June 1976	−3.153	0.078	−3.346

The variance ratio at twelve months, the variance of the twelve-month change divided by twelve times the variance of the one-month change, is 1.511 for the monthly average and 1.099 for the month-end close.[22] A ratio above one indicates positive serial dependence in monthly changes. Most of the departure of the monthly average from one is the averaging itself, as the next subsection shows.

4.2 Time Averaging and the Two Benchmarks

Suppose the daily yield were a random walk, so that no information forecasts its changes. The monthly average would still be forecastable. Consecutive changes in the average share the daily movements inside the month they have in common, and their first-order autocorrelation ρ approaches one quarter as the number of days averaged grows.[23] In the sample, the change in the monthly average has a first-order autocorrelation of 0.336; the change in the month-end close has 0.102. Table 2 and Figure 2 report the comparison.

Table 2. Serial dependence of the monthly change in the 10-year yield under two samplings of the daily series, February 1982 through September 2026. The monthly average is the published average of business days; the month-end close is the yield on the final business day of the month. The variance ratio is the variance of the 12-month change divided by 12 times the variance of the one-month change; standard deviations are in percentage points. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, each from FRED and its real-time archive, Federal Reserve Bank of St. Louis, accessed 5 October 2026; author's calculations.

Statistic	Monthly average	Month-end close
First-order autocorrelation of the monthly change	0.336	0.102
Variance ratio, 12 months	1.511	1.099
Standard deviation of the monthly change	0.256	0.305

[22] Andrew W. Lo and A. Craig MacKinlay, "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test," *Review of Financial Studies* 1, no. 1 (1988): 41–66.

[23] Working, "Note on the Correlation," 916–918.

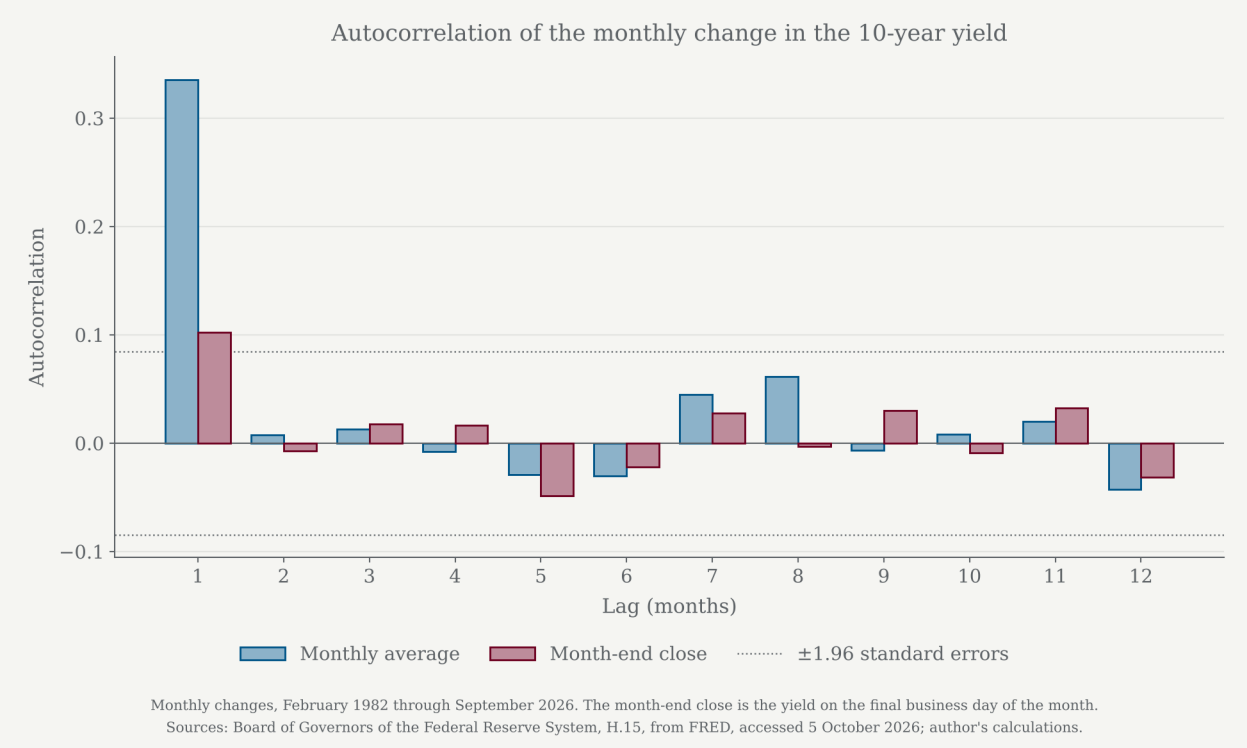


Figure 2. Autocorrelation of the monthly change in the 10-year yield (vertical axis) at lags of one through twelve months (horizontal axis), February 1982 through September 2026. Deep blue bars are for the monthly average and maroon bars for the yield at the month-end close; dotted gray lines mark ± 1.96 standard errors under no autocorrelation. Sources: Board of Governors of the Federal Reserve System, H.15, from FRED, accessed 5 October 2026; author's calculations.

The same reasoning gives the forecast. If the daily yield is a random walk, the expected value of every future daily yield, and so of every future monthly average, is the last yield observed:

$$\mathrm{E}\,[\,y_{t+h} \mid \text{information at the origin}\,] = c_t, \qquad h = 1, 2, \ldots \tag{1}$$

The no-change forecast of the average, y_t, discards the part of the origin month's path that lies between its average and its close. The gap $c_t - y_t$ has a standard deviation of 0.157 percentage points in the sample, and its correlation with the next month's change in the average is 0.689.

The paper therefore reports two benchmarks. The *no-change forecast of the average* is y_t; it is the benchmark a comparison of monthly averages conventionally uses, and the one against which published results are stated. The *month-end close* is c_t; it is the forecast that equation (1) implies. A procedure that improves on the first and not on the second has not been shown to carry information beyond the path of the yield inside the origin month. Claims about information beyond the yield's own path are read against the close.

5 Cointegration and Error Correction

This section estimates the long-run relation on which the error correction form rests.

Under the expectations hypothesis a long yield reflects the expected path of the short rate and a term premium, so the two rates move together over long spans. The yield and the federal funds rate f_t are tied by a long-run relation,

$$y_t = \alpha + \beta f_t + z_t, \tag{2}$$

whose residual z_t is the deviation the model corrects. The full-sample estimates are α = 1.961 and β = 0.846: a change of one percentage point in the funds rate is associated with a change of less than one percentage point in the yield, which is what a long rate that averages expected short rates implies when the short rate reverts toward its mean. Term structure models that separate expected short rates from term premia give the framework in which such a relation is read.[24]

Table 1 reports the cointegration statistics.[25] Over the sample, the Engle-Granger statistic for the yield and the funds rate is -3.273 against a 5 percent critical value of -3.348 (p = 0.059). The null of no cointegration is not rejected at the 5 percent level, and the out-of-sample evaluation does not depend on it: the relation is re-estimated inside every estimation window, and Section 8 evaluates the equation that carries it. The in-sample t-statistic of the error correction term does depend on it, and Section 7 reads that statistic against a bootstrap distribution.

6 Supply-Side versus Demand-Side Inflation Signals

This section compares the two inflation signals the model could carry.

The companion paper on inflation evaluates supply-side indicators as leading indicators of consumer inflation, and the authors of the GSCPI relate inflation outcomes to their index, most of all at the level of producer prices.[26] A bond market that looks forward would then respond to producer prices before the same pressure reaches the consumer price measure. Figure 3 shows the PPI rate and the PCE rate as first published, with the yield.

[24] Tobias Adrian, Richard K. Crump, and Emanuel Moench, "Pricing the Term Structure with Linear Regressions," *Journal of Financial Economics* 110, no. 1 (2013): 110–138.

[25] Engle and Granger, "Co-Integration and Error Correction," 251–276.

[26] Benigno, di Giovanni, Groen, and Noble, "The GSCPI."

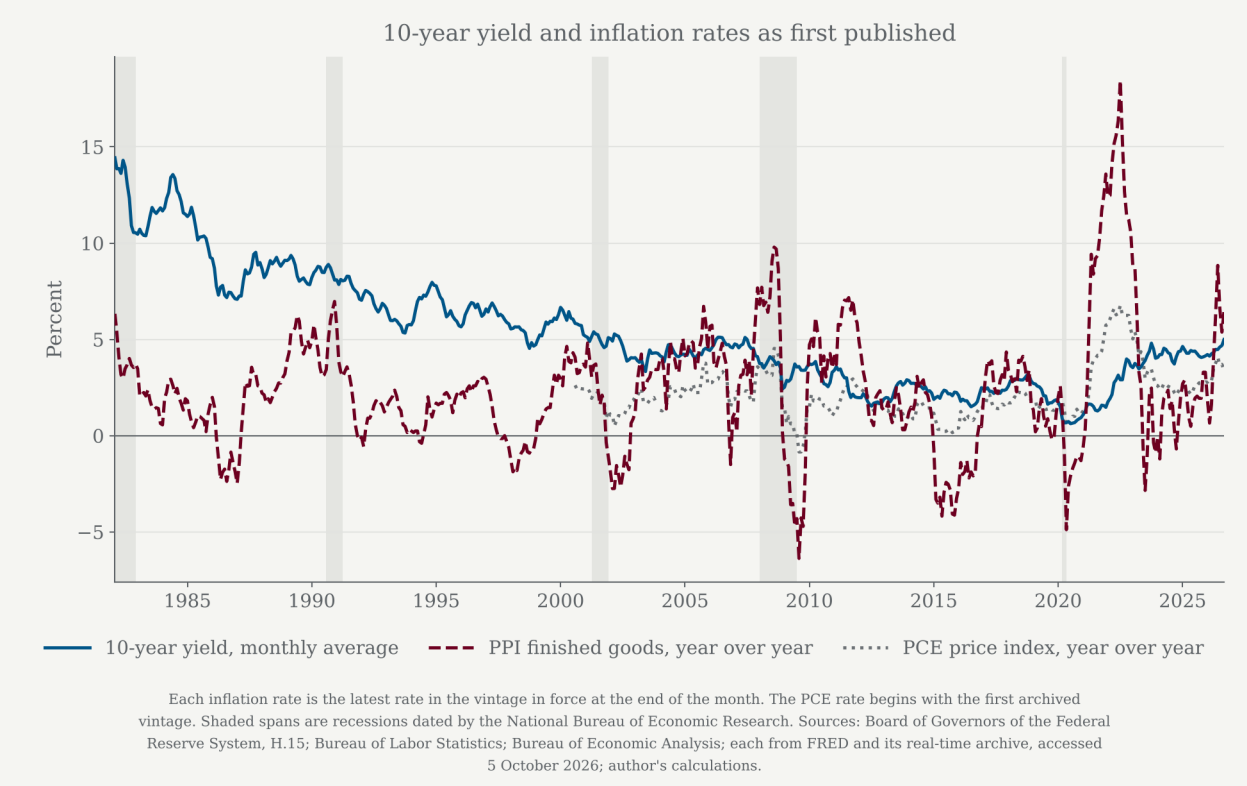


Figure 3. Monthly average of the 10-year yield (solid deep blue line), the year-over-year rate of the PPI for finished goods (dashed maroon line), and the year-over-year rate of the PCE price index (dotted gray line), in percent on the vertical axis. Each inflation rate is the latest rate in the vintage in force at the end of the month on the horizontal axis; the PCE rate begins with the first archived vintage. Shaded spans are recessions dated by the National Bureau of Economic Research. Sources: Board of Governors of the Federal Reserve System, H.15; Bureau of Labor Statistics; Bureau of Economic Analysis; each from FRED and its real-time archive, accessed 5 October 2026; author's calculations.

Table 3 estimates the one-month equation of Section 7 under each signal, holding every other regressor, on the origins from August 2000, where a PCE vintage exists. Each specification is also read by its Bayesian information criterion (BIC), the lower value preferred.[27] The coefficient on the PPI term, the latest published change in the PPI rate, has a t-statistic of 0.18 (p = 0.859) and that on the PCE term, the same change in the PCE rate, in its place -0.84 (p = 0.401). The two specifications have adjusted R^2 of 0.086 and 0.088 and BIC of -72.8 and -73.4. Neither coefficient is statistically distinguishable from zero in the yield equation. The BIC is lower with the PCE term, by less than one in either form of the model, and the model keeps the PPI term, whose content is the hypothesis under test. Section 8.4 repeats the comparison out of sample.

[27] Gideon Schwarz, "Estimating the Dimension of a Model," *The Annals of Statistics* 6, no. 2 (1978): 461–464.

Table 3. The inflation term of the one-month equation under a supply-side and a demand-side signal, on the 312 origins from August 2000 for which a vintage of the PCE price index is archived. Each row holds every other regressor and replaces the inflation term: the latest published change in the year-over-year rate of the PPI for finished goods, or of the PCE price index, each read from the vintage in force at the origin. BIC is the Bayesian information criterion; the lower value is preferred. Standard errors are Newey-West. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, each from FRED and its real-time archive, Federal Reserve Bank of St. Louis, accessed 5 October 2026; author's calculations. PCE: Bureau of Economic Analysis, from the same archive.

Inflation signal	Coefficient	Std. error	t	p-value	Adj. R^2	BIC
Equation						
PPI finished goods	0.0022	0.0123	0.18	0.859	0.086	−72.8
PCE price index	−0.0279	0.0332	−0.84	0.401	0.088	−73.4
Equation with the close						
PPI finished goods	−0.0012	0.0089	−0.13	0.893	0.392	−195.3
PCE price index	−0.0114	0.0268	−0.42	0.671	0.393	−195.5

7 Model Specification and Estimation

This section states the model, the tests that compare forecasts, the estimates, and the diagnostics.

7.1 Specification

For horizon h the forecasting equation is

$$y_{t+h} - y_t = \delta_0 + \delta_1 \Delta y_t + \delta_2 \Delta y_{t-1} + \delta_3 \Delta f_t + \delta_4 \Delta \pi_t^P + \delta_5 \Delta u_t + \delta_6 \Delta s_t + \gamma z_t + \varepsilon_{t+h}, \tag{3}$$

where π_t^P is the latest published year-over-year PPI rate, whose term $\delta_4 \Delta \pi_t^P$ the paper calls the PPI term, u_t the latest published unemployment rate, s_t the spread between the monthly averages of the 10-year and 2-year yields, and z_t the deviation from the long-run relation (2). Each horizon has its own coefficients; the forecast is direct and requires no assumption about the future path of any regressor. The supply-side hypothesis is $\delta_4 > 0$, and error correction is $\gamma < 0$. The spread change contains the yield change, so δ_1 and δ_6 are not separately interpretable.

The *equation with the close* adds the gap of Section 4.2:

$$y_{t+h} - y_t = \delta_0 + \cdots + \gamma z_t + \varphi \, (c_t - y_t) + \varepsilon_{t+h}. \tag{4}$$

The month-end close benchmark is the special case $\varphi = 1$ with every other coefficient zero. The paper calls equation (3) the equation and equation (4) the equation with the close; the two are the forms of the model.

7.2 Tests and Reference Distributions

Forecasts are compared by two tests. Where one procedure is a restricted case of another, the comparison uses the Clark-West statistic, which adjusts the difference in squared errors for the

estimation noise of the larger model and is read one-sided.[28] Where neither nests the other, it uses the Diebold-Mariano statistic with the Harvey-Leybourne-Newbold correction, two-sided.[29] Variances are Newey-West.[30] The Clark-West statistic and the ratio of RMSEs answer different questions: the first tests whether the larger model's additional terms have content in population, and the second reports whether the estimated model forecast more accurately. A model can pass the first and fail the second when its additional coefficients are estimated with enough noise to offset their content.

Reference distributions for the forecast comparisons against the close and for the t-statistic of γ are taken by bootstrap, with 4,999 replications for each statistic. Forecast errors at horizons beyond one month overlap, and the asymptotic distributions of both statistics are approximations in a record of this length. For nested comparisons the paper uses the fixed-regressor wild bootstrap of Clark and McCracken: the residuals of the larger model, fitted on all origin-target pairs, are given a moving-average structure of order $h - 1$ estimated by the regressions of Hannan and Rissanen; the fitted innovations are multiplied by independent draws of $+1$ or -1; and the rebuilt residuals are added to the fitted values of the smaller model. The out-of-sample exercise is then repeated on each replicate with every regressor held at its sample value.[31] Where neither procedure nests the other, the differential of squared errors is centered and resampled in circular blocks of length $\max(2h, \lceil P^{1/3} \rceil)$, P being the number of forecasts.[32] The t-statistic of γ does not have a normal distribution when the yield and the funds rate are not cointegrated.[33] Its reference distribution is taken from a recursive wild bootstrap under $\gamma = 0$: the monthly average is rebuilt from the sign-randomized residuals of the equation without the term, and the long-run relation is re-estimated on each rebuilt series.[34] The other regressors are held at their sample values in every replicate. Where a table reports a bootstrap p-value it gives the asymptotic p-value in parentheses, and the text reads the bootstrap value. One-month comparisons with the no-change forecast, whose errors do not overlap, are read against the asymptotic distribution and are labeled as such.

7.3 Estimates

Table 4 reports the estimates of the one-month equation on the full sample, with and without the close. In the equation, the coefficient on the latest change in the average is 0.3763 with a t-statistic of 9.13, and removing that regressor lowers the R^2 of 0.142 by 0.104. With the close in the equation, the same coefficient is 0.0606 (t = 1.44), the coefficient on the gap is 1.0851, and the R^2 is 0.488.

[28] Todd E. Clark and Kenneth D. West, "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models," *Journal of Econometrics* 138, no. 1 (2007): 291–311.

[29] Francis X. Diebold and Roberto S. Mariano, "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics* 13, no. 3 (1995): 253–263; David Harvey, Stephen Leybourne, and Paul Newbold, "Testing the Equality of Prediction Mean Squared Errors," *International Journal of Forecasting* 13, no. 2 (1997): 281–291.

[30] Whitney K. Newey and Kenneth D. West, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* 55, no. 3 (1987): 703–708.

[31] Todd E. Clark and Michael W. McCracken, "Reality Checks and Comparisons of Nested Predictive Models," *Journal of Business & Economic Statistics* 30, no. 1 (2012): 53–66; E. J. Hannan and J. Rissanen, "Recursive Estimation of Mixed Autoregressive-Moving Average Order," *Biometrika* 69, no. 1 (1982): 81–94.

[32] Hans R. Künsch, "The Jackknife and the Bootstrap for General Stationary Observations," *The Annals of Statistics* 17, no. 3 (1989): 1217–1241; Dimitris N. Politis and Joseph P. Romano, "A Circular Block-Resampling Procedure for Stationary Data," in *Exploring the Limits of Bootstrap*, ed. Raoul LePage and Lynne Billard (New York: Wiley, 1992), 263–270.

[33] Anindya Banerjee, Juan J. Dolado, and Ricardo Mestre, "Error-Correction Mechanism Tests for Cointegration in a Single-Equation Framework," *Journal of Time Series Analysis* 19, no. 3 (1998): 267–283.

[34] Sílvia Gonçalves and Lutz Kilian, "Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form," *Journal of Econometrics* 123, no. 1 (2004): 89–120.

The hypothesis that the gap's coefficient equals one is not rejected (p = 0.237). The serial correlation that the lagged change captures in the first panel is the time averaging of Section 4.2: once the close is present, the lagged change has no coefficient distinguishable from zero.

Table 4. Estimates of the one-month equation, the dependent variable being the change in the monthly average of the 10-year yield from the origin month to the next, in percentage points; 535 origins, February 1982 through August 2026. Regressors are dated at the origin: the two latest changes in the average, the change in the funds rate, the latest published change in the PPI rate and in the unemployment rate, the change in the spread, the deviation z from the long-run relation, and, in the right panel, the origin's close less the origin month's average. The final row gives the one-sided bootstrap p-value of the t-statistic of z under a zero coefficient. Standard errors are Newey-West with bandwidth 5. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, each from FRED and its real-time archive, Federal Reserve Bank of St. Louis, accessed 5 October 2026; author's calculations.

	Equation			Equation with the close		
Regressor	Coefficient	Std. error	t	Coefficient	Std. error	t
Constant	−0.0121	0.0103	−1.17	0.0037	0.0081	0.46
Δy_t	0.3763	0.0412	9.13	0.0606	0.0421	1.44
Δy_{t-1}	−0.1409	0.0474	−2.98	−0.0295	0.0391	−0.75
Δf_t	0.0515	0.0637	0.81	0.0402	0.0497	0.81
$\Delta \pi_t^P$	0.0039	0.0112	0.35	0.0027	0.0079	0.34
Δu_t	0.0054	0.0081	0.67	0.0016	0.0065	0.24
Δs_t	−0.0317	0.0993	−0.32	−0.0601	0.0804	−0.75
z_t	−0.0203	0.0095	−2.14	−0.0187	0.0073	−2.55
$c_t - y_t$	—	—	—	1.0851	0.0720	15.07
Observations		535			535	
R^2		0.142			0.488	
Adjusted R^2		0.130			0.481	
Residual standard error		0.239			0.185	
Breusch-Godfrey, 12 lags, p		0.137			0.879	
ARCH LM, 12 lags, p		0.012			0.001	
Bootstrap p-value of z_t, one-sided		0.346			0.082	

The remaining short-run terms add little. In the equation, the change in the funds rate has a coefficient of 0.0515 (t = 0.81), the change in the unemployment rate 0.0054 (t = 0.67), and the change in the spread −0.0317 (t = −0.32). Removing any one of the three lowers the R^2 by no more than 0.002; removing the second lag of the yield change lowers it by 0.016, and removing the error correction term by 0.009. With the close in the equation the three t-statistics are 0.81, 0.24, and −0.75.

The PPI term has a coefficient of 0.0039 (t = 0.35) in the equation and 0.0027 (t = 0.34) with the close. Neither is statistically distinguishable from zero; the coefficients are in percentage points of yield per percentage point of change in the PPI rate. Removing the term leaves the R^2 unchanged at three decimals.

The error correction term has t-statistics of −2.14 and −2.55 in the two panels, and its coefficient with the close is −0.0187 per month per percentage point of deviation from the long-run relation. Each statistic lies beyond the 5 percent point of the normal distribution. Under the null that the coefficient is zero, the bootstrap distribution of the statistic lies to the left of the normal: its 5th percentile is −3.23 in the equation and −2.81 with the close. The first of the two equals, at

two decimals, the 5 percent critical value that Banerjee, Dolado, and Mestre tabulate for one regressor and a constant in a sample of 500 observations.[35] The bootstrap p-values, one-sided in the direction of error correction, are 0.346 and 0.082, so the term is not statistically distinguishable from zero at the 5 percent level in either panel, in agreement with the Engle-Granger statistic of Section 5. The joint hypothesis that every coefficient other than the gap's is zero has an asymptotic p-value of 0.075.

7.4 Diagnostics

Figure 4 shows the residuals of the equation. Their first-order autocorrelation is -0.005, and the Breusch-Godfrey test does not reject the absence of serial correlation ($p = 0.137$ for the equation and 0.879 with the close). The residuals have tails heavier than the normal distribution's: excess kurtosis is 1.99 and skewness -0.24, and 5.2 percent of the residuals lie beyond two residual standard errors. The largest, 1.25 percentage points in absolute value, belongs to the origin of September 1982. The autoregressive conditional heteroskedasticity (ARCH) Lagrange multiplier (LM) test rejects constant variance ($p = 0.012$ and 0.001); the standard errors are robust to both features, and the forecast intervals of Section 10 are taken from empirical percentiles for the same reason.[36]

[35] Banerjee, Dolado, and Mestre, "Error-Correction Mechanism Tests," 267–283.

[36] T. S. Breusch, "Testing for Autocorrelation in Dynamic Linear Models," *Australian Economic Papers* 17, no. 31 (1978): 334–355; L. G. Godfrey, "Testing against General Autoregressive and Moving Average Error Models When the Regressors Include Lagged Dependent Variables," *Econometrica* 46, no. 6 (1978): 1293–1301; Robert F. Engle, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica* 50, no. 4 (1982): 987–1007.

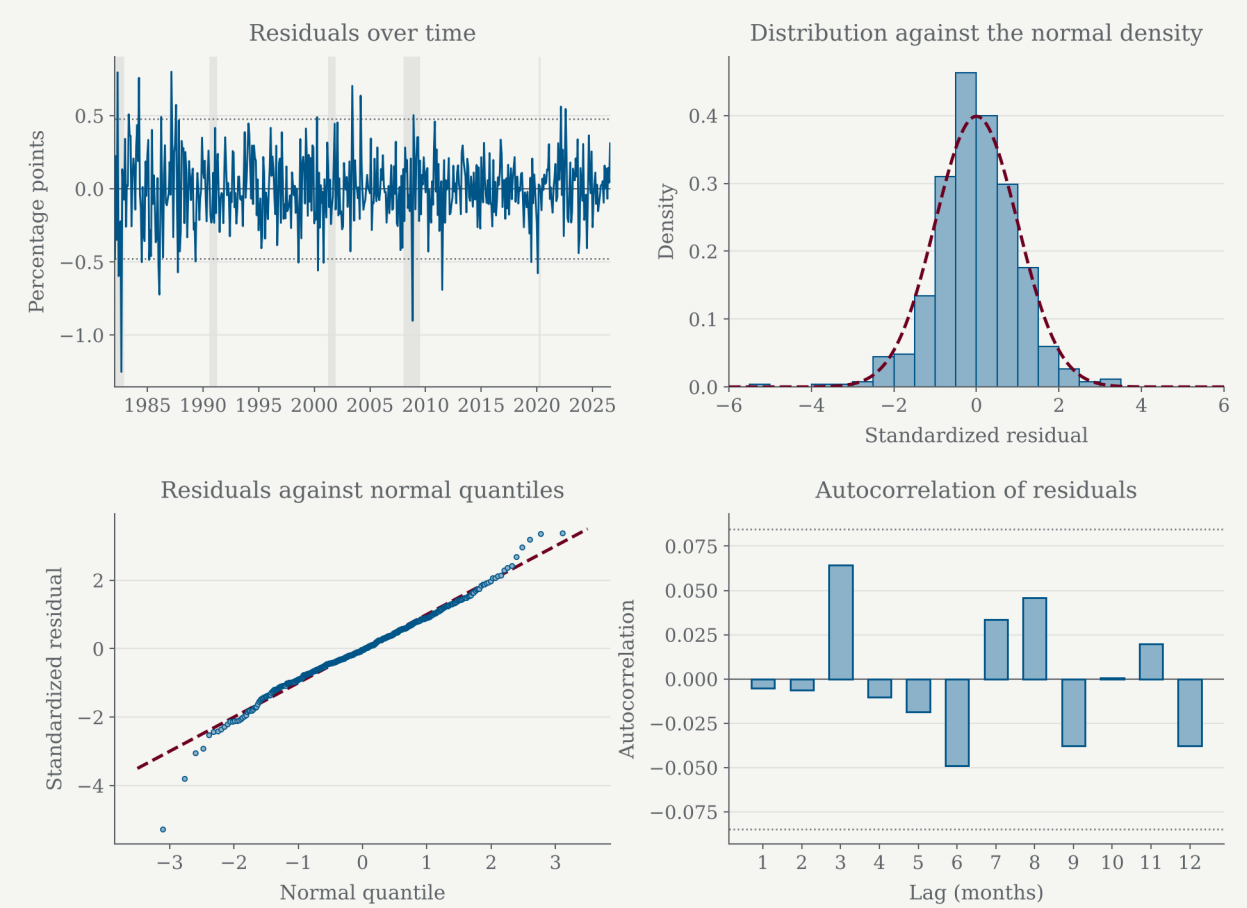


Figure 4. Diagnostics of the one-month equation, February 1982 through August 2026. Upper left: residuals in percentage points by origin, with dotted lines at two residual standard errors and shaded spans for recessions dated by the National Bureau of Economic Research. Upper right: histogram of the standardized residuals with the normal density (dashed maroon line). Lower left: standardized residuals against normal quantiles, with the line of equal quantiles (dashed maroon line). Lower right: autocorrelation of the residuals at lags of one through twelve months, with dotted lines at ± 1.96 standard errors under no autocorrelation. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, from FRED and its real-time archive, accessed 5 October 2026; author's calculations.

8 Out-of-Sample Performance and Comparison with the Literature

This section evaluates the forecasts. Section 8.1 states the scheme; Section 8.2 reports accuracy by horizon; Section 8.3 sets the results beside published ones; Section 8.4 isolates the PPI term; Section 8.5 reports robustness; Section 8.6 states the limitations.

8.1 Scheme

At each origin, the coefficients of the direct equation for horizon h are estimated on the 120 most recent origin-target pairs whose target was already known, and the long-run relation (2) is re-estimated on the same window, so no forecast uses a coefficient estimated on later data. The procedures compared are the two benchmarks; an autoregression in the two latest changes in the average; the equation (3); a regression on the gap alone; and the equation with the close (4). One-month forecasts number 415, from the origin of February 1992 through that of August 2026; twelve-month forecasts number 393.

8.2 Forecast Accuracy

Table 5 and Figure 5 report the results. The month-end close has the lowest RMSE of the six procedures at each of the twelve horizons. At one month its RMSE is 0.163 percentage points and that of the no-change forecast of the average is 0.217; equal accuracy of the two is rejected (bootstrap p <0.001). At three months the ratio of the close's RMSE to the no-change forecast's is 0.966 and at twelve months 0.983: the advantage of the close is the within-month path, and it is a smaller share of the forecast error as the horizon lengthens.

Table 5. Out-of-sample accuracy of direct forecasts of the monthly average of the 10-year yield, by horizon. Each forecast is made at the close of the final business day of the origin month from coefficients estimated on the 120 most recent origin-target pairs whose target was known at the origin. The first two rows are root mean squared errors (RMSE) in percentage points; the next block divides each procedure's RMSE by that of the month-end close, so a value above 1 is a less accurate forecast than the close; a later block divides by the RMSE of the no-change forecast of the average, the benchmark of the literature. The share of forecasts with the sign of the realized change is taken over the forecasts whose realized change is not zero. One-month origins run from February 1992 through August 2026. Each p-value is the bootstrap value, with the asymptotic value in parentheses. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, each from FRED and its real-time archive, Federal Reserve Bank of St. Louis, accessed 5 October 2026; author's calculations.

Horizon (months)	1	3	6	12
RMSE, month-end close	0.163	0.422	0.619	0.856
RMSE, no change of the average	0.217	0.437	0.630	0.870
RMSE relative to the month-end close				
Autoregression	1.284	1.044	1.040	1.060
Equation	1.352	1.111	1.146	1.271
Equation without the PPI term	1.340	1.100	1.143	1.261
Regression on the close alone	1.002	1.008	1.019	1.043
Equation with the close	1.105	1.091	1.126	1.252
Same, without the PPI term	1.090	1.087	1.122	1.241
p-value of equal accuracy with the close (Diebold-Mariano, two-sided)				
Autoregression	<0.001 (<0.001)	0.031 (0.017)	0.258 (0.181)	0.146 (0.176)
Equation	<0.001 (<0.001)	0.005 (0.002)	0.013 (0.005)	0.018 (0.004)
Regression on the close alone	0.794 (0.818)	0.525 (0.467)	0.498 (0.432)	0.286 (0.327)
Equation with the close	0.245 (0.030)	0.015 (0.006)	0.023 (0.012)	0.020 (0.007)
p-value of no gain over the close (Clark-West, one-sided)				
Regression on the close alone	0.110 (0.141)	0.318 (0.378)	0.391 (0.446)	0.434 (0.510)
Equation with the close	0.218 (0.239)	0.127 (0.126)	0.184 (0.166)	0.353 (0.352)
RMSE relative to the no-change forecast of the average				
Autoregression	0.966	1.008	1.022	1.042
Equation	1.016	1.073	1.127	1.250
Equation with the close	0.831	1.054	1.107	1.231
Month-end close	0.752	0.966	0.983	0.983
Share of forecasts with the sign of the realized change, percent				
Equation	52.2	49.8	49.6	47.8
Equation with the close	67.6	54.4	51.4	50.9
Month-end close	66.3	57.6	54.8	55.5
Forecasts	415	411	405	393

Against the no-change forecast of the average, the model's record depends on the horizon and on whether it is given the close. The equation with the close has an RMSE of 0.181 percentage points at one month, 0.831 times that of no change, and equal accuracy with no change is rejected there (Diebold-Mariano p = 0.001, asymptotic). The equation without the close has an RMSE of 0.221, a ratio of 1.016, and the autoregression a ratio of 0.966. At three, six, and twelve months the ratios are 1.054, 1.107, and 1.231 for the equation with the close and 1.073, 1.127, and 1.250 for the equation: beyond one month neither form of the model improves on no change.

The equation with the close forecasts the sign of the next month's change in the average in 67.6 percent of months, the close in 66.3 percent, and the equation in 52.2 percent. Mean absolute

errors at one month order the procedures as the RMSEs do: 0.128 percentage points for the close, 0.138 for the equation with the close, 0.167 for no change, and 0.170 for the equation.

The equation is less accurate than the close at every horizon. Its RMSE relative to the close is 1.352 at one month, 1.111 at three, 1.146 at six, and 1.271 at twelve, and equal accuracy is rejected at each (bootstrap $p < 0.001$, 0.005, 0.013, and 0.018). The autoregression's gain over the no-change benchmark is the time averaging of Section 4.2, and the equation's additional coefficients cost more accuracy than they add.

Giving the equation the close narrows the distance and does not remove it. The regression on the gap alone has an RMSE of 1.002 times the close's at one month; the equation with the close has 1.105 at one month and 1.252 at twelve. The Clark-West test of the equation with the close against the close itself does not reject at any of the four horizons (bootstrap p = 0.218, 0.127, 0.184, and 0.353).

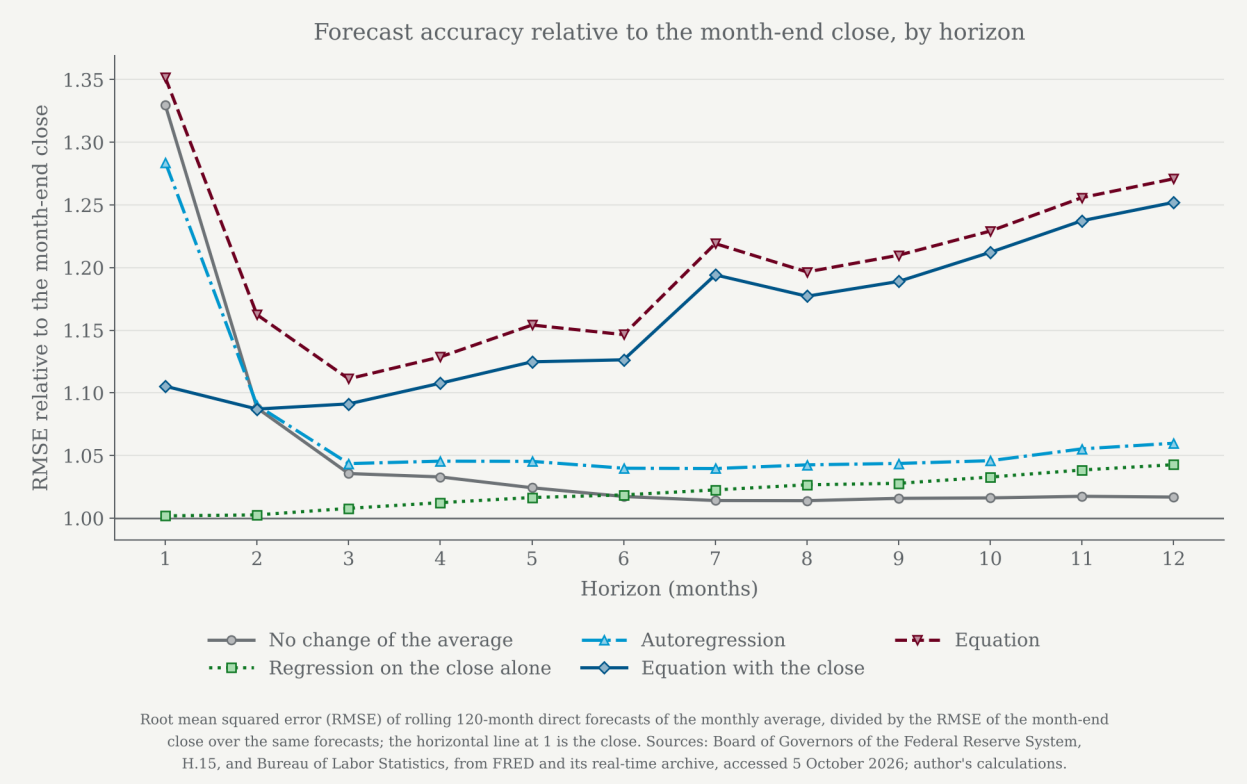


Figure 5. RMSE of each procedure divided by the RMSE of the month-end close (vertical axis), at horizons of one through twelve months (horizontal axis), over rolling-window forecasts of the monthly average of the 10-year yield. The horizontal line at 1 is the close. Solid gray line with circles: the no-change forecast of the average; dotted green line with squares: the regression on the close alone; dash-dotted bright blue line with triangles: the autoregression; solid deep blue line with diamonds: the equation with the close; dashed maroon line with inverted triangles: the equation. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, from FRED and its real-time archive, accessed 5 October 2026; author's calculations.

8.3 Comparison with Published Benchmarks

Table 6 sets the results beside those of the studies reviewed in Section 2. The published results are stated in words, because the studies differ in sample, maturity, data vintage, and design, and a ratio taken from one is not comparable with a ratio taken from another.

Table 6. Forecasts of long-term Treasury yields against the no-change forecast: published results and this paper. The upper block states each study's finding in words, as the study or, for the first row, a later study reports it; samples, maturities, data vintages, and evaluation designs differ across the rows, and the block is a guide to the literature and not a ranking. The lower block gives this paper's root mean squared error (RMSE) relative to the no-change forecast of the monthly average, on real-time data and rolling windows of 120 months; a ratio below 1 is a more accurate forecast than no change. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, each from FRED and its real-time archive, Federal Reserve Bank of St. Louis, accessed 5 October 2026; author's calculations.

Study	Procedure	Horizon	Result for long-term yields
Duffee	Affine term structure models	Several horizons	Less accurate than no change
Kolb and Stekler	Analysts' forecasts	Six months	Not more accurate than no change
Diebold and Li	Dynamic three-factor model	One month	Not distinguishable from no change
Diebold and Li	Dynamic three-factor model	Twelve months	More accurate than no change
Moench	Factor-augmented vector autoregression	Six to twelve months	More accurate than no change
Baghestani	Autoregression with expected inflation	Several horizons	More accurate than no change
Bauer and Rudebusch	Shifting trends in inflation and the real rate	Long horizons	More accurate than no change
This paper: equation	Monthly average, real-time data	One month	RMSE ratio 1.016
This paper: equation	Monthly average, real-time data	Twelve months	RMSE ratio 1.250
This paper: equation with the close	Monthly average, real-time data	One month	RMSE ratio 0.831
This paper: equation with the close	Monthly average, real-time data	Twelve months	RMSE ratio 1.231
This paper: month-end close	Monthly average, real-time data	One month	RMSE ratio 0.752
This paper: month-end close	Monthly average, real-time data	Twelve months	RMSE ratio 0.983

At one month, the equation with the close improves on the no-change forecast, where the three-factor model of Diebold and Li is not distinguishable from it. The improvement is not evidence of information in the model's macroeconomic terms. The target here is a monthly average, whose no-change benchmark is weak by the construction of Section 4.2, and the month-end close reaches a ratio of 0.752 against that benchmark with no estimation at all. At twelve months, where the factor models and the models with macroeconomic trends report their gains, both forms of the model have ratios above one.

8.4 The PPI Term

Table 7 isolates the PPI term. The Clark-West test of the equation against the same equation without the PPI term does not reject at any horizon (bootstrap p = 0.740 at one month and 0.653 at twelve), nor does the test with the close in both (bootstrap p = 0.856 and 0.754). Removing the PPI term lowers the equation's RMSE relative to the close from 1.352 to 1.340 at one month.

Table 7. The PPI term out of sample, rolling window of 120 months, by horizon. The upper block tests each equation against the same equation without the PPI term, the latest published change in the PPI rate. The lower block replaces that term with the latest published change in the PCE rate, both read from the vintage in force at the origin, on the origins from August 2010, the first whose rolling window lies within the archived PCE vintages; RMSE is in percentage points. Each p-value is the bootstrap value, with the asymptotic value in parentheses. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, each from FRED and its real-time archive, Federal Reserve Bank of St. Louis, accessed 5 October 2026; author's calculations. PCE: Bureau of Economic Analysis, from the same archive.

Horizon (months)	1	3	6	12
p-value of no gain from the PPI term (Clark-West, one-sided)				
In the equation	0.740 (0.781)	0.645 (0.677)	0.491 (0.481)	0.653 (0.742)
In the equation with the close	0.856 (0.889)	0.463 (0.315)	0.492 (0.475)	0.754 (0.820)
PPI rate against PCE rate, origins from August 2010				
The equation, PPI rate	0.202	0.426	0.647	0.953
The equation, PCE rate	0.202	0.423	0.649	0.982
Month-end close	0.159	0.388	0.564	0.820
Diebold-Mariano p-value, PPI against PCE	0.816 (0.840)	0.432 (0.449)	0.685 (0.746)	0.423 (0.388)
Forecasts in this block	193	189	183	171

Replacing the PPI rate with the PCE rate, on the origins from August 2010, gives an equation whose accuracy is not distinguishable from that of the equation with the PPI rate: at one month the two have RMSEs of 0.202 and 0.202 percentage points, against 0.159 for the close, and equal accuracy of the two is not rejected at any horizon. Neither equation, with the PPI rate or with the PCE rate, is as accurate as the close at any horizon.

8.5 Robustness

Table 8 divides the one-month record into three subperiods. The close has the lower RMSE in each, and the equation's RMSE relative to the close is 1.398, 1.337, and 1.273. Figure 6 shows the RMSE over trailing windows of 24 months, and Figure 7 the cumulative difference in squared errors from the close. Over the record the equation's cumulative squared error ends 9.15 squared percentage points above the close's, that of the no-change forecast of the average 8.50 above, and that of the equation with the close 2.45 above.

Table 8. One-month forecasts by subperiod of the rolling record. Columns: the number of forecasts; the RMSE of the month-end close and of the no-change forecast of the average, in percentage points; and the RMSE of the equation and of the equation with the close, each divided by the RMSE of the close. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, each from FRED and its real-time archive, Federal Reserve Bank of St. Louis, accessed 5 October 2026; author's calculations.

Origins	Forecasts	Close	No change	Equation	With close
February 1992 through August 2008	199	0.160	0.224	1.398	1.036
September 2008 through December 2019	136	0.159	0.205	1.337	1.098
January 2020 through August 2026	80	0.177	0.220	1.273	1.244

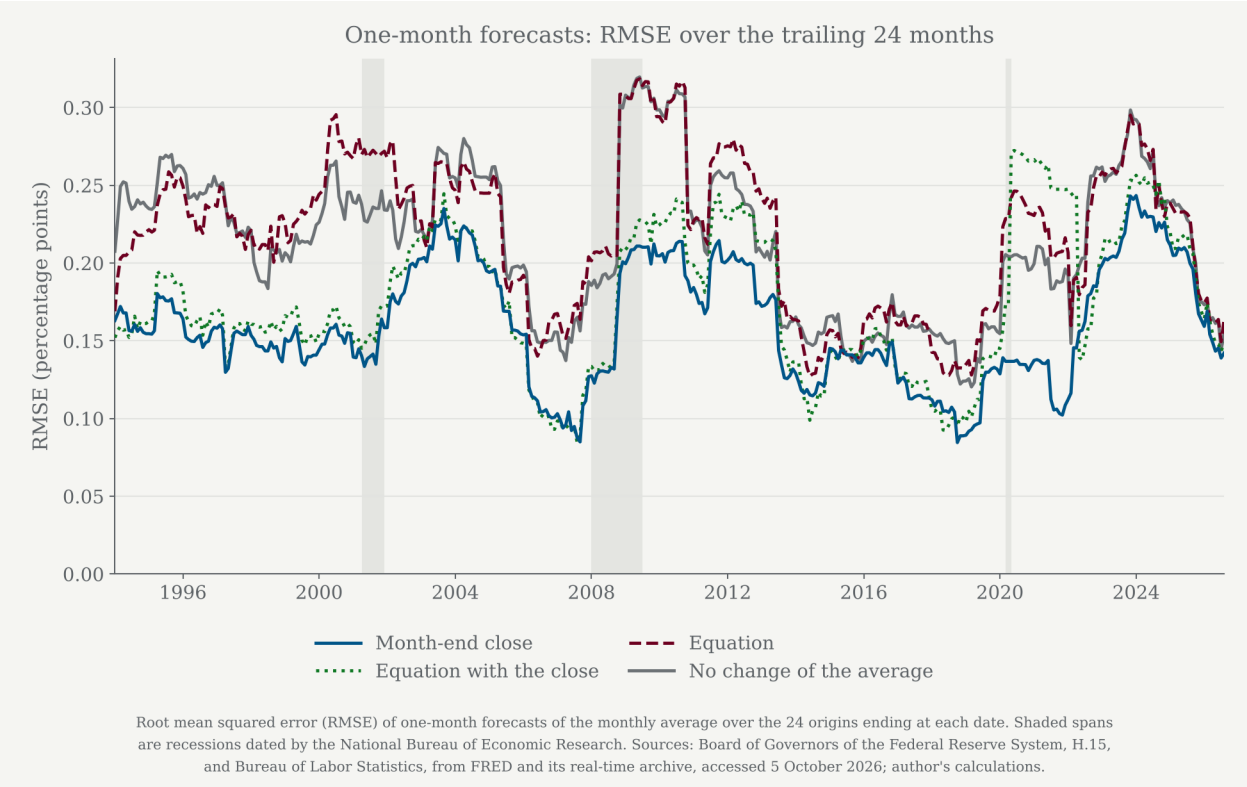


Figure 6. RMSE of one-month forecasts of the monthly average of the 10-year yield over the 24 origins ending at each date (horizontal axis), in percentage points (vertical axis). Solid deep blue line: the month-end close; dotted green line: the equation with the close; dashed maroon line: the equation; solid gray line: the no-change forecast of the average. Shaded spans are recessions dated by the National Bureau of Economic Research. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, from FRED and its real-time archive, accessed 5 October 2026; author's calculations.

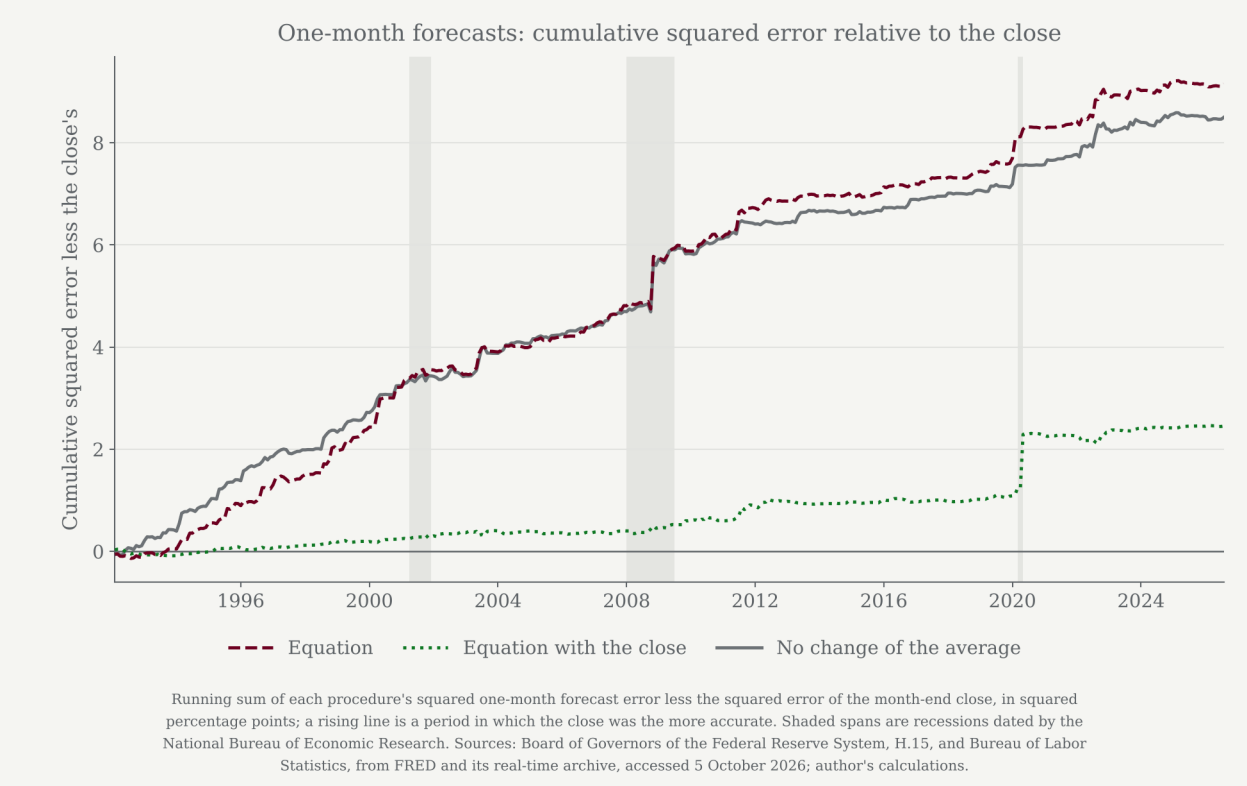


Figure 7. Running sum, over origins (horizontal axis), of each procedure's squared one-month forecast error less the squared error of the month-end close, in squared percentage points (vertical axis). A rising line is a period in which the close was the more accurate. Dashed maroon line: the equation; dotted green line: the equation with the close; solid gray line: the no-change forecast of the average. Shaded spans are recessions dated by the National Bureau of Economic Research. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, from FRED and its real-time archive, accessed 5 October 2026; author's calculations.

The plot of cumulative squared-error differences follows the practice of the return-predictability literature.[37] Figure 8 shows the coefficient on the PPI term in each rolling window of the equation with the close. It ranges from -0.031 through 0.056 and is positive in 186 of the 416 windows; its absolute t-statistic exceeds 1.96 in 13 of them.

[37] Ivo Welch and Amit Goyal, "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," *Review of Financial Studies* 21, no. 4 (2008): 1455–1508.

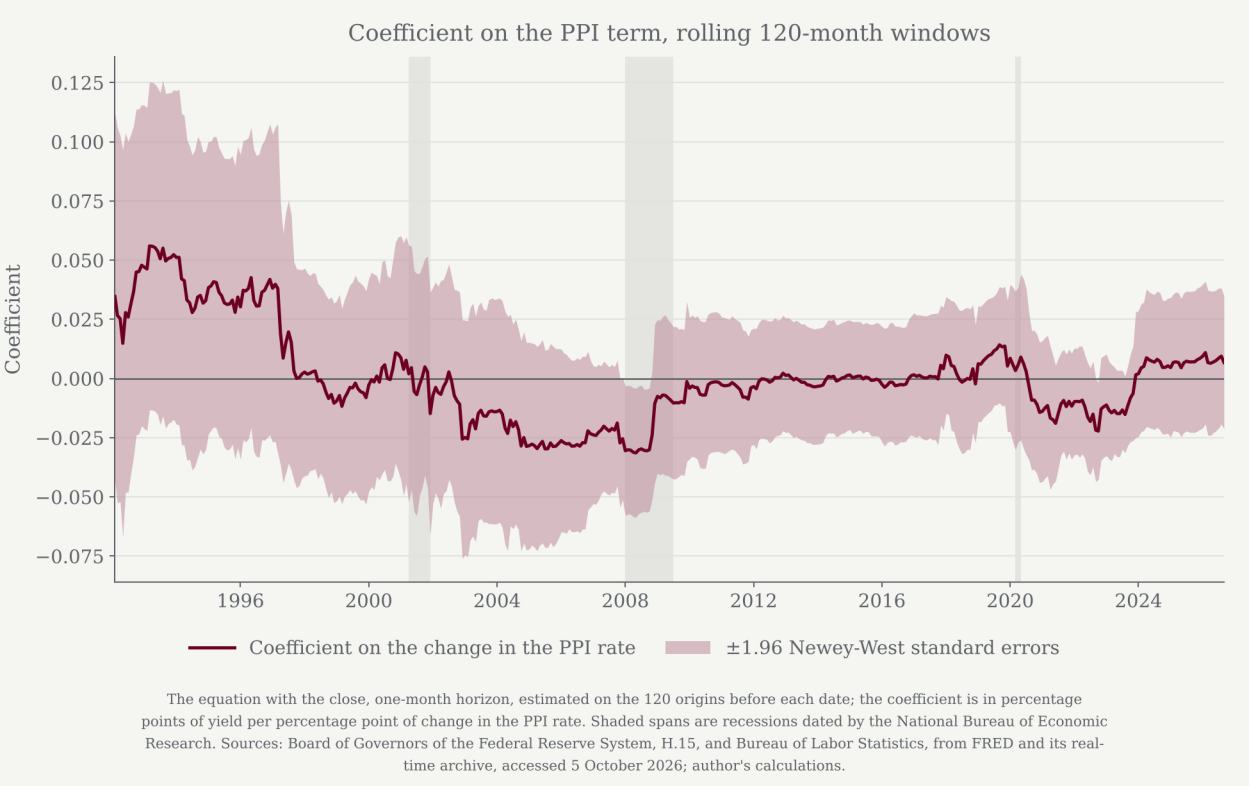


Figure 8. Coefficient on the latest published change in the PPI rate in the one-month equation with the close (solid maroon line, vertical axis, in percentage points of yield per percentage point of change in the PPI rate), estimated on the 120 origins before each date (horizontal axis). The band is ± 1.96 Newey-West standard errors; the horizontal line marks zero. Shaded spans are recessions dated by the National Bureau of Economic Research. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, from FRED and its real-time archive, accessed 5 October 2026; author's calculations.

Table 9 changes one element of the design at a time. With an expanding window in place of the rolling one, the equation's RMSE relative to the close is 1.317 at one month and 1.124 at twelve, and the equation with the close has 1.025 and 1.117. With latest-vintage data in place of real-time data the one-month ratio is 1.357, and with the PPI rate in level in place of its change it is 1.345.

Table 9. Alternative designs, by horizon. The upper block keeps the monthly-average target and changes one element of the design at a time: the window, the data vintage, or the form in which the PPI rate enters. The lower block changes the target to the yield at the close of the final business day of the target month, with the equation's rate terms sampled at month-end closes; its benchmark is no change of the close. A ratio above 1 is a less accurate forecast than the benchmark of its block. Each p-value is the bootstrap value, with the asymptotic value in parentheses. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, each from FRED and its real-time archive, Federal Reserve Bank of St. Louis, accessed 5 October 2026; author's calculations.

Horizon (months)	1	3	6	12
Monthly-average target: RMSE relative to the month-end close				
Equation, expanding window	1.317	1.050	1.088	1.124
Equation with the close, expanding window	1.025	1.028	1.091	1.117
Equation, latest-vintage data, rolling window	1.357	1.128	1.144	1.269
Equation, PPI rate in level, rolling window	1.345	1.116	1.182	1.292
Month-end target: RMSE relative to no change of the close				
Autoregression, rolling window	1.002	1.016	1.022	1.044
Equation, rolling window	1.063	1.102	1.133	1.243
Equation, expanding window	1.015	1.023	1.078	1.117
Clark-West p: PPI term, rolling window	0.759 (0.806)	0.459 (0.357)	0.820 (0.856)	0.895 (0.928)

The lower block changes the target to the yield at the close of the final business day of the target month. That target has no time averaging, and its benchmark is no change of the close. The equation's RMSE relative to that benchmark is 1.063 at one month and 1.243 at twelve, and its in-sample R^2 is 0.037. Under either target the equation does not improve on the no-change forecast of its target, that of the monthly average under the first and that of the close under the second. Under the month-end target the PPI term has a t-statistic of -0.13 in sample, and the Clark-West test of the equation against the same equation without it does not reject at any of the four horizons.

8.6 Caveats and Limitations

Four limitations bear on the reading of these results. The comparison of Table 6 is indicative: the cited studies differ from this one and from each other in sample, target, and design. The model is a reduced form and imposes no restriction from the absence of arbitrage, which limits its use for pricing. Its fit in sample comes mostly from the yield's own path, the lagged changes in the average and, in the equation with the close, the close less the average; no macroeconomic term is individually distinguishable from zero. And the first origin is set by the first archived vintage of the PPI, so the evaluation does not reach the years before February 1982.

9 The Inflation-Targeting Regimes

This section asks whether the result depends on the monetary regime. The supply-side hypothesis is conditional on a flat Phillips curve, and a flat curve has been attributed to inflation targeting,[38] so an estimate pooled over a sample that begins in February 1982 is not by itself a test of the hypothesis.

[38] McLeay and Tenreyro, "Optimal Inflation," 199–255.

Two dates are available from the record of the FOMC, and the paper uses both without choosing between them. At its meeting of 2–3 July 1996, the Committee discussed a numerical objective for inflation. The transcript records the Chairman's remark on the first day that the members had "now all agreed on 2 percent," and his request on the second day that the figure not be made public.[39] In January 2012 the Committee announced a longer-run goal of 2 percent.[40] The first date marks an implicit target and the second an official one. They divide the sample into three segments: 173 origins before July 1996, the origins between the two dates, and 177 origins from January 2012. Neither date was chosen by inspection of the data.

Table 10. The PPI term of the one-month equation by segment of the sample. Each row estimates the equation, and the equation with the close, on the origin-target pairs whose origin and target month both lie inside the segment, the long-run relation estimated inside the segment; the columns report the coefficient on the latest published change in the PPI rate, its Newey-West standard error, and its t-statistic. The lower block tests equality of every slope coefficient across each date; the statistic is chi-squared with 7 degrees of freedom for the equation and 8 with the close. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, each from FRED and its real-time archive, Federal Reserve Bank of St. Louis, accessed 5 October 2026; author's calculations.

		Equation			Equation with the close		
Segment	Pairs	Coefficient	Std. error	t	Coefficient	Std. error	t
Before July 1996	172	0.0171	0.0401	0.43	0.0262	0.0317	0.83
From July 1996, before January 2012	185	0.0065	0.0182	0.36	−0.0028	0.0126	−0.23
From January 2012	176	0.0010	0.0144	0.07	0.0041	0.0121	0.33
From July 1996	362	0.0039	0.0121	0.32	0.0009	0.0085	0.11
Equality of all slope coefficients across the date: Wald statistic and p-value							
Split at July 1996		15.33 (0.032)			18.98 (0.015)		
Split at January 2012		1.42 (0.985)			2.28 (0.971)		

Table 10 reports the coefficient on the PPI term of the one-month equation in each segment, with the long-run relation estimated inside the segment. Its t-statistic is 0.43 before July 1996, 0.36 between the two dates, and 0.07 from January 2012; with the close in the equation the three are 0.83, −0.23, and 0.33. The coefficient is not statistically distinguishable from zero in any segment under either equation. In the equation without the close, its point estimate is smallest in the official regime, where the hypothesis places its effect.

The equation does differ across the first date. Equality of all slope coefficients before and from July 1996 is rejected (p = 0.032 for the equation and 0.015 with the close), while equality before and from January 2012 is not (p = 0.985 and 0.971). The interaction of the PPI term with the split at July 1996 is not statistically distinguishable from zero in the equation (t = −0.31). The R^2 of the equation is 0.235 before July 1996 and 0.090 from it, and the slope of the long-run relation is 0.740 before and 0.554 from it. Two dates and three segments are several looks at one equation, and a rejection at one of two dates is read with that count in mind.

[39] Federal Open Market Committee, *Meeting of the Federal Open Market Committee, July 2–3, 1996*, transcript (Board of Governors of the Federal Reserve System), 63, https://www.federalreserve.gov/monetarypolicy/files/FOMC19960703meeting.pdf.

[40] Federal Open Market Committee, "Statement on Longer-Run Goals and Monetary Policy Strategy," adopted effective January 24, 2012 (Washington: Board of Governors of the Federal Reserve System, 2012).

Table 11. Out-of-sample accuracy inside each inflation-targeting regime, by horizon. Coefficients are estimated on origins of the regime alone, on an expanding window that starts at the regime date and holds at least 60 origin-target pairs. RMSE ratios divide by the RMSE of the month-end close over the same forecasts. The first Clark-West row tests the equation with the close against the regression on the close alone; the second tests it against the same equation without the PPI term. Each p-value is the bootstrap value, with the asymptotic value in parentheses. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, each from FRED and its real-time archive, Federal Reserve Bank of St. Louis, accessed 5 October 2026; author's calculations.

Horizon (months)	1	3	6	12
Estimated on origins from July 1996 alone				
RMSE relative to the close: equation	1.307	1.077	1.081	1.191
Same: equation with the close	1.036	1.042	1.053	1.162
Same: regression on the close alone	1.006	1.012	1.018	1.063
Clark-West p: terms beyond the close	0.018 (0.016)	0.010 (0.005)	0.021 (0.006)	0.058 (0.054)
Clark-West p: PPI term, with the close	0.528 (0.674)	0.482 (0.570)	0.199 (0.253)	0.899 (0.934)
Forecasts	302	298	292	280
Estimated on origins from January 2012 alone				
RMSE relative to the close: equation	1.379	1.182	1.242	1.243
Same: equation with the close	1.176	1.091	1.192	1.221
Same: regression on the close alone	0.995	1.003	1.011	1.051
Clark-West p: terms beyond the close	0.841 (0.886)	0.285 (0.284)	0.509 (0.568)	0.287 (0.208)
Clark-West p: PPI term, with the close	0.500 (0.467)	0.227 (0.164)	0.437 (0.360)	0.975 (0.963)
Forecasts	116	112	106	94

Table 11 repeats the out-of-sample evaluation with coefficients estimated on the regime's own origins, on an expanding window that starts at the regime date. From July 1996 this yields 302 one-month forecasts and from January 2012 116. In both regimes the equation and the equation with the close have a higher RMSE than the close at each of the four horizons: the equation's RMSE relative to the close at one month is 1.307 on data from July 1996 and 1.379 on data from January 2012. In either regime, the Clark-West test of the equation with the close against the same equation without the PPI term does not reject at any of the four horizons.

On data from July 1996, the equation with the close comes nearer to the close than in the pooled record, with an RMSE ratio of 1.036 at one month, and the Clark-West test of that equation against the regression on the close alone rejects at one, three, and six months (bootstrap p = 0.018, 0.010, and 0.021; at twelve months p = 0.058). The terms beyond the close therefore have content in that regime. No one of them is identified as its source: in the equation estimated on those origins the deviation from the long-run relation has the largest t-statistic in absolute value among them (-2.55), and its bootstrap p-value is 0.135; no other term of that equation is statistically distinguishable from zero in sample. The content does not make the estimated equation the more accurate forecast. On data from January 2012 the same test does not reject (bootstrap p = 0.841 at one month); that record has 94 overlapping twelve-month forecasts.

10 Conditional Forecast

This section reports the forecasts made at the close of 30 September 2026, conditional on what had been published by that day.

The yield closed September at 5.29 percent, 0.30 percentage points above the month's average of 4.99 percent. Table 12 and Figure 9 report the equation's forecast for each of the twelve following monthly averages, with intervals taken from the percentiles of the equation's own out-of-sample errors at each horizon, and beside it the forecast of the equation with the close and the close itself. The equation places the average of October 2026 at 5.15 percent, with a 95 percent interval of 4.71 through 5.54 percent, and the average of September 2027 at 5.54 percent, with a 68 percent interval of 4.13 through 6.27 and a 95 percent interval of 3.51 through 7.39 percent. The intervals widen with the horizon, as the forecast errors of a persistent series do.

Table 12. Conditional forecast of the monthly average of the 10-year yield made at the close of 30 September 2026, in percent. The equation's forecast is direct, from the rolling window ending at that origin; its intervals add the 16th and 84th, and the 2.5th and 97.5th, percentiles of the equation's own out-of-sample errors at that horizon to the forecast. The final two columns are the forecast of the equation with the close and the month-end close, which is that day's yield at every horizon. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, each from FRED and its real-time archive, Federal Reserve Bank of St. Louis, accessed 5 October 2026; author's calculations.

Target month	Equation	68% interval	95% interval	With the close	Month-end close
October 2026	5.15	[4.92, 5.33]	[4.71, 5.54]	5.38	5.29
November 2026	5.23	[4.82, 5.51]	[4.47, 5.91]	5.48	5.29
December 2026	5.29	[4.74, 5.64]	[4.28, 6.14]	5.47	5.29
January 2027	5.33	[4.63, 5.71]	[4.10, 6.26]	5.48	5.29
February 2027	5.36	[4.55, 5.81]	[3.98, 6.47]	5.56	5.29
March 2027	5.41	[4.50, 5.90]	[3.97, 6.49]	5.66	5.29
April 2027	5.45	[4.43, 5.97]	[3.91, 6.75]	5.78	5.29
May 2027	5.55	[4.47, 6.06]	[3.92, 6.85]	5.91	5.29
June 2027	5.59	[4.41, 6.10]	[3.82, 7.21]	5.97	5.29
July 2027	5.57	[4.23, 6.12]	[3.73, 7.21]	5.97	5.29
August 2027	5.52	[4.14, 6.13]	[3.60, 7.39]	5.90	5.29
September 2027	5.54	[4.13, 6.27]	[3.51, 7.39]	5.90	5.29

The forecast of the equation with the close for September 2027 is 5.90 percent, and the close carried forward is 5.29 percent. Both forms of the model place the average of September 2027 above the close. Section 8 governs how the three are weighed: at the twelve-month horizon their RMSEs over the evaluation record are 1.088 and 1.072 percentage points for the two forms of the model, against 0.856 for the close.



Figure 9. Monthly average of the 10-year yield from January 2022 (solid deep blue line) and forecasts of the twelve monthly averages after September 2026, in percent (vertical axis). Dashed maroon line: the equation's conditional forecast; the darker and lighter bands are its 68 and 95 percent intervals, from percentiles of the equation's own out-of-sample errors at each horizon. Dotted green line: the equation with the close. The marker is the yield at the close of the last origin, and the thin deep blue line continuing from it is the close carried forward. Sources: Board of Governors of the Federal Reserve System, H.15, and Bureau of Labor Statistics, from FRED and its real-time archive, accessed 5 October 2026; author's calculations.

11 Yield Decomposition

This section divides the yield into a real and an inflation component.

Figure 10 shows the 10-year yield with the yield on 10-year inflation-protected Treasury securities and breakeven inflation, the difference between the two, at month-end closes from January 2003. From January 2003 to the low of the nominal yield in July 2020, at 0.55 percent, the nominal yield changed by −3.45 percentage points; the inflation-protected yield changed by −3.19 and breakeven inflation by −0.26. The inflation-protected yield was below zero in 47 months, the first December 2011 and the last March 2022, and reached −1.16 percent in July 2021. From the low to the close of the last origin the nominal yield rose by 4.74 percentage points, of which 3.93 was the inflation-protected yield and 0.81 breakeven inflation. At that close the nominal yield of 5.29 percent divided into an inflation-protected yield of 2.93 and breakeven inflation of 2.36.

In both movements the inflation-protected yield accounts for the larger part of the change. The division is by market prices and not by a model; a term structure model would divide the yield differently, into expected short rates and a term premium.[41]

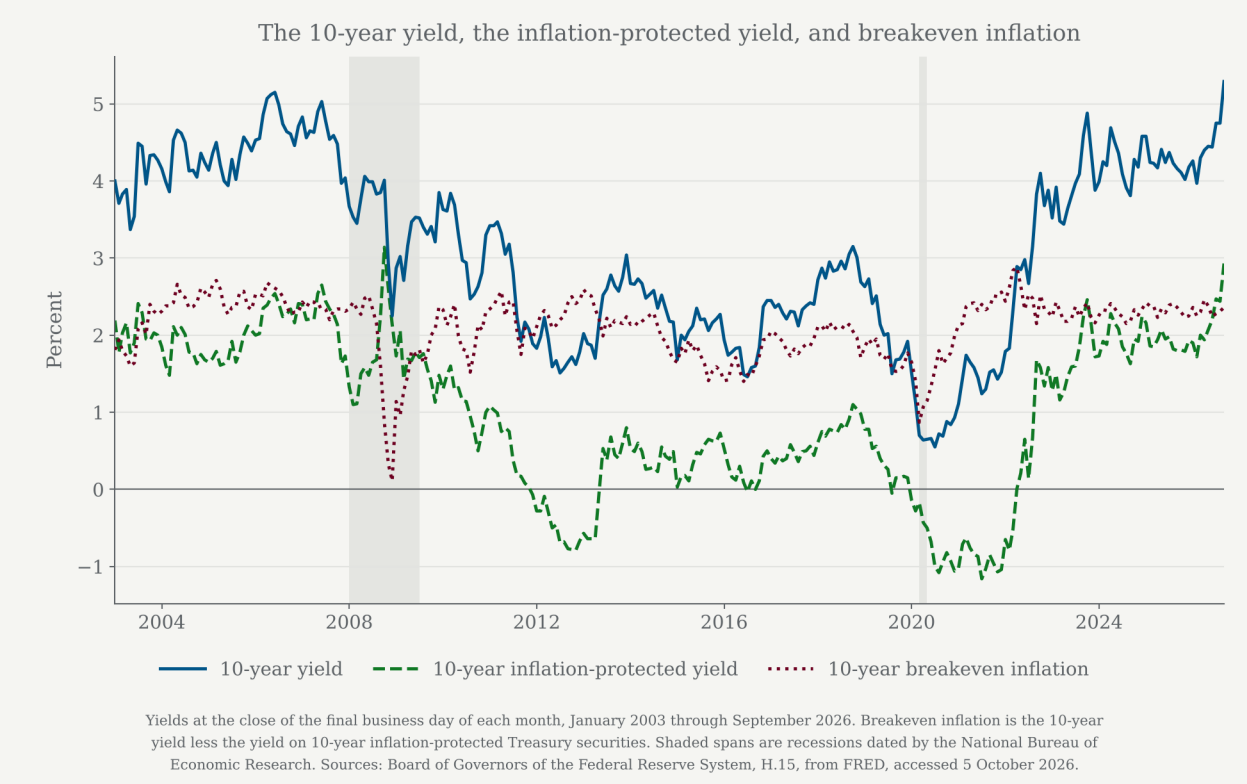


Figure 10. The 10-year Treasury yield (solid deep blue line), the yield on 10-year inflation-protected Treasury securities (dashed green line), and 10-year breakeven inflation (dotted maroon line), in percent (vertical axis), at the close of the final business day of each month, January 2003 through September 2026. Breakeven inflation is the 10-year yield less the inflation-protected yield. Shaded spans are recessions dated by the National Bureau of Economic Research. Sources: Board of Governors of the Federal Reserve System, H.15, from FRED, accessed 5 October 2026.

12 Connection to the Taylor Rule and Observational Equivalence

This section states what the findings do and do not imply for the companion papers.

The companion papers concern monetary policy rules and inflation: which accounts of the policy rate real-time data can separate, and whether supply-side indicators forecast consumer inflation when the Phillips curve is flat. The long-run relation of Section 5 ties the yield to the policy rate that the first of them studies, with a slope of 0.846. This paper concerns the yield. Its finding on the PPI term is stated out of sample and in sample. Out of sample, removing the term lowers the RMSE of the forecasts of the monthly averages of the 10-year yield at each of the four horizons, with and without the yield's own close in the equation; the Clark-West test of the equation against

[41] Adrian, Crump, and Moench, "Pricing the Term Structure," 110–138.

the same equation without the term does not reject at any horizon tested, and in either regime neither does that of the equation with the close. In sample, the term's coefficient is not statistically distinguishable from zero in the full sample or in any segment. That finding is compatible with a PPI that leads consumer inflation, under an account this paper does not test: a signal that forecasts inflation and is public can be priced by a bond market when it is published, and a signal already in the price does not forecast the price. The evaluation here does not test whether yields respond to PPI releases on the day of release; it tests whether the PPI term improves the forecasts of the monthly averages that follow.

Figure 11 sets the yield beside the GSCPI, another supply-side indicator that the companion paper on inflation evaluates. The index reached its highest value, 4.43 standard deviations above its mean, in December 2021, when the monthly average of the yield was 1.47 percent. One year later the yield averaged 3.62 percent, and two years later 4.02 percent. The GSCPI enters no forecast of this paper; the out-of-sample test of a supply-side signal here is that of the PPI term in Sections 8.4, 8.5, and 9.

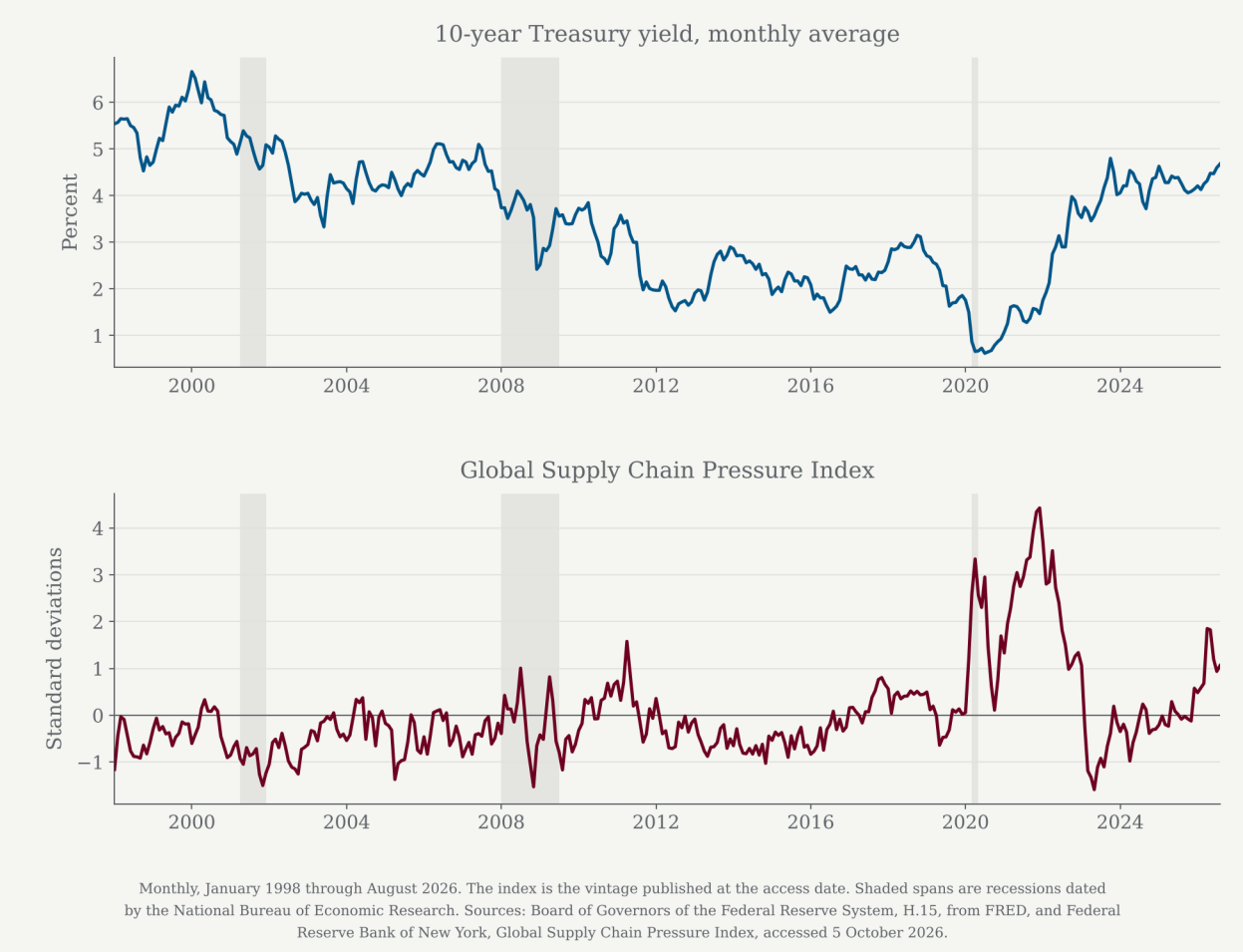


Figure 11. Upper panel: monthly average of the 10-year Treasury yield, in percent. Lower panel: the GSCPI, in standard deviations from its mean, as published at the access date. Monthly, January 1998 through August 2026. Shaded spans are recessions dated by the National Bureau of Economic Research. Sources: Board of Governors of the Federal Reserve System, H.15, from FRED; Federal Reserve Bank of New York; each accessed 5 October 2026.

On the origins from August 2010, replacing the PPI rate with the PCE rate leaves the equation's accuracy not distinguishable from that of the equation with the PPI rate at any horizon (Table 7). What content the model has beyond the close is found on data from July 1996, in its terms taken jointly. The paper does not attribute that content to the long-run relation between the yield and the funds rate: the coefficient on the deviation from that relation is not statistically distinguishable from zero under the bootstrap.

13 Conclusion

This paper developed a forecasting model for the 10-year Treasury yield, an error correction equation with a supply-side inflation signal, and evaluated it in real time. Given the yield at the month-end close, the model forecasts the next month's average with an RMSE 0.831 times that of the no-change forecast and with the sign of the change in 67.6 percent of months. Beyond one month it does not improve on no change, and it does not forecast the monthly average of the yield more accurately than the yield at the month-end close at any horizon from one month through twelve under a rolling window, nor at the four horizons tested under an expanding window or on latest-vintage data.

The model's fit comes from the yield's own dynamics. The coefficient on the PPI term is not statistically distinguishable from zero in the full sample or in any of three segments defined by the implicit and the official adoption of an inflation target, and the Clark-West test of the equation with the close against the same equation without the term does not reject in either regime. In the full sample the coefficients on the funds rate, the spread, and the unemployment rate are not distinguishable from zero either, nor, on the origins from August 2000, that on the PCE term in the PPI term's place.

One finding qualifies the reading of forecast comparisons for averaged interest rates. A monthly average is predictable from its own last change by construction, with a first-order autocorrelation of 0.336 in this sample, so an equation that includes that change improves on the no-change forecast of the average without information beyond the yield's own path; the month-end close already carries that predictability, so against the close the improvement does not appear. Separately, equality of the equation's slope coefficients before and from July 1996 is rejected (p = 0.032), at one of the two dates tested, which bears on estimates of this equation that pool the two periods.

The model is not without content beyond the close. On data from July 1996 the Clark-West test of the equation with the close against the regression on the close alone rejects at horizons of one, three, and six months under bootstrap reference distributions, so that its terms beyond the close jointly carry information beyond the regression on the close alone; on data from January 2012 it does not reject at any horizon. No single term is identified as its source: the coefficient on the deviation of the yield from its long-run relation with the funds rate, on which the error correction form rests, is not statistically distinguishable from zero once its t-statistic is read against the distribution it has when cointegration is not established. The equation that carries those terms has the higher out-of-sample RMSE all the same, 1.036 times the close's at one month on those data. At the last origin the close stood at 5.29 percent; the model's conditional forecast of the monthly average twelve months ahead is 5.54 percent, with a 95 percent interval of 3.51 through 7.39.

Bibliography

Adrian, Tobias, Richard K. Crump, and Emanuel Moench. "Pricing the Term Structure with Linear Regressions." *Journal of Financial Economics* 110, no. 1 (2013): 110–138.

Ang, Andrew, and Monika Piazzesi. "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables." *Journal of Monetary Economics* 50, no. 4 (2003): 745–787.

Baghestani, Hamid. "Forecasting the 10-Year US Treasury Rate." *Journal of Forecasting* 29, no. 8 (2010): 673–688.

Banerjee, Anindya, Juan J. Dolado, and Ricardo Mestre. "Error-Correction Mechanism Tests for Cointegration in a Single-Equation Framework." *Journal of Time Series Analysis* 19, no. 3 (1998): 267–283.

Bauer, Michael D., and Glenn D. Rudebusch. "Interest Rates under Falling Stars." *American Economic Review* 110, no. 5 (2020): 1316–1354.

Benigno, Gianluca, Julian di Giovanni, Jan J. J. Groen, and Adam I. Noble. "The GSCPI: A New Barometer of Global Supply Chain Pressures." Staff Report no. 1017. Federal Reserve Bank of New York, May 2022.

Breusch, T. S. "Testing for Autocorrelation in Dynamic Linear Models." *Australian Economic Papers* 17, no. 31 (1978): 334–355.

Campbell, John Y., and Robert J. Shiller. "Yield Spreads and Interest Rate Movements: A Bird's Eye View." *Review of Economic Studies* 58, no. 3 (1991): 495–514.

Christensen, Jens H. E., Francis X. Diebold, and Glenn D. Rudebusch. "The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models." *Journal of Econometrics* 164, no. 1 (2011): 4–20.

Cieslak, Anna. "Short-Rate Expectations and Unexpected Returns in Treasury Bonds." *Review of Financial Studies* 31, no. 9 (2018): 3265–3306.

Clark, Todd E., and Michael W. McCracken. "Reality Checks and Comparisons of Nested Predictive Models." *Journal of Business & Economic Statistics* 30, no. 1 (2012): 53–66.

Clark, Todd E., and Kenneth D. West. "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models." *Journal of Econometrics* 138, no. 1 (2007): 291–311.

Croushore, Dean, and Tom Stark. "A Real-Time Data Set for Macroeconomists." *Journal of Econometrics* 105, no. 1 (2001): 111–130.

Diebold, Francis X., and Canlin Li. "Forecasting the Term Structure of Government Bond Yields." *Journal of Econometrics* 130, no. 2 (2006): 337–364.

Diebold, Francis X., and Roberto S. Mariano. "Comparing Predictive Accuracy." *Journal of Business & Economic Statistics* 13, no. 3 (1995): 253–263.

Diebold, Francis X., Glenn D. Rudebusch, and S. Boragan Aruoba. "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach." *Journal of Econometrics* 131, nos. 1–2 (2006): 309–338.

Duffee, Gregory R. "Term Premia and Interest Rate Forecasts in Affine Models." *Journal of Finance* 57, no. 1 (2002): 405–443.

Ellwanger, Reinhard, and Stephen Snudden. "Forecasts of the Real Price of Oil Revisited: Do They Beat the Random Walk?" *Journal of Banking & Finance* 154 (2023): 106962.

Engle, Robert F. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50, no. 4 (1982): 987–1007.

Engle, Robert F., and C. W. J. Granger. "Co-Integration and Error Correction: Representation, Estimation, and Testing." *Econometrica* 55, no. 2 (1987): 251–276.

Federal Open Market Committee. *Meeting of the Federal Open Market Committee, July 2–3, 1996*. Transcript. Board of Governors of the Federal Reserve System. https://www.federalreserve.gov/monetarypolicy/files/FOMC19960703meeting.pdf.

Federal Open Market Committee. "Statement on Longer-Run Goals and Monetary Policy Strategy." Adopted effective January 24, 2012. Washington: Board of Governors of the Federal Reserve System, 2012.

Godfrey, L. G. "Testing against General Autoregressive and Moving Average Error Models When the Regressors Include Lagged Dependent Variables." *Econometrica* 46, no. 6 (1978): 1293–1301.

Gonçalves, Sílvia, and Lutz Kilian. "Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form." *Journal of Econometrics* 123, no. 1 (2004): 89–120.

Hannan, E. J., and J. Rissanen. "Recursive Estimation of Mixed Autoregressive-Moving Average Order." *Biometrika* 69, no. 1 (1982): 81–94.

Harvey, David, Stephen Leybourne, and Paul Newbold. "Testing the Equality of Prediction Mean Squared Errors." *International Journal of Forecasting* 13, no. 2 (1997): 281–291.

Hazell, Jonathon, Juan Herreño, Emi Nakamura, and Jón Steinsson. "The Slope of the Phillips Curve: Evidence from U.S. States." *Quarterly Journal of Economics* 137, no. 3 (2022): 1299–1344.

Kolb, R. A., and H. O. Stekler. "How Well Do Analysts Forecast Interest Rates?" *Journal of Forecasting* 15, no. 5 (1996): 385–394.

Künsch, Hans R. "The Jackknife and the Bootstrap for General Stationary Observations." *The Annals of Statistics* 17, no. 3 (1989): 1217–1241.

Lo, Andrew W., and A. Craig MacKinlay. "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test." *Review of Financial Studies* 1, no. 1 (1988): 41–66.

McLeay, Michael, and Silvana Tenreyro. "Optimal Inflation and the Identification of the Phillips Curve." *NBER Macroeconomics Annual* 34 (2020): 199–255.

Moench, Emanuel. "Forecasting the Yield Curve in a Data-Rich Environment: A No-Arbitrage Factor-Augmented VAR Approach." *Journal of Econometrics* 146, no. 1 (2008): 26–43.

Newey, Whitney K., and Kenneth D. West. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55, no. 3 (1987): 703–708.

Politis, Dimitris N., and Joseph P. Romano. "A Circular Block-Resampling Procedure for Stationary Data." In *Exploring the Limits of Bootstrap*, edited by Raoul LePage and Lynne Billard, 263–270. New York: Wiley, 1992.

Said, Said E., and David A. Dickey. "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order." *Biometrika* 71, no. 3 (1984): 599–607.

Schmid, Frank. *Observational Equivalence in Taylor Rule Estimation: What Is Identified Depends on the Output Gap and the Sample Period*. Working paper, 26 July 2026.

Schmid, Frank. *Supply-Side Inflation Forecasting When the Phillips Curve Is Flat: A Real-Time Data Approach for the U.S. Inflation-Targeting Era, 2012:Q1–2025:Q4*. Working paper, 19 March 2026.

Schwarz, Gideon. "Estimating the Dimension of a Model." *The Annals of Statistics* 6, no. 2 (1978): 461–464.

Wang, Y.-F. "Comparing Machine Learning Models for Short-Term U.S. Treasury Yield Forecasting." *Concurrency and Computation: Practice and Experience* (2025): e70265.

Welch, Ivo, and Amit Goyal. "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction." *Review of Financial Studies* 21, no. 4 (2008): 1455–1508.

Working, Holbrook. "Note on the Correlation of First Differences of Averages in a Random Chain." *Econometrica* 28, no. 4 (1960): 916–918.

Data Sources

Board of Governors of the Federal Reserve System, H.15 Selected Interest Rates: the 10-year and 2-year constant-maturity Treasury yields, daily and monthly; the effective federal funds rate; the yield on 10-year inflation-protected Treasury securities. Bureau of Labor Statistics: the Producer Price Index for finished goods and the unemployment rate, each in every vintage of the real-time archive of FRED, Federal Reserve Bank of St. Louis. Bureau of Economic Analysis: the price index of personal consumption expenditures, in the same archive. Federal Reserve Bank of Philadelphia, Real-Time Data Set for Macroeconomists: monthly vintages of the Producer Price Index for finished goods. Federal Reserve Bank of New York: the Global Supply Chain Pressure Index. Federal Reserve Bank of St. Louis, FRED: 10-year breakeven inflation. National Bureau of Economic Research: recession dates. All series were accessed on 5 October 2026.