Federal Reserve Bank of Cleveland Working Paper Series

Low Interest Rates and the Predictive Content of the Yield Curve

Michael D. Bordo and Joseph G. Haubrich

Working Paper No. 20-24R

December 2021

Suggested citation: Bordo, Michael D., and Joseph G. Haubrich. 2021. "Low Interest Rates and the Predictive Content of the Yield Curve." Working Paper No. 20-24R. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202024r.

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

Low Interest Rates and the Predictive Content of the Yield Curve

By Michael D. Bordo and Joseph G. Haubrich*

Does the yield curve's ability to predict future output and recessions differ when interest rates and inflation are low, as in the current global environment? We explore the issue using historical data going back to the 19th century for the US. This paper is similar in spirit to Ramey and Zubairy (2018) who look, at the government spending multiplier in times of low interest rates. If anything, the yield curve tends to predict output growth better in low interest rate environments, though this result is stronger for RGDP than for IP.
JEL: E32, N10, G01
Keywords: Low Interest Rates, Policy, and the Predictive Content of the Yield Curve

* Bordo: Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, NBER and Hoover Institution, michael.bordo@gmail.com. Haubrich: Federal Reserve Bank of Cleveland, PO Box 6387, Cleveland, OH 44101-1387, 216 579 2802 jhaubrich@clev.frb.org. The views expressed here are solely those of the authors and not necessarily those of the Federal Reserve Bank of Cleveland or the Board of Governors of the Federal Reserve System. We thank George Nurisso and Rachel Widra for research assistance, and Todd Clark, Andy Filardo, and Andrew Martinez for valuable comments, and we give special thanks to Florens Odendahl, Barbara Rossi, and Tatevik Sekhposyan for sharing their MATLAB program.

I. Introduction

Does the yield curve's ability to predict future output and recessions differ when interest rates are low, as in the current global environment? Despite a variety of work that examines the predictive content of the yield spread accounting for the level of rates (Ang, Piazzesi, and Wei, 2006), monetary policy (Cooper, Fuhrer, and Olivei, 2020), and the persistence of inflation (Benati and Goodhart, 2008), this question remains open. Using historical data on the low inflation, low interest rate environments of the late 1800s, the 1930s and the 1940s, and the period following the 2007-2008 financial crisis, a series of predictive regressions and encompassing tests shows that the yield curve contains information about future IP and real GDP in low rate environments, though the evidence is stronger for the quarterly data.

The yield curve inversion in the summer and fall of 2019 gave the matter renewed urgency. As it turns out, this inversion did foreshadow a recession in 2020, though some may think that the advent of COVID-19 made this example a coincidence. In fact, the monetary policy response to the pandemic makes the question particularly relevant, with expectations of an extended period of rates at the effective lower bound, quantitative easing, and even possible yield curve control (Belz and Wessel, 2020). We bring historical evidence to bear on this question, and find that the yield curve often, but not always, can predict output in three low interest rate environments since the 1870s. Though using very different techniques and asking different questions, this paper is similar in spirit to Ramey and Zubairy (2018), who look at the government spending multiplier in times of low interest rates.

In our data set, the US has had three extended periods of low inflation and low interest rates: during the secular price decline of 1870–1897, in the 1930s and 1940s, and in the years during and following the Great Recession. (the post-pandemic period, though undoubtedly one of low rates, does not yet count as extended, in our mind). These episodes stand out in Figure 1, which plots the two short rates we use: the commercial paper rate and the target fed funds rate. Ramey and Zubairy (2018) look for episodes when interest rates are near the zero lower bound or times of "extended monetary accommodation," which they define as either very low rates or when rates "stay constant rather than follow the Taylor rule." They find that this happens in two periods: 1932Q2–1951Q1 and 2008Q4–2015Q4. We compare the predictive content of the yield curve in and out of these low rate periods, and although the exact results depend on the time period and data set used, in general the yield curve shows predictive power even in low rate environments.

This paper contributes to the literature in several ways. First, it contributes to our understanding of how the yield curve predicts output in low interest rate environments. Fendael, Mai, and Mohr (2019) directly address the issue; relative to them, we look at future output, instead of recessions; we look at the US, instead of the EU; and we use data from 1876–2021, instead of 2001–2017. Concentrating

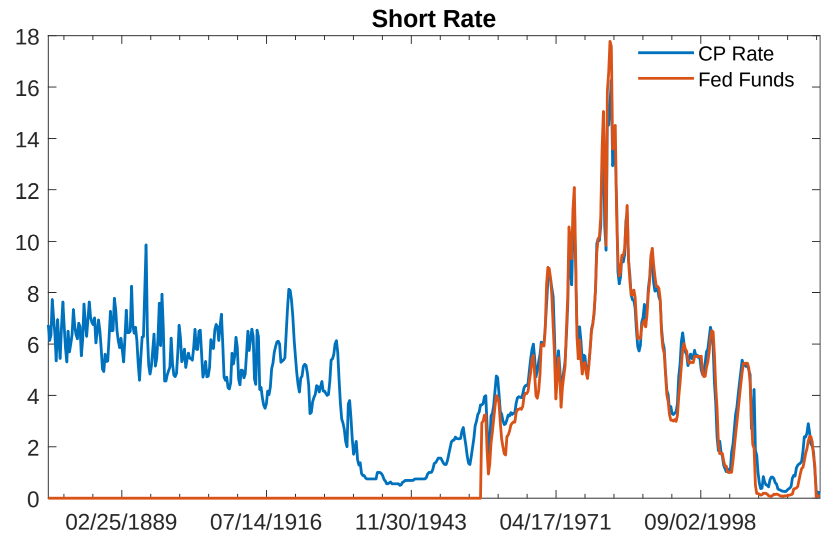

Figure 1. Short Rates

on low rates is in turn a special case of whether the level or the slope of the yield curve predicts output. Plosser and Rouwenhorst (1994) find that the slope of the term structure has information beyond short-term rates, as did Estrella and Mishkin (1997). Ang, Piazzesi, and Wei (2006) find that the predictive content of the spread resides in the short rate, but Bordo and Haubrich (2008b) add in a short rate separately and find that the yield curve still matters. Liu and Moench (2016) find that in-sample, adding the short rate does not help predict recessions, but out-of-sample, it can. Bauer and Mertens (2018) include the natural level of real interest rates, r^*, but find it has little impact on predictions. In turn, considering low rate periods is related to the question of how monetary regimes affect the predictive ability of the yield curve: Bordo and Haubrich (2008a,b) have suggested that the monetary regime may play a role, Giacomini and Rossi (2006) provide evidence that changes in monetary policy have led to breakdowns in predictive accuracy, and Benati and Goodhart (2008) show that changing inflation persistence matters, a finding that has extensive overlap with the first question, as short rates are often used as an indicator of monetary policy. These are all part of the rich literature that documents and explains the predictive ability of the yield curve; see the surveys by Estrella (2005), Wheelock and Wohar (2009) and Haubrich (2021). As such it also contributes to understanding the economics of low interest rate environments, as in Ramey and Zubairy (2018), Sims and Wu (2020), and Eggertsson and Woodford (2003).

The rest of the paper is structured as follows: Section II describes the empirical techniques used, and Section III describes the data. Section VII describes the results, and Section V includes robustness tests and Section VI concludes.

II. Empirical Strategy

We test for predictive content by running a series of in-sample and pseudo-out-of-sample predictive regressions, using quarterly real GDP and monthly industrial production as the output variables. The literature has used both approaches, and we can do no better than to quote Rossi (forthcoming) on the comparative advantages of each:

> Indeed, evaluating models' performance in-sample faces the risks of over-fitting, data snooping and lack of robustness to the presence of instabilities: evaluating models in terms of their out-of-sample forecasting ability helps alleviate these problems. On the other hand, Diebold (2015) argues in favor of a more cautionary use of forecast tests: while they are appropriate for evaluating forecasts, they may not always be the best option for evaluating specific features of a model, as they may involve a loss of power

In our case, in-sample has the added advantage that we can use the early deflationary period at the start of our data. In the pseudo-out-of-sample or rolling regression results, the initial estimation period covers most of the initial deflationary period, which is lost to the out-of-sample results.

Historically, the literature has used both approaches: Estrella and Hardouvelis (1991) use both in-sample and (pseudo) out-of-sample regressions, as did Harvey (1988) and Chinn and Kucko (2015). Hamilton and Kim (2002), Rudebusch, Sack and Swanson (2007), Benati and Goodhart (2008) and McMillan (2021) all use in-sample techniques (although they are econometrically sophisticated along other dimensions).

A. In-Sample

In-sample, we start with a regression of future growth against past growth, the yield curve, and a short rate. This provides a baseline for the predictive ability of the yield curve, and is particularly important in our case, because to create historically consistent series, we end up using somewhat non-standard term spreads (these particular spreads have appeared in the literature, but not as commonly as the 10-year three-month Treasury spread). Equation (1) describes the in-sample basic predictive regression, where y_t denotes the measure of output growth, *Spread* stands for the particular yield spread used and *Short* denotes the short-term interest rate used, with k denoting the lags used.

$$y_t = \alpha + \beta y_{t-k} + \gamma Spread_{t-k} + \delta Short_{t-k} + \epsilon_t \tag{1}$$

The lagged output is to see if the specification can at least do better than a random walk alternative. The short rate aims to control for the possible impact of monetary policy, and for the possibility that the predictive ability of the yield curve lies in its level, not its slope. We then account for the effects of low interest rate environments in two ways. First, we add a dummy δ_L for low interest rate periods and an interaction between the low rate dummy and the spread. The dummy and interaction terms are to directly test if things look different in a low rate environment, allowing for changes in both the level and the slope. Equation (2) shows this.

$$(2) \qquad y_t = \alpha + \beta y_{t-k} + \gamma Spread_{t-k} + \delta Short_{t-k} + \delta_L + \delta_L \times Spread_{t-k} + \epsilon_t$$

Second, we use a set of in-sample regressions to split the sample into periods of high rates and low rates. This has the advantage of allowing the coefficients on output and the short rate to vary across time, and does not impose the same coefficients on the yield curve variables across samples.

B. Pseudo Out-of-Sample

In-sample regressions use all the available data, but can be misleading if the underlying relationship shifts over time, as technology, financial practices, and regulations develop. A standard alternative is a rolling regression, where the sample changes over time. Thoma (1994) champions the forward expanding window (though he calls it a rolling window) which expands the window by adding successive observations, making the window larger as later observations are included. We prefer the rolling window, used by Swanson (1998), which keeps the window the same size by dropping observations at the beginning as later observations are added (hence the term rolling).

We can also formally test for the impact of low rate environments using tests from a recent paper by Odendahl, Rossi, and Sekhposyan, (2021). That paper presents several tests of forecast ability with state dependence, whereas in our case the state dependence is on the level of the short-term interest rate. Broadly speaking, their procedure first compares two forecasts to determine if a challenger forecast is more accurate than a base forecast. If it is, the procedure then fits a non-linear model to see if the forecast performance is a function of a conditioning variable, with the specification usually being some sort of threshold model, with the threshold endogenously determined. In our case, we compare the rolling out-of-sample predictions from equation (1) with the predictions from

$$(3) \qquad y_t = \alpha + \beta y_{t-k} + \gamma Spread_{t-k} + +\delta_L + \delta_L \times Spread_{t-k} + \epsilon_t$$

Letting $f^{(1)}_{t=h|t,1}$ denote the first or base forecast, and $f^{(2)}_{t=h|t,1}$ denote the challenger forecasts, then the forecast errors are

$$\epsilon_{t=h|t,1} = Y_{t+h} - f^{(1)}_{t=h|t,1} \tag{4}$$

$$\epsilon_{t=h|t,2} = Y_{t+h} - f^{(2)}_{t=h|t,1}, \tag{5}$$

Using squared errors as the loss function, $L_{t=h|t,i} = \epsilon^2_{t+h|t,i}$ for $i = 1, 2$ allows us to then define the loss differential as $\mathcal{L}_t = \Delta L_{t+h|t} = L_{t+h|t,1} - L_{t+h|t,2}$. The "local" or state-contingent forecasting performance is estimated as a non-linear model of the loss differential, in this case a threshold model of the difference in squared errors between the two models.

$$E\Delta(L_{t+h|t}|S_t) = \alpha + \theta\mathbb{1}(S_t > \gamma). \tag{6}$$

A positive and significant value for α indicates that the second model has better forecast ability than the first.[1]. If the predictive ability of the second forecast is sensitive to the state S_t, then θ will be significant and large relative to α.

Just because one model forecasts more accurately does not mean that the second model has no additional information. Tests for encompassing look at whether the competing forecast contains any useful information beyond the preferred forecast (e.g., whether the preferred forecast "encompasses" the competing forecast). Rejecting encompassing means there is useful information in the competing forecast. We look at whether equation (1) encompasses equation (2), that is, if a regression with lagged output and the short rate encompasses a regression that also includes the yield curve. Encompassing is a stricter test than predictive accuracy. Reformulating $\mathcal{L}_t$ as

$$\mathcal{L}_t = \epsilon^2_{t+h|t,1} - \epsilon_{t+h|t,1}\epsilon_{t+h|t,2} \tag{7}$$

allows the ORS procedure to test for encompassing.

Formal comparisons of the yield curves' predictive accuracy have been around for a while, with Dotsey (1998), Moneta (2005), and of course Chinn and Kucko (2015), among others, using the Diebold-Mariano test. Tests of encompassing are more rare, but were used effectively by Giacomini and Rossi (2006) in their study.

III. Data

We look at the question using both quarterly real GDP data and monthly industrial production data. GDP is the more comprehensive measure of output and the data go back further in time. IP is noisier, more variable, and its relationship to GDP varies over time. The GDP component that most closely matches

[1] A positive intercept indicates that the average differential is positive, which means the base forecast has higher squared errors on average.

IP (BEAhttps://www.bea.gov/help/faq/73) has decreased from 51.8 percent of GDP in 1947 to 29.5 percent in 2020 (FRED (A353RC1Q027SBEA)/GDP). IP has the advantage of being monthly, however.

Quarterly data. — For quarterly data, we produce a reasonably consistent data set that extends into the late 19th century. Looking at early data provides a richer set of monetary regimes, including the gold standard and a time without a central bank, but it also presents challenges. We start with the quarterly data set of Balke and Gordon (1986), which covers 1875-1983, extended with the data from Ramey and Zubairy (2018), which we further extend to 2021. For output we use quarterly year-over-year real GNP/GDP growth, 1876:Q1–2021:Q1. A major challenge is that a risk-free Treasury yield curve is not reliably available, both because Treasury bills did not become standard until the 1920s, and because it was not obvious that even longer-term Treasuries were the risk-free benchmark—for most of the 19th century, railroad bonds had at least as great a claim. Following Balke and Gordon (1986), we use the spread difference between the Baa corporate bond rate and the commercial paper rate. Though neither yield is risk free, it is hoped that their risk premiums will be roughly comparable; for more discussion see Bordo and Haubrich (2008a). Figure 2 plots the Baa-cp and the 10-year 3-month spreads, and for the half century in which they overlap, the correlation between them is 90 percent. Over this period there is not an obvious interest rate that can count as the monetary policy rate; indeed, for much of the period, the United States did not have a central bank. We use the short rate as a rough equivalent, as a proxy for the stance of conditions in the money market.

See the data appendix for additional details on construction.

For periods of low rates, we start with the period of secular price decline, which Bordo, Landon-Lane and Redish (2009) put as 1870–1896, and we extent the period slightly to 1897:Q2, when, according to our data, the secular price decline ends, though we note there is some disagreement about inflationary expectations after greenbacks became convertible in 1879 (Calomiris, 1993). For later periods, we use the definition of Ramey-Zubairy (2018), 1932:Q2 to 1951:Q1 and 2008:Q4 to 2015:Q4, translated for monthly data as 1932 April to 1951 March, and 2008 October to 2015 December. The FOMC increased the lower bound of the target federal funds range above zero on December 16, 2015, so we judge the Ramey-Zubairy ending date as appropriate. Post-COVID, the FOMC again lowered rates to the zero lower bound in April 2020, which remained in effect as of November 2021. These periods are clearly visible in Figure 1, which plots the commercial paper rate and the target fed funds rate. The US experience with yield curve control lasted from 1942 to 1951, with caps on Treasury bills ending in July 1947 and caps on longer rates expiring with the Treasury-Fed Accord of February 1951 (Garbade, 2020). Thus the yield curve control period is contained in the period of low rates, and we cannot distinguish between their effects.

Monthly data. — Monthly industrial production data begin in 1919, and that forms the start of our data, and we use year-over year growth. To provide a consistent series at the short end of the yield curve, we use a monetary policy rate. This differs from the 3-month or two-year Treasury yield most commonly used in post-WWII data, but it has its precedents: the early paper of Laurent (1988) looked at the 20-year Treasury FFR spread, Stock and Watson (2003) looked at a long-overnight spread, and in 2005 the Conference Board added the difference between the 10-year Treasury note and the fed funds rate to its index of leading economic indicators (Zarnowitz and Lee, 2005). For the policy rate, we initially use the discount rate from the Federal Reserve Bank of New York, available since 1914. Before the New Deal reforms, discount rates at Federal Reserve District banks varied, but New York, as the financial center of the nation, is the most obvious single rate to choose. In February of 1963, we start using the effective federal funds rate (EFFR), which in that month coincidentally(?) is at the same rate as the discount rate. The EFFR has a better claim to being a representative short rate, and using it also avoids dealing with the change in discount window policy in 2003, when the adjustment credit program, extended at a below market rate, was replaced by the primary credit program, set at a penalty rate above the federal funds rate. This was not a change in the stance of monetary policy, but rather in the administration of the discount window (Madigan and Nelson, 2002). For robustness we also use the discount rate up until the present, and note that the two series move closely together: the correlation in the overlap period is 0.97. For long rates we use long-term Treasury bonds, putting together several surprisingly consistent series. This gives us just over a century of monthly data, from 1919 to 2021.

Summary Statistics are in Table 1.

IV. Results

We end up with four separate sets of results, in and out-of-sample, for IP and GDP. Though patterns emerge, the results are not all identical.

A. *Quarterly in-sample results*

The long quarterly data set lets us look at the US economy before there was a Federal Reserve System, and also lets us look at an extended deflationary environment, albeit one under the gold standard.

Table 2 presents the results of the full sample in sample predictive regressions for quarterly real GNP. In the regression with lags of RGDP, the short rate and the spread, the spread is significant. Accounting for low rate environments noticeably improves the R^2, and while the spread by itself is no longer significant, the interaction is highly significant and positive: the F-test for both LOW (the dummy for low rate periods) and lospread (the interaction between low rate periods and the level of the short rate) being zero is 7.11, significant at the 1 percent

Table 1—Summary Statistics

Quarterly data					
Series	Obs	Mean	Std Error	Minimum	Maximum
RGDPGROW	581	0.047	0.083	-0.336	0.341
CP	581	4.60	2.77	0.19	16.21
BAA	581	6.69	2.44	0.00	17.11
SPREAD	581	2.09	1.96	-6.54	8.50
RGROW	581	4.74	8.34	-33.61	34.12
LOSPREAD	581	0.79	1.73	-3.40	8.50
Monthly data					
Series	Obs	Mean	Std Error	Minimum	Maximum
SPLICEDLONGRATE	1226	4.77	2.59	0.623	14.14
DISCOUNT	1226	3.78	2.66	0.25	14.00
IP	1225	45.50	35.05	3.74	112.05
CPI	1226	86.06	80.08	12.60	263.01
EFFR	1226	37.46	45.02	0.05	99.00
POLRATE	1226	3.94	3.23	0.05	19.10
SPREAD	1226	0.99	1.20	-3.12	4.43
NEWSPREAD	1226	0.82	1.49	-7.43	4.38
IPGROWTH	1213	0.030	0.108	-0.409	0.479

Source: Authors' calculation from Appendix

A. *Note:* RGDPGROW is year-over-year growth rate of Real GDP, CP is the commercial paper rate, BAA is the BAA rate, RGROW is RGDPGROWx100, SPREAD is BAA-CP, LOSPREAD is SPREAD times a dummy for low interest rate periods. SPLICED LONG RATE is an interest rate on long-maturity US Treasury securities, DISCOUNT is the discount rate at the Federal Reserve Bank of New York, IP is the level of industrial production, CPI is the level of the consumer price index, EFFR is the effective federal funds rate, POLRATE is the discount rate from 1919 to February 1963, and EFFR after, Spread is SPLICEDLONGRATE minus DISCOUNT, NEWSPREAD is SPLICEDLONGRATE minus POLRATE, and IPGROWTH is year-over-year growth of IP.

(actually, the 0.1 percent) level. This suggests that it is precisely in low interest rate environments that the yield spread has predictive power.

Another way to address the question is to compare in-sample predictive regressions across low and high interest rate environments. Table 3 provides some additional evidence that the yield curve predicts best in low rate periods, breaking out periods of high and low interest rates. If the regression includes the short rate, the coefficient on the yield curve is positive and significant in both the deflationary period and the first (1932–1951) low rate period. Only the most recent low rate period (2016–2020) shows an insignificant coefficient. The evidence is even stronger if the short rate is excluded.

B. Quarterly out-of-sample

The in-sample regressions have the advantage of using all available data, but may be inappropriate if the underlying relationship shifts over time, as technology, financial practices, and regulations develop over time. For that, an alternative is out-of-sample forecasts, or given a historical data set, pseudo-out-of-sample, that use data only up to a particular date to make a forecast, so that one does not use future information to help predict the future.

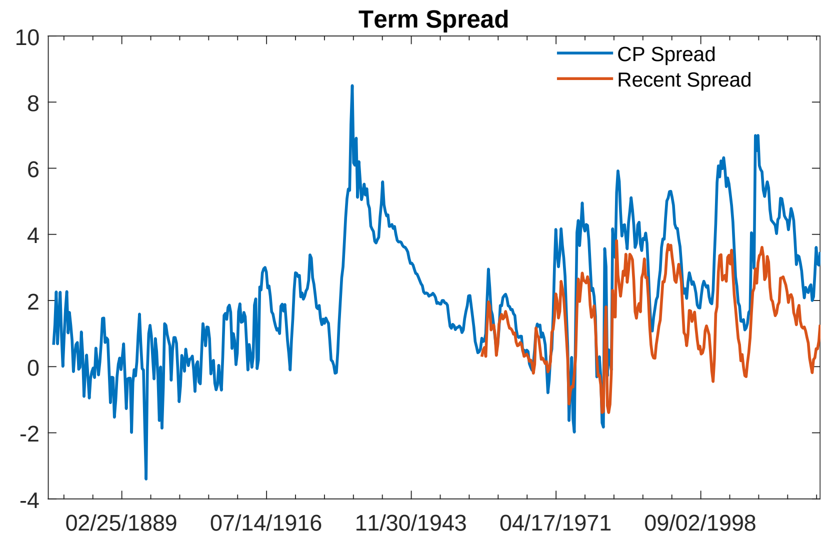

Figure 2. Corporate Commercial Paper and 10-Year 3-Month Spread

Table 2—In-Sample Regressions: Quarterly Real GDP growth.

Variable	Coeff	
1. Constant	3.968***	3.467***
	(4.32)	(3.45)
2. RGROW5	0.126*	0.118*
	(1.93)	(1.89)
3. SPREAD5	0.316*	-0.115
	(1.77)	(0.53)
4. CP rate5	-0.098	0.113
	(0.94)	(1.02)
4. LOW5		-1.457
		(1.39)
5. LOSPREAD5		1.198***
		(3.58)
R^2	0.022	0.042

Quarterly Data From 1876:Q1 to 2021:Q1
N=576, d.f.=570,

Note: This table lists the regression output for1 and 2 with real GDP growth as the dependent variable. RGROW5 is year-over-year RGDP growth lagged 5 quarters, SPREAD5 is the BAA-CP spread, lagged 5 quarters; CP rate5 is the commercial paper rate lagged 5 quarters, LOW5 is a dummy for low rate periods (1876:Q1-1897:Q2, 1932:Q2-1951:Q1, and 2008:Q4-2015:Q4) lagged 5 quarters, and LOSPREAD5 is LOW interacted with SPREAD, lagged 5 quarters. Robust regressions with the Newey-West standard errors. *,**, and *** denote significance at 10%, 5%, and 1%, respectively. t-statisitics in parenthesis. Source: Authors' calculations

Table 3—Spread coefficients by time period; In-sample regressions, Quarterly RGNP data,

	Time period	Spread	Spread, no short rate
deflationary	1876:Q1–1897:Q2	3.822*	1.601*
		(1.32)	(1.64)
	1897:Q3–1932:Q1	-0.820	0.534
		(0.89)	(0.64)
low rate	1932:Q2–1951:Q1	2.065***	1.572*
		(2.64)	(1.89)
	1951:Q2–2008:Q3	0.084	0.086
		(0.65)	(0.67)
low rate	2008:Q4–2015:Q4	- 0.203	1.352***
		(0.82)	(6.08)
	2016:Q1–2020:Q1	-0.085**	-0.041
		(2.33)	(0.24)

Note: This table reports the coefficient on the spread variable for time periods classified and deflationary, low rate, and other. The independent variable is RGDP growth. The first column reports the coefficient on spread when a short rate is included in the regression, and the second when the short rate is excluded. *,**, and *** denote significance at 10%, 5%, and 1%, respectively. t-statistics in parenthesis. Source: authors' calculations

A rolling regression serves as an ocular first test; Figure 3 graphs the coefficient on the lagged spread and two standard error bands from a rolling regression of RGDP growth over the next year against lagged RGDP growth, lagged CP rate, and lagged yield spread, with a rolling window of 40 quarters (10 years). The coefficient on the yield spread shows up positive and significant in several periods, around 1910-1915, 1925 to the early 1930s, and the mid-1930s to 1940 or so, with additional periods in the 1960s, 1970s, 1990s, and 2000s. There is some overlap with low rate periods, but the coefficient is also positive in more normal periods. What jumps out is the decreased variability starting in the early 1960s (c.f. Romer 1986).

Taking a more formal approach, we use the ORS procedure to compare the RGDP forecasts from rolling regressions with and without the spread variable. Table 4 reports the results. Panel A compares a regression with lagged RGDP growth (the restricted model) with one that also includes the yield spread (unrestricted model). The initial tests for forecast (the sup-, ave- and exp-W statistics) find no significant difference between the restricted and unrestricted models. However, the positive and significant α in the loss differential regression indicates that the restricted model has on average higher squared forecast errors, and the negative θ indicates that at high interest rates (above the threshold), the forecast using the yield spread does relatively worse. Thus, though this evidence is not strong, it suggests that, if anything, the yield curve predicts better at low rates.

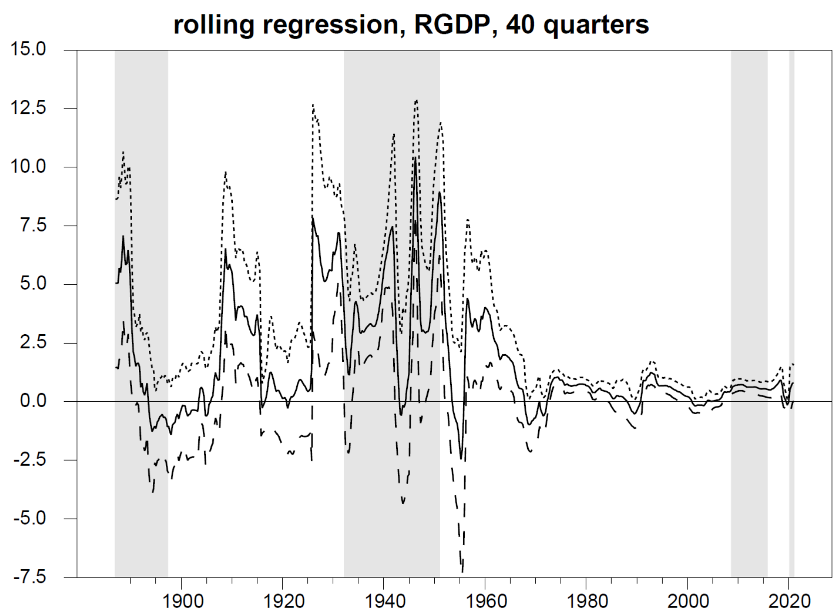

Figure 3. Coefficient on spread, Quarterly,low rate periods shaded

The endogenously determined threshold, $\gamma = 0.94$, marks the low rate environment as starting just below 1 percent. The encompassing tests indicate some benefit from using the yield spread (via the ave-W statistic), but the loss differential regression is not significant. The results using a 40-quarter window are quite similar, with the intriguing difference that the low rate threshold appears to be much lower, at about 26 basis points. These results reinforce the message of the in-sample regressions in Tables (2) and (3) that the slope of the yield curve has a *closer* relationship with future GDP growth in low rate periods.

Table 4—State-Dependent Forecast Tests; RGDP growth

20-quarter window	Forecast		Encompassing	
Statistic	statistic	p-value	statistic	p-value
sup-W	8.6608	0.4492	14.1199	0.2126
ave-W	3.0247	0.4084	**11.7065**	0.0900
exp-W	2.0464	0.4410	6.0030	0.1296
γ	0.940		8.090	
	$\alpha = 46.2384$	θ=**-66.0112**	$\alpha = 30.9160$	θ=5.6825
	(2.1707)	(-1.7513)	(0.8008)	(0.2989)
N=560,				
40-quarter window				
Statistic	statistic	p-value	statistic	p-value
sup-W	10.1168	0.3734	**24.7177**	0.0956
ave-W	2.1500	0.6206	10.2783	0.1694
exp-W	1.6465	0.6012	7.8938	0.1422
γ	0.260		0. 260	
	$\alpha = 6.4600$	θ=-10.1328	$\alpha = 4.4592$	θ=9.0437
	(0.5012)	(-0.7328)	(0.4736)	(0.7983)
Quarterly Data 1876:2 To 2021:1				

Note: The first colum reports the ORS W statistics of forecast ability based on regression 1 of future RGDP growth against lagged RGDP growth, and a regression of future RGDP growth against lagged RGDP growth and the cp-baa yield spread. The second column reports the p-value of the statistics. The third and fourth columns repeat for tests of encompassing. γ is the estimated threshold level of the short rate. α and θ are the coefficients from equation (6). The first panel reports the results using a rolling window of 20 quarters, the second for 40 quarters. **Bold** denotes significance at 10%. t-statistics in parenthesis.
Source: Authors' calculations.

C. Monthly IP data: In-sample

This section reports the results of in-sample regressions for equations (1) and (2) for Industrial Production, that is, a regression of the future IP growth against lagged IP growth rate, a short rate, and the yield spread, and then adds a dummy for the low interest rate environment periods and an interaction term between the spread and the low interest rate periods, which captures a shift in both the intercept and the slope. We use a lag of 13 months as a simple attempt to account for "real-time" issues, since IP is not reported fully contemporaneously. For the short rate, we use both the policy rate, which is the discount rate until 1963 and then the federal funds rate, and also separately the discount rate.

Table 5 reports the results. The first two columns report the coefficients and t-statistics for equations (1) and (2) using the policy rate as the short rate. The spread coefficient is highly significant and positive indicating that a steeper slope indicates higher future growth. Is it economically significant? A one standard deviation increase in the spread predicts an increase in expected IP growth of 1.5 percent, which is 15 percent of IP growth's standard deviation. After adding a dummy for low rates and an interaction term, the coefficient on the spread is still significant, and actually slightly larger. Dummy and interaction are both significant. The interaction term comes in as negative and larger in absolute value than the coefficient on the spread by itself. This suggests that while the yield curve retains its predictive ability in low interest rate environments, a steeper yield curve indicates lower growth in those periods. The third and fourth columns repeat the exercise using the discount rate as the short rate: the results are qualitatively quite similar.

Another way to address the question is to compare in-sample predictive regressions across low and high interest rate environments. Table 6 shows the coefficients of the lagged spread from in-sample regressions of future IP growth, breaking out periods of high and low interest rates. It separately reports the results using both the merged discount rate/fed funds rate and the discount rate as the short rate.

This weakly confirms the full sample analysis, that in low rate periods the spread has either a negative (1932:4–1951:2) or small and insignificant (2008:10–2015:12) coefficient. This also confirms one of the highly robust conclusions of the yield curve literature that its predictive ability depends sensitively on the time period.

D. Monthly data–Pseudo-out-of-sample

As in the quarterly case, a rolling regression can serve as an ocular first test; Figure 4 graphs the coefficient on lagged spread and two standard error bands from a rolling regression of IP over the next year against lagged IP growth, lagged policy rate, and lagged yield spread, using a window of 5 years (60 months). The coefficient is almost uniformly positive after 1960, though the size and significance vary over time. Prior to that, the coefficient is quite variable. Somewhat in contrast to the in-sample results, the coefficient is positive and significant in the late 1930s and early 1940s, before it turns significantly (in both senses) negative in the 1950s.

To make a more formal comparison, we again turn to the Odendahl, Rossi, and

Table 5—In-sample regressions; IP growth.

Variable	Coeff			
Equation	(1)	(2)	(3)	(4)
Constant	0.031***	0.005	0.040***	0.016**
	(4.13)	(0.58)	(5.54)	(2.14)
IPGROWTH13	-0.122***	-0.153***	-0.128***	-0.149***
	(-2.81)	(3.72)	(2.95)	(3.58)
SPREAD13	0.011***	0.015***		
	(5.80)	(8.17)		
POLRATE13	-0.002*	0.002**		
	(1.89)	(2.04)		
DiscSPREAD13			0.010***	0.012***
			(4.96)	(5.86)
Discount Rate13			-0.004***	-0.001
			(4.58)	(0.93)
LOW13		0.117***		0.099***
		(6.62)		(5.56)
LOSPREAD13		-0.050***		-0.056***
		(7.24)		(5.16)
R^2	0.049	0.092	0.047	0.073

Monthly Data From 1921:02 to 2021:01
N=1200, d.f.=1196,1194,

Note: The first column reports the output of regression 1 of future IP growth against lagged IP growth rate, the yield spread, and the policy rate. SPREAD13 is a measure of long-term US Treasury rate minus POLRATE, the discount rate until February 1963 and the EFFR after. The second column reports the output of 2, and adds a dummy for the low interest rate environment periods, and an interaction term between the spread and the low interest rate periods. Columns (3) and (4) report the results using the discount rate as the short rate. DiscSPREAD is the spread between the long-term rates and the discount rate. Estimated in RATS, with Newey-West corrected errors for autocorrelations and heteroskedasticity. *,**, and *** denote significance at 10%, 5%, and 1%,
respectively. t-statistics in parenthesis.
Source: Authors' calculations.

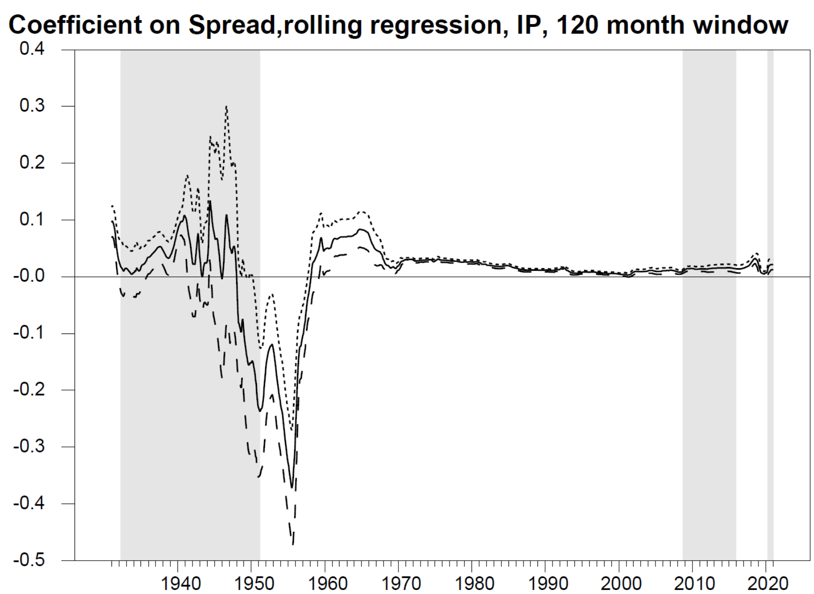

Figure 4. Coefficient on spread, monthly IP, low rate period shaded

Table 6—Spread coefficients by time period; IP growth.

	Time period	Spread	DiscSpread
	1920:1–1932:3	0.074**	same
		(2.26)	
low rate	1932:4–1951:3	-0.102**	same
		(2.29)	
	1951:4–2008:10	0.008***	0.005***
		(5.95)	(4.08)
low rate	2008:10–2015:12	0.003	-0.0004
		(0.70)	(0.09)
	2016:1–2020:2	-0.036***	same
		(3.19)	

Note: This reports the output of regressions 1 for different time periods. The column Spread uses the 10-year Treasury yield less the policy rate as spread, while the column DiscSpread uses the 10-year Treasury less the Federal Reserve discount rate as the spread. *,**, and *** denote significance at 10%, 5%, and 1%, respectively. t-statistics in parenthesis. Source: Authors' calculations

Sekhposyan (2021) tests of conditional forecast accuracy and encompassing. Table 7 reports the results. For industrial production over this time period, the yield curve shows little predictive power beyond forecasts using only lagged IP. None of the W statistics are significant in the forecasting equations, and only one for the ecompassing equations. The α and θ parameters are all quite small and only one is of marginal significance, indicating that there is little difference between the squared errors of the base forecast and the forecast using the yield spread. The one significant value of θ provides a hint that at interest rates above the threshold, the base forecast has larger squared errors, and therefore base forecast does worse relative to the forecast including the yield curve. In a backhanded way, this suggests that the forecast using the yield curve does relatively better in times of low rates. The value of θ is not large, and the estimated value of the threshold, $\gamma = 5.31$, suggests that this is not a low rate phenomenon.

V. Robustness

The quarterly in-sample results suggest that, if anything, the yield curve has more predictive power in low interest rate environments. The monthly results tell a different story. To try and explain the difference, we run the in-sample quarterly regressions over the same time period as the monthly IP regressions (1919-1921) and also using the same yield spread used in the monthly IP regressions.

Table 8 reports the results using the same time period as the IP numbers, but keeping the Baa-cp spread. Table 9 reports the results for the sub-samples of high and low interest rate environments. Over the entire 1919-2021 sample, the spread

Table 7—State-dependent forecast tests; IP growth.

60-month window				
	Forecast		Encompassing	
Statistic	statistic	p-value	statistic	p-value
sup-W	18.6606	0.4884	36.5386	0.1396
ave-W	5.7873	0.5764	30.4652	**0.0302**
exp-W	6.3188	0.5108	16.1648	0.1146
γ	5.31		0.400	
	$\alpha = 6.1x10^{-4}$	θ=**0.0022**	$\alpha = 3.9x10^{-4}$	θ=0.0015
	(0.4177)	(1.7399)	(0.4404)	(1.5566)
N = 1152				
120-month window				
Statistic	statistic	p-value	statistic	p-value
sup-W	18.7481	0.5546	29.8588	0.2840
ave-W	6.1628	0.6088	12.8552	0.2676
exp-W	7.1416	0.5098	12.1638	0.2720
γ	0.1900		0.190	
	$\alpha = 8.9x10^{-4}$	$\theta = 6.4x10^{-4}$	$\alpha = 6.2x10^{-4}$	$\theta = -6.0x10^{-4}$
	(0.7255)	(-0.4984)	(0.9460)	(-8.4$x10^{-4}$)
N=1092,				
Monthly Data 1921:02 to 2021:01				

Note: The first column reports the W statistics for forecast performance of ORS based on regression 1 of future IP growth against lagged IP growth rate, and a regression of future IP growth against lagged IP growth rate and and the yield spread using the policy rate. The second column reports the probability value of the statistic. Third and fourth columns report the statistics and p-values for encompassing tests. The first panel reports the results for a rolling window of 60 months, the second for 120 months. γ is the threshold determined in the ORS tests. α and θ are the coefficients from equation (6). **Bold** denotes significance at the 10% level.
Source: Authors' calculations.

remains significant, although the significance disappears when a dummy for low rate periods and an interaction term are added, though neither is individually significant. In the sub-sample regressions, the spread remains significant in the low rate period of the 1930s and 1940s, but for the post-crisis period only if the short rate is not included in the regression. So as before, to the extent that the evidence points in any one direction, it is that the yield curve predicts better, if anything, in low rate environments.

Table 8—In-Sample regression real GDP growth, short sample.

Variable	Coeff	
1. Constant	1.930*	0.267
	(1.76)	(0.27)
2. RGROW5	0.221**	0.178**
	(2.44)	(2.04)
3. SPREAD5	0.585***	0.239
	(2.92)	(1.10)
4. CP rate5	-0.017	0.351***
	(0.017)	(3.61)
4. LOW5		4.167
		(1.59)
5. LOSPREAD5		0.213
		(0.36)
R^2	0.053	0.085

Quarterlyly Data From 1919:Q1 to 2021:Q1
N=404, d.f.=399,

Note: This table lists the regression output for1 and 2 with real GDP growth as the dependent variable and the baa-cp yield spread. The sample is restricted to the same period for which IP data are available. Robust regressions with the Newey-West standard errors. *,**, and *** denote significance at 10%, 5%, and 1%, respectively. t-statistics in parenthesis. Source: Authors' calculations

Now, using the Treasury-policy rate spread (the same spread used in the IP regressions), we find some evidence that the term spread may do worse in low interest rate periods. In the full-sample regression, displayed in Table 10, though the spread is positive and significant in a regression with a low-rate dummy and interaction term, the interaction term is negative, indicating that in low rate environments, the relationship between the term spread and future output is less positive. In sub-samples, Table 11, the coefficient on the yield spread is sometimes negative in the low rate environments of 1931-51 and 2008-2015.

The ORS (2021) results on contingent predictability for quarterly real GDP restricted to 1919-2021 generally show little evidence that the yield curve adds predictive content; the tables are reported in the robustness appendix.

Table 9—Spread coefficients, in-sample regressions; RGDP growth, short sample.

	Time period	Spread	Spread, no short rate
	1919:Q1–1932:Q1	5.379***	5.1444***
		(2.77)	(3.75)
low rate	1932:Q2–1951:Q1	2.064***	1.571*
		(2.64)	(1.89)
	1951:Q2–2008:Q3	0.084	0.085
		(0.65)	(0.67)
low rate	2008:Q4–2015:Q4	-0.203	1.352***
		(0.81)	6.08)
	2016:Q1–2020:Q1	-0.086**	-0.041
		(2.34)	(0.241)

Note: This table reports the coefficient on the spread variable for time periods classified and deflationary, low rate, and other. The independent variable is RGDP growth. The first column reports the coefficient on spread when a short rate is included in the regression, and the second when the short rate is excluded. *,**, and *** denote significance at 10%, 5%, and 1%, respectively. t-statistics in parenthesis. Source: Authors' calculations

Table 10—In-sample regression: Real GDP growth, using PolRate.

Variable	Coeff	
1. Constant	3.056***	-0.503
	(2.87)	(0.45)
2. RGROW5	0.194**	0.141*
	(2.14)	(1.78)
3. SPREAD5	0.355	0.790***
	(1.28)	(2.86)
4. Policy rate5	0.066	0.574***
	(0.61)	(4.85)
4. LOW5		12.839***
		(6.34)
5. LOSPREAD5		-4.678***
		(5.66)
R^2	0.039	0.142

Quarterlyly Data From 1919:Q1 to 2021:Q1
N=404, d.f.=398,

Note: This table lists the regression output for1 and 2 with real GDP growth as the dependent variable. The spread and policy rate are quarterly averages, using the spread used in IP regressions. Robust regressions with the Newey-West standard errors. *,**, and *** denote significance at 10%, 5%, and 1%, respectively. t-statistics in parenthesis. Source: Authors' calculations

Table 11—Spread coefficients, in-sample regressions, RGNP growth, short sample using CP as short rate.

	Time period	Spread	Spread, no short rate
	1919:Q1–1932:Q1	5.083	4.092
		(1.41)	(1.58)
low rate	1932:Q2–1951:Q1	-18.237***	-0.938
		(3.31)	(1.46)
	1951:Q2–2008:Q3	- 0.190	-0.110
		(1.19)	(0.77)
low rate	2008:Q4–2015:Q4	-0.044	1.554***
		(0.18)	(5.39)
	2016:Q1–2020:Q1	-0.107	0.204
		(0.18)	(0.43)

Note: This table reports the coefficient on the spread variable for time periods classified and deflationary, low rate, and other. The independent variable is RGDP growth. Using quarterly averages of the policy rate as the short rate, both by itself and in the spread.The first column reports the coefficient on the spread when a short rate is included in the regression, and the second when the short rate is excluded. *,**, and *** denote significance at 10%, 5%, and 1%, respectively. t-statistics in parenthesis. Source: Authors' calculations

VI. Conclusion

So does the yield curve predict output when interest rates are low? It depends on the particular question; this work re-confirms that the predictive content of the yield curve varies over time and across data sets. Accounting for episodes of low interest rates can improve the ability of the yield curve to predict output. The reason, however, varies: for monthly industrial production, the slope of the yield curve has less of a relationship with future output growth when rates are low, but for quarterly real GDP, the relationship is stronger. We find the GDP results to be more interesting, both because it is the broader, less noisy measure and because the results include the secular price decline ending in 1897. The initial draft of this paper was written before the COVID-19 outbreak, and it was a matter of some concern whether the inversion in the summer of 2019 presaged a recession; although a recession did occur, the question remains as to whether the conjunction is just a coincidence. This paper is not meant to be a forecasting exercise, but it does suggest that the current low interest rate environment should not be a reason to dismiss concerns over recent inversions of the yield curve.

Our results raise as many questions as they answer: If the level of rates matters, does it matter why rates are low, whether inflation expectations are anchored under a gold standard or a price level target, are explicitly pegged, or held down by forward guidance or quantitative easing? Do aspects of inflation such as persistence matter? These are questions for another day, though perhaps our results can contribute to the answers.

References

Ang, Andrew, Monika Piazzesi, and Min Wei. 2006. "What Does the Yield Curve Tell Us about GDP Growth?" Journal of Econometrics 131 (12): 359–403. https://doi.org/10.1016/j.jeconom.2005.01.032.

Balke, Nathan, and Robert J. Gordon. 1986. "Appendix B: Historical Data." In The American Business Cycle: Continuity and Change, 781–850. University of Chicago Press. https://www.nber.org/chapters/c10036.

Bauer, Michael D., and Thomas M. Mertens. 2018. "Economic Forecasts with the Yield Curve." FRBSF Economic Letter. https://econpapers.repec.org/article/fipfedfel/00158.htm.

Belz, Sage, and David Wessel. 2020. "What Is Yield Curve Control?" Brookings (blog). June 5, 2020. https://www.brookings.edu/blog/up-front/2020/06/05/what-is-yield-curve-control/.

Benati, Luca, and Charles Goodhart. 2008. "Investigating Time-Variation in the Marginal Predictive Power of the Yield Spread." Journal of Economic Dynamics and Control 32 (4): 1236–72. https://doi.org/10.1016/j.jedc.2007.05.005.

Bordo, Michael D., and Joseph G. Haubrich. 2008a. "The Yield Curve as a Predictor of Growth: Long-Run Evidence, 18751997." Review of Economics and Statistics 90 (1): 182–85. https://doi.org/10.1162/rest.90.1.182.

Bordo, Michael D., and Joseph G. Haubrich. 2008b. "Forecasting with the Yield Curve; Level, Slope, and Output 18751997." Economics Letters 99 (1): 48–50. https://doi.org/10.1016/j.econlet.2007.05.026.

Bordo, Michael D., Angela Redish, and John Landon Lane. 2009. "Good versus Bad Deflation: Lessons from the Gold Standard Era." In Monetary Policy in Low-Inflation Economies, edited by David E. Altig and Ed Nosal. Cambridge University Press. https://doi.org/10.1017/CBO9780511605475.010.

Bordo, Michael D., and Pierre Siklos. 2015. "Central Bank Credibility: An Historical and Quantitative Exploration." Working paper 20824. National Bureau of Economic Research. https://doi.org/10.3386/w20824.

Calomiris, Charles W. 1993. "Greenback Resumption and Silver Risk: The Economics and Politics of Monetary Regime Change in the United States, 1862–1900." In Monetary Regimes in Transition, edited by Michael D. Bordo and Forrest Capie, 1st ed., 86–132. Studies in Macroeconomic History. Cambridge University Press. https://doi.org/10.1017/CBO9780511664564.

Chinn, Menzie, and Kavan Kucko. 2015. "The Predictive Power of the Yield Curve Across Countries and Time." International Finance 18 (2): 129–56. https://doi.org/10.1111/infi.12064.

Cooper, Daniel H., Jeffrey C. Fuhrer, and Giovanni P. Olivei. 2020. "Predicting Recessions Using the Yield Curve: The Role of the Stance of Monetary Policy." 87522. Current Policy Perspectives. Federal Reserve Bank of Boston. https://ideas.repec.org/p/fip/fedbcq/87522.html.

Diebold, Francis X. 2015. "Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold-Mariano Tests." Journal of Business & Economic Statistics 33 (1): 1–9. https://doi.org/10.1080/07350015.2014.983236.

Dotsey, Michael. 1998. "The Predictive Content of the Interest Rate Term Spread for Future Economic Growth." Economic Quarterly (Federal Reserve Bank of Richmond) 84 (3): 31–51. https://ideas.repec.org/a/fip/fedreq/y1998isump31-51.html.

Eggertsson, Gauti B., and Michael Woodford. 2003. "The Zero Bound on Interest Rates and Optimal Monetary Policy." Brookings Papers on Economic Activity 34 (1): 139–235. https://doi.org/10.1353/eca.2003.0010.

Estrella, Arturo. 2005. "The Yield Curve as a Leading Indicator: Frequently Asked Questions." Federal Reserve Bank of New York. https://www.newyorkfed.org/medialibrary/media/research/capital_markets/ycfaq.pdf.

Estrella, Arturo, and Gikas A. Hardouvelis. 1991. "The Term Structure as a Predictor of Real Economic Activity." The Journal of Finance 46 (2): 555–76. https://doi.org/10.1111/j.1540-6261.1991.tb02674.x.

Estrella, Arturo, and Frederic S. Mishkin. 1997. "The Predictive Power of the Term Structure of Interest Rates in Europe and the United States: Implications for the European Central Bank." European Economic Review 41 (7): 1375–1401. https://doi.org/10.1016/S0014-2921(96)00050-5.

Fendel, Ralf, Nicola Mai, and Oliver Mohr. 2019. "Predicting Recessions Using Term Spread at the Zero Lower Bound: The Case of the Euro Area." VoxEU/CEPR Policy Portal (blog). January 17, 2019. https://voxeu.org/article/predicting-recessions-using-term-spread-zero-lower-bound.

Garbade, Kenneth. 2020. "Managing the Treasury Yield Curve in the 1940s." Staff Report 913. Federal Reserve Bank of New York. https://ideas.repec.org/p/fip/fednsr/87456.html.

Giacomini, Raffaella, and Barbara Rossi. 2006. "How Stable Is the Forecasting Performance of the Yield Curve for Output Growth?" Oxford Bulletin of Economics and Statistics 68 (s1): 783–95. https://doi.org/10.1111/j.1468-0084.2006.00456.x.

Hamilton, James, and Dong Heon Kim. 2002. "A Reexamination of the Predictability of Economic Activity Using the Yield Spread." Journal of Money, Credit and Banking 34 (2): 340–60. https://doi.org/10.1353/mcb.2002.0040.

Harvey, Campbell R. 1988. "The Real Term Structure and Consumption Growth." Journal of Financial Economics 22 (2): 305–33. https://doi.org/10.1016/0304-405X(88)90073-6.

Haubrich, Joseph G. 2021. "Does the Yield Curve Predict Output?" Annual Review of Financial Economics, 13:341–362 https://doi.org/10.1146/annurev-financial-100620-065648

Laurent, Robert D. 1988. "An interest rate-based indicator of monetary policy." Federal Reserve Bank of Chicago Economic Perspectives. vol. 12 pp. 3–14.

Liu, Weiling, and Emanuel Moench. (2016) "What Predicts U.S. Recessions?" International Journal of Forecasting, Vol. 32 No. 4, October 2016, pp. 1138–1150.

Mackintosh, James. 2020. "The Markets Favorite Recession Signal Probably Has It Wrong." Wall Street Journal, April 2, 2020, sec. Markets - Streetwise. https://www.wsj.com/articles/the-markets-favorite-recession-signal-probably-has-it-wrong-11580824310.

Madigan, Brian, and William R. Nelson. 2002. "Proposed Revision to the Federal Reserves Discount Window Lending Programs." Federal Reserve Bulletin 88 (7): 313–19. https://doi.org/10.17016/bulletin.2002.88-7.

McDonald, John H. 2014. Handbook of Biological Statistics. 3rd ed. Baltimore, MD: Sparky House Publishing.

McMillan, David G. 2021. "When and Why Do Stock and Bond Markets Predict US Economic Growth?" The Quarterly Review of Economics and Finance 80 (May): 331–43. https://doi.org/10.1016/j.qref.2021.03.004.

Odendahl, Florens, Barbara Rossi, and Tatevik Sekhposyan. 2021. "Evaluating Forecast Performance with State Dependence." Working paper 1800. Department of Economics and Business, Universitat Pompeu Fabra. https://ideas.repec.org/p/upf/upfgen/1800.html.

Okimoto, Tatsuyoshi. 2019. "Trend Inflation and Monetary Policy Regimes in Japan." Journal of International Money and Finance 92 (April): 137–52. https://doi.org/10.1016/j.jimonfin.2018.12.008.

Plosser, Charles I., and K. Geert Rouwenhorst. 1994. "International Term Structures and Real Economic Growth." Journal of Monetary Economics 33 (1): 133–55. https://doi.org/10.1016/0304-3932(94)90017-5.

Ramey, Valerie A., and Sarah Zubairy. 2018. "Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data." Journal of Political Economy 126 (2): 850–901. https://doi.org/10.1086/696277.

Romer, Christina D. 1986. "Is the Stabilization of the Postwar Economy a Figment of the Data?" The American Economic Review 76 (3): 314–34. https://www.jstor.org/stable/1813353.

Rossi, Barbara. forthcoming. "Forecasting in the Presence of Instabilities: How Do We Know Whether Models Predict Well and How to Improve Them." Journal of Economic Literature. https://doi.org/10.1257/jel.20201479.

Rudebusch, Glenn D., Brian P. Sack, and Eric T. Swanson. 2007. "Macroeconomic Implications of Changes in the Term Premium." Review (Federal Reserve Bank of St. Louis) 89 (Jul): 241–70. https://ideas.repec.org/a/fip/fedlrv/y2007ijulp241-270nv.89no.4.html.

Sims, Eric, and Jing Cynthia Wu. 2020. "Evaluating Central Banks Tool Kit: Past, Present, and Future." Journal of Monetary Economics, April. https://doi.org/10.1016/j.jmoneco.2020.03.018.

Swanson, Norman R. 1998. "Money and Output Viewed through a Rolling Window." Journal of Monetary Economics 41 (3): 455–74. https://doi.org/10.1016/S0304-3932(98)00005-1.

Thoma, Mark A. 1994. “Subsample Instability and Asymmetries in Money-Income Causality.” Journal of Econometrics 64 (12): 279–306. https://doi.org/10.1016/0304-4076(94)90066-3.

Wheelock, David C., and Mark E. Wohar. 2009. “Can the Term Spread Predict Output Growth and Recessions? A Survey of the Literature.” Review (Federal Reserve Bank of St. Louis) 91 (Sep): 419–40. https://ideas.repec.org/a/fip/fedlrv/y2009isepp419-440nv.91no.5.html.

Zarnowitz, Victor, and D. Lee. 2005. “The New Treatment of the Yield Spread in the TCB Composite Index of Leading Indicators.” The Conference Board. https://www.conference-board.org/data/bci/index.cfm?id=2161.

Data Appendix

We extend the B-G data set as follows.

1) Yield spread: The CP and Baa from Balke and Gordon can be easily updated as far as it goes, 1998. Source: Federal Reserve H-15, 6-MONTH PRIME COMMERCIAL PAPER - AVERAGE DEALER OFFERING RATE QUOTED ON DISCOUNT BASIS. To this we append data from Bloomberg, DCPB180D Index, US Commercial Paper Placed Top 180 Day Discount.
2) We update Baa by using Moody’s Baa till 1998, and then DBAA from FRED, Moody’s Seasoned Baa Corporate Bond Yield.
3) Real GNP. Use BG till 1983. Take natural logs to get YOY growth. From 1983 on, use the Ramey RGDP data, taking logs to get YOY growth. Specifically, the 1984 growth should use the Ramey 1983 data. The growth rates of the two series match pretty well, even if the levels dont, and this avoids the jump when the series splice together. After that, follow RZ and use quarterly NIPA accounts.
4) Inflation, again use the deflator from BG till 83, then use Ramey (again YOY log diffs), again using Ramey 83 levels to compute 84 growth.

We produce an “early” monthly data set as follows.

1) Industrial production is monthly since January 1919, from FRED.
2) CPI, not seasonally adjusted, from FRED, since 1913.
3) For the policy rate (alternatively the short-term interest rate) we use the discount rate at the Federal Reserve Bank of New York. Prior to the New Deal reforms, discount rates were not uniform across the Districts, but New York, as the financial center, has a certain priority (even post-New Deal reforms, there are occasional brief differences between District banks). This is available from November 1914 until 1969:7 from the St. Louis FRED database, series M13009USM156NNBR. In February 1963, we switch to

using the effective fed funds rate, from FRED. In that month both the discount rate and the effective federal funds rate are an even 3.00 percent.

4) For the long interest rate, we start with the Yield on Long-Term United States Bonds, from the NBER Macro History Database, m13033a starting in January 1919. Starting in December 1935 this is replaced with the COMPOSITE YIELD ON U.S. TREASURY BONDS WITH MATURITY OVER TEN-YEARS from the Federal H-15 discontinued series. For the months between January 1925 and November 1935, the NBER and H15 rates match exactly, but diverge somewhat after that. That series ends in June of 2000 (2000:6) and we continue with the 10-year Treasury constant maturity rate, which has a correlation of 0.988 over the period during which both series exist (April 1953 to June 2000).

5) In the robustness section, we use quarterly averages of the policy rate and the spliced long rate in regressions using RGDP data.

The intersection of available dates gives us a data set running from January 1919 to January 2021. At 103 years, this is substantially longer than the 58 years that the more recent monthly data cover, and also allows insight into the early period of low inflation and low interest rates.

Robustness Appendix

This appendix collects some results using different spreads, time periods, and specifications of the ORS test.

Table B1 reports the ORS (2021) tests for GDP using the shortened sample, still using the Baa-CP spread.

Table B2 reports the ORS (2021) tests for GDP, but using the same spread and short rate as used in IP regressions, that is, using the policy rate as the short rate.

The standard and preferred form of the ORS tests account for a linear factor of the conditioning variable in addition to the non-linear threshold regression on the loss difference. It is possible to turn off the linear factor, which in our case means looking at the predictive content of the yield spread without accounting for the level of the short rate. We report the results here for completeness, but the results merely reinforce the results above: that some of the predictive power of the yield spread resides in the short rate, and ignoring the short rate makes the yield curve look better. Now we turn off the linear factor in the ORS tests.

Table B3 reports the results of the ORS test without accounting for the short rate, for IP growth using the policy rate in the spread.

Table B4 reports the results of the ORS test without accounting for the short rate, for IP growth using the discount rate in the spread.

Table B5 reports the results of the ORS test without accounting for the short rate, for RGDP growth using the cp rate in the spread.

Table B1—State-Dependent Forecast Tests; RGDP Growth, Short Sample Using cp-baa Spread.

20-quarter window	Forecast		Encompassing	
Statistic	statistic	p-value	statistic	p-value
sup-W	9.3515	0.5868	20.0572	0.2340
ave-W	3.8186	0.4854	14.0568	0.1014
exp-W	2.4492	0.5684	7.4407	0.2124
γ	3.68		0.260	
	$\alpha = 16.326$	θ=8.441	$\alpha = 3.594$	θ=28.196
	(1.151)	(0.678)	(0.180)	(1.043)
N=388,				
40-quarter window				
Statistic	statistic	p-value	statistic	p-value
sup-W	10.0049	0.4276	24.9425	0.1022
ave-W	1.8930	0.7094	10.3038	0.1752
exp-W	1.6017	0.6562	7.9911	0.1538
γ	0.260		0. 260	
	f$\alpha = 6.5129$	θ=-8.8451	$\alpha = 4.4742$	θ=9.8697
	(0.4784)	(-0.6233)	(0.4468)	(0.8182)
N=368,				
Quarterly Data 1919:1 to 2021:1				

Note: The first colum reports the W statistics of ORS based on regression (1) of future RGDP growth against lagged RGDP growth, and a regression of future IP growth against lagged IP growth rate and the yield spread. The second column reports the p-value of the statistics. γ is the estimated threshold level of the short rate. α and θ are the coefficients from equation (6). **Bold** denotes significance at 10%. t-statistics in parenthesis.
Source: Authors' calculations.

Table B2—State-Dependent Forecast Tests; RGDP Growth, Short Sample, using Policy Rate Spread

20-quarter window	Forecast		Encompassing	
Statistic	statistic	p-value	statistic	p-value
sup-W	4.7273	0.7122	12.1495	0.2612
ave-W	1.4758	0.6822	4.7794	0.2122
exp-W	1.1156	0.6318	3.5257	0.2752
γ	1		1.6733	
	$\alpha = 1.8486$	θ=-13.3159	$\alpha = 3.3249$	θ=9.0239
	(0.0759)	(0.4969)	(0.5350)	(0.8120)
N=388,				
40-quarter window				
Statistic	statistic	p-value	statistic	p-value
sup-W	8.7055	0.4774	13.444	0.2218
ave-W	3.5349	0.4462	6.0872	0.2052
exp-W	2.4309	0.4898	4.1771	0.2434
γ	9.2867		9.2867	
	$\alpha = -1.7954$	θ=14.4906	$\alpha = 4.0106$	θ=8.1705
	(0.1255)	(1.5414)	(0.5181)	(1.6065)
N=368,				
Quarterly Data 1919:1 to 2021:1				

Note: The first colum reports the W statistics of ORS based on regression 1 of future RGDP growth against lagged RGDP growth, The second column reports the p-value of the statistics. γ is the estimated threshold level of the short rate. α and θ are the coefficients from equation (6). **Bold** denotes significance at 10%. t-statistics in parenthesis.
Source: Authors' calculations.

Table B3—State-Dependent Forecast Tests; IP growth, without Short Rate, Using Policy Rate in Spread .

60-month window	Forecast		Encompassing	
Statistic	statistic	p-value	statistic	p-value
sup-W	**49.1399**	0.0308	**155.3278**	0
ave-W	16.5443	0.1220	**117.2887**	0
exp-W	**21.2893**	0.0326	**74.5760**	0
γ	5.29		9.370	
	$\alpha = 2.4x10^{-4}$	$\theta = 0.0011$	**0.0015**	$7.7x10^{-4}$
	(0.5492)	(1.1174)	(2.8745)	(1.1927)
120-month window				
Statistic	statistic	p-value	statistic	p-value
sup-W	24.0206	0.4088	106.3292	$4x10^{-4}$
ave-W	13.6894	0.2530	**74.6691**	0.0012
exp-W	10.0420	0.3630	**49.9844**	$4x10^{-4}$
γ	5.22		9.190	
	$\alpha = 2.6x10^{-4}$	$\theta = 8.3x10^{-4}$	$8.9x10^{-4}$	**0.0015**
	(0.5399)	(0.8631)	(1.2923)	(1.9876)
Monthly Data From 1921:02 to 2021:01				
N=1152,				

Note: The first colum reports the W statistics of ORS based on regression 1 of future RGDP growth against lagged RGDP growth, The second column reports the p-value of the statistics. γ is the estimated threshold level of the short rate. α and θ are the coefficients from equation (6). **Bold** denotes significance at 10%. t-statistics in parenthesis.
Source: Authors' calculations.

Table B4—State-Dependent Forecast Tests; IP Growth, without Short Rate, using Discount Rate in Spread.

60-month window	Forecast		Encompassing	
Statistic	statistic	p-value	statistic	p-value
sup-W	14.6005	0.6014	**110.4174**	$4x10^{-4}$
ave-W	5.0711	0.6368	**85.0883**	$2x10^{-4}$
exp-W	5.0186	0.5964	**51.5194**	$4x10^{-4}$
γ	6		6	
	$\alpha = 8.7x10^{-4}$	θ=0.0010	0.0013	$6.8x10^{-4}$
	(0.1277)	(1.0109)	(0.1277)	(1.0109)
120-month window				
Statistic	statistic	p-value	statistic	p-value
sup-W	33.8669	0.1800	**84.0015**	0.0048
ave-W	5.9314	0.6406	**55.5978**	0.0056
exp-W	13.0755	0.2164	**39.0076**	0.0050
γ	0.500		0.500	
	$\alpha =$ **0.0048**	θ=**-0.0046**	**0.0031**	**-0.0024**
	(7.5973)	(-2.5326)	(8.9883)	(2.5028)
Monthly Data From 1921:02 to 2021:01				
N=1152,				

Note: The first colum reports the W statistics of ORS based on regression 1 of future RGDP growth against lagged RGDP growth, The second column reports the p-value of the statistics. γ is the estimated threshold level of the short rate. α and θ are the coefficients from equation (6). **Bold** denotes significance at 10%. t-statistics in parenthesis.
Source: Authors' calculations.

Table B5—State-Dependent Forecast Tests; RGDP Growth, without short rate, Using cp-baa Spread.

20-quarter window	Forecast		Encompassing	
Statistic	statistic	p-value	statistic	p-value
sup-W	6.8520	0.5920	**21.9507**	0.0628
ave-W	2.6034	0.3770	**15.7155**	0.0290
exp-W	1.6157	0.5030	**6.250**	0.0614
γ	0.940		0. 940	
	α = 48.2041	θ=-51.2806	**17.3933**	-9.0836
	(2.1707)	(-1.7513)	(4.8374)	(-1.1093)
40-quarter window				
Statistic	statistic	p-value	statistic	p-value
sup-W	11.0007	0.3226	**24.3078**	0.0780
ave-W	3.9420	0.3176	9.5344	0.1368
exp-W	2.3851	0.4388	7.6811	0.1228
γ	0.260		0. 260	
	fα = 6.4301	θ=-10.7524	4.0634	0.8431
	(1.1966)	(-1.3087)	(1.1977)	(0.2188)
Monthly Data From 1876:02 to 2021:1				
N=560,				

Note: The first colum reports the W statistics of ORS based on regression 1 of future RGDP growth against lagged RGDP growth, and a regressions of future IP growth against lagged IP growth rate and the yield spread. The second column reports the p-value of the statistics. γ is the estimated threshold level of the short rate. α and θ are the coefficients from equation (6). **Bold** denotes significance at 10%. t-statistics in parenthesis.
Source: Authors' calculations.