Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

The Near-Term Forward Yield Spread as a Leading Indicator: A Less Distorted Mirror

Eric C. Engstrom and Steven A. Sharpe

2018-055

Please cite this paper as:
Engstrom, Eric C., and Steven A. Sharpe (2018). "The Near-Term Forward Yield Spread as a Leading Indicator: A Less Distorted Mirror," Finance and Economics Discussion Series 2018-055. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2018.055r1.

The Near-Term Forward Yield Spread as a Leading Indicator: A Less Distorted Mirror

Eric C. Engstrom and Steven A. Sharpe[1]

Original Draft: July 2018

This Draft: February 2019

The spread between the yield on a 10-year Treasury note and the yield on a shorter maturity security, such as a 2-year Treasury note, is commonly used as an indicator for predicting U.S. recessions. We show that such "long-term spreads" are statistically dominated in models that predict recessions or GDP growth by an economically more intuitive alternative, a "near-term forward spread." The latter can be interpreted as a measure of the market's expectations for the near-term trajectory of conventional monetary policy rates. Its predictive power suggests that, when market participants expected—and priced in—a monetary policy easing over the subsequent year and a half, a recession was quite likely in the offing. We also find that the near-term spread predicts four-quarter GDP growth with greater accuracy than survey consensus forecasts and that it has substantial predictive power for stock returns. Yields on bonds maturing beyond 6-8 quarters are shown to have no added value for forecasting either recessions, GDP growth, or stock returns.

JEL codes: E52, G12

Keywords: Yield Spread, Recession Forecast, Monetary Policy, Policy Path, Equity Return Predictability, Portfolio Strategy

[1] Board of Governors of the Federal Reserve System. The views in this document do not necessarily reflect those of the Federal Reserve System, its Board of Governors, or staff.

Introduction

Commonly cited measures of the term spread, such as the difference between the 10-year and 2-year nominal Treasury yields, dropped to nearly zero by the start of 2019, following several years of decline (Figure 1, blue line), raising concerns and provoking extensive commentary in the financial press of an impending recession. Those concerns owe to the statistical power that low levels of term spreads have shown for predicting recessions over the past several decades. In particular, many studies have documented this predictive power of the term structure, such as Estrella, and. Mishkin (1998) and Rudebush and Williams (2009), to name just a couple. Recently, Bauer and Mertens (2018) and Johansson and Meldrum (2018) have shown that the predictive power of term spreads remains undiminished of late, and is robust to the inclusion of additional predictors.

In this paper we show that, for predicting recessions, measures of a "long-term spread"—the spread in yields between a far-off maturity such as 10 years and a shorter maturity such as 1 or 2 years—are statistically inferior to an economically more intuitive measure of the term structure of interest rates. In particular, we introduce the "near-term forward spread," which can be interpreted as a measure of the market's expectations for the trajectory of conventional near-term monetary policy. When negative, it indicates that market participants expect monetary policy to ease over the next several quarters, presumably because they expect monetary policymakers to respond to the threat or onset of a recession. The predictive power of our near-term forward spread indicates that, when market participants have expected—and priced in—a monetary policy easing over the next year and half or so, their fears were validated more often than not.[2]

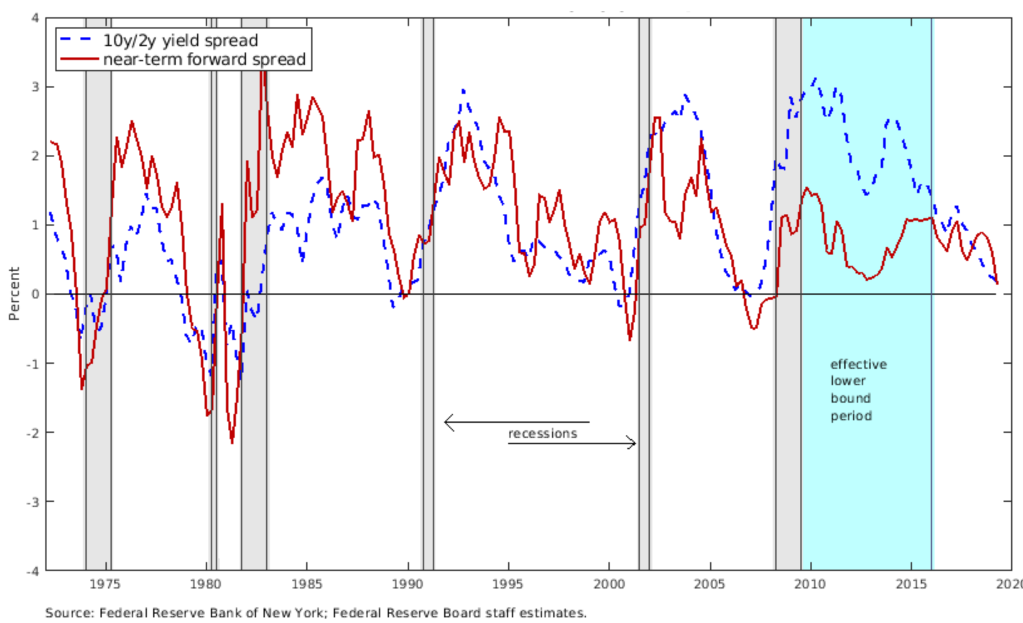
At some level, our findings merely serve to demystify the historical predictive content of the yield curve. We show that the predictive component of yield spreads largely is highly correlated with market expectations for monetary policy over the coming several quarters. The bond market serves to aggregate the heterogeneous perceptions of traders and investors, and thereby signal the expectations of the "representative" market participant. So, in some sense, for the typical market participant to infer the likelihood of recession by looking at the yield curve—particularly the near-term forward spread—is akin to that participant looking in the mirror at himself to see if he fears a recession. Even so, the yield curve does provide onlookers and policy makers with a tool for gauging the expectations of bond market participants.

Beyond presaging recessions, we find that our near-term forward spread appears to embed other related information for economists and market observers. For one thing, we show that it has predictive power for the incidence of recession even after conditioning on economists' own consensus forecast of recession probabilities as measured by the Survey of Professional Forecasters (SPF). Perhaps even more interesting, we show that the near-term forward spread has considerable power for forecasting GDP growth over the subsequent four quarters, and here

[2] Benzoni et al (2018), in a paper circulated subsequent to the initial draft of this one, also find support for this interpretation.

too it outperforms the SPF consensus GDP growth forecast. In contrast, traditional long-term spreads do relatively poorly at forecasting GDP growth. Finally, we test whether the near-term forward spread has power to predict excess returns on equities over the year ahead, and we find that it has substantial power for predicting market downturns, indeed, a fair bit more than conventional financial ratios.

Figure 1: Long-term Yield Spread and Near-term Forward Spread



Defining Near-Term Forward Spreads

Like a standard term spread measure using yield to maturity, a forward spread gauges the slope of the term structure of Treasury rates. However, using forward rates should help identify more precisely than yields-to-maturity where on the maturity spectrum the signal for a recession lies.[3] The forward rate at a given maturity can be interpreted as representing the market's expected short rate at that horizon, plus a term premium. In contrast, because a yield is an average of the forward rates spanning the period to maturity, yields tend to blur the signal embedded in forward rates.

We define the near-term forward spread on any given day as the difference between the implied interest rate expected on a three-month Treasury bill six quarters ahead and the current yield on a three-month Treasury bill.[4] Our six-quarter horizon is a common intermediate-term horizon for

---

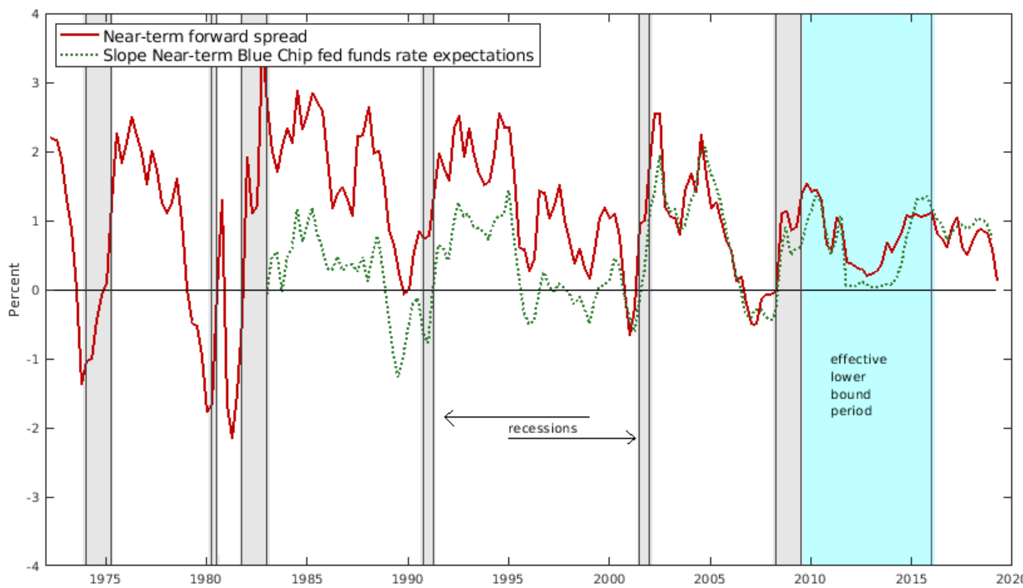
[3] Another interesting approach for decomposing the yield curve that appears to dominate the long-term spread, used by Johansson and Meldrum (2018), is to examine the predictive power of the three principal components of yield curve.

[4] As will described more fully later, the forward rate 6 quarters ahead is inferred from the yields to maturity on zero-coupon Treasury notes maturing 6 quarters ahead and 7 quarters ahead. In particular, it is the rate that would have to

forecasting, and nearly matches the horizon for which Blue Chip consensus forecasts of short-term interest rates are available.[5] To match the quarterly frequency of most macroeconomic data, we construct the near-term forward spread as a quarterly average of daily values, which is plotted in red alongside the long-term yield spread in Figure 1.[6]

Arguably, changes in this forward spread should be driven largely by changes in the market's expectations for the path of interest rates to be set by monetary policymakers over the next six quarters, departing from this only to the extent that differences in term premiums or liquidity premiums embedded in shorter-term Treasury rates change over time. Figure 2 plots the near-term spread alongside a survey-based measure of the expected trajectory of the federal funds rate. In particular the dotted green line plots the difference between the Blue Chip consensus forecasts for the average federal funds rate 5 quarters out and the current quarter. As can be seen, these two lines have moved nearly in lock step since 2001.[7] Thus, the near-term forward spread does indeed appear to be a pretty good gauge of expectations regarding monetary policy. In particular, when the near-term forward spread is negative, it signals that investors expect the Federal Reserve to ease monetary policy over this horizon.

Figure 2: Near-term Forward Spread and Market-Expected Paths of Short Rates



Source: Federal Reserve Bank of New York; Blue Chip Financial Forecasts; Federal Reserve Board staff estimates.

---

be earned on a 3-month Treasury bill purchased six quarters from now that would equate these two investment strategies: simply investing in a Treasury note that matures in 7 quarters, or investing in a Treasury note that matures 6 quarters from now and reinvesting proceeds in that 3-month Treasury bill.

[5] Inferences are nearly identical if we use any horizon ranging between 5 and 8 quarters.

[6] When we use end-of-quarter measures for interest rate spreads, such as average values over the last month of the quarter, our inferences are unaffected although the predictive value of spreads is a bit lower.

[7] Starting in 1997, the survey-based measure is the Blue Chip expected federal funds rate 5-quarters ahead minus the Treasury bill yield. Prior to 1997, the 5-quarter ahead forecast is not available, so a 4-quarter ahead forecast scaled by a factor of 5/4 is used instead.

When do investors expect monetary policy easing? Presumably, when they anticipate a substantial slowing or decline in economic activity. Consequently, assuming market participants have some foresight, it seems quite logical that low readings for the near-term forward spread tend to precede (and thus can be used statistically to forecast) recessions. This interpretation implies that inversions of the near-term forward spread do not cause recessions. Rather, they merely reflect something that market analysts already track closely—investors' expectations for monetary policy over the next several quarters and, by extension, the economic conditions driving those expectations. While, measures of the long-term spread also impound this information, they are likely to be affected by other factors unimportant for forecasting recessions, which would degrade their forecasting power.[8]

Horse Race for Recession Predictions

The data used in our analysis is quarterly and spans the period from 1972:Q1 to 2019:Q1[9]. Our macroeconomic data are standard. We use data published by the NBER to define quarters as periods of either recession or expansion. For GDP growth, we use the four-quarter log difference of real GDP as published by the Bureau of Economic Activity. Our financial data are also fairly standard. All of our spread measures begin with daily estimates of the continuously-compounded zero-coupon nominal U.S. Treasury curve estimated as in Gurkaynak, Sack and Wright (2007). The yield curve for each day is composed of yields at maturities from one to 40 quarters out. We take quarterly averages of the daily yield data. We calculate forward rates from the zero coupon curve using the standard formula

$$f_t^{n,1} = (n+1)\, y_t^{n+1} - n\, y_t^n$$

where $f_t^{n,1}$ is the (average) forward rate from quarter (*n*) to (*n+1*) that prevailed during quarter t, and y_t^n is the zero-coupon yield for maturity *n* (expressed at annual rate). The forward spread is calculated as the difference between $f_t^{n,1}$ and the (average daily) 3-month Treasury bill yield during quarter t. See appendix A for a detailed guide to constructing forward spreads.

Following the long-standing academic literature, our recession prediction analysis is based on a probit model. One minor departure from most previous studies is that we drop from the estimation any observations in which the economy was already in recession in the previous quarter. As a result, we directly estimate the probability of *transitioning* into recession. Arguably, one should allow for a somewhat different model for estimating the probability of *remaining* in recession. Nonetheless, our main results are not highly sensitive to this choice. As in some other recent studies, we also drop observations during which the effective lower bound

[8] One such factor could be a secular decline in the inflation risk premium on long-term bonds.
[9] Data for 2019:Q1 are available through January only.

on the federal funds (around zero) rate was binding; during those quarters the near-term forward spread is effectively constrained to be nonnegative.

The statistical results of our probit analysis are shown in Table 1. In the first and second columns, we show a comparison of the power of our near-term forward spread and the popular long-term spread (10-year yield minus 2-year yield) for forecasting future recessions. As shown in column (1), a one standard deviation decrease in the near-term spread increases the probability of recession by 47 percentage points, an economically large effect that is statistically significant with a p-value of 1 percent.[10] As shown at the bottom of the table, the average value of the model-implied probability of a recession was 64 percent for those observations when recessions did occur within the specified four-quarter window, compared to 10 percent average probability preceding periods in which no recession occurred. This suggests that the model delivers a fairly discriminating signal. In comparison, the results of specification (2) show that the 10-year/2-year spread delivers a smaller effect, along with a notably less accurate in-sample fit.

Table 1: Near-term Spread versus Far-term spreads for Forecasting Recessions

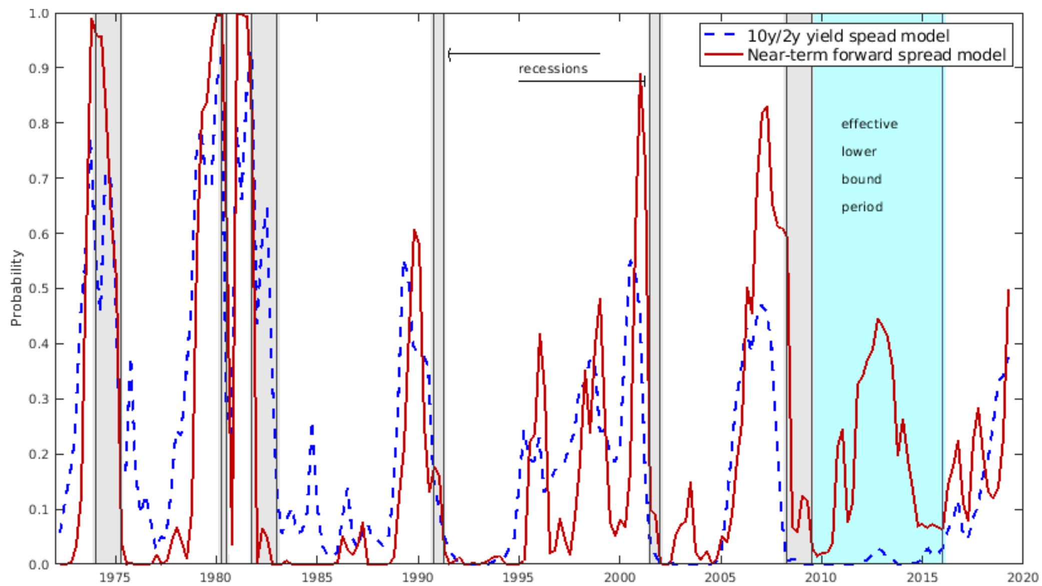
Explanatory variables	1	2	3	4	5	6	7	8
Near-term forward spread	-0.47			-0.40	-0.28	-0.40		-0.38
	(<0.01)			(<0.01)	(0.04)	(0.01)		(<0.01)
10y/2y yield spread		-0.36		-0.04				
		(<0.01)		(0.48)				
10y/1q yield spread			-0.41		-0.09			
			(<0.01)		(0.35)			
10y fwd / 6q fwd spread						-0.07		
						(0.23)		
Survey-expected prob. of recession							0.53	0.29
							(<0.01)	(<0.01)
mean fitted prob | future recession	0.64	0.50	0.62	0.64	0.64	0.65	0.58	0.72
mean fitted prob | no future recession	0.10	0.15	0.11	0.10	0.10	0.10	0.12	0.08

Notes. Data are quarterly 1972:Q1-2018:Q4. Observations for which the economy is already in recession are dropped, as are observations during the ZLB period (2009:Q1-2015:Q4). Results are for probit regressions in which the dependent variable is an indicator equal to 1 if the economy transitions from an expansion to a recession 1, 2, 3, or 4 quarters in the future. Reported are "sensitivity statistics." When the explanatory variable is a term spread, sensitivity is defined as the increase in probability of recession that is estimated to occur when the spread falls by one standard deviation from its unconditional average value, while the other explanatory variables remain at their unconditional means. For the survey-expected probabiltiy of recession, sensitivity is defined as the amount by which the probability of recession increases when the survey-based measure increases by one standard devation from its unconditional mean, holding all other explanatory variables constant at their respective means. Reported in parentheses are the bootstrapped significance level for a Wald test that the sensitivity is significantly different from zero (bootstrapped under null hypothesis of no predictability). The bottom two rows report the mean value for the fitted probability of recession in the quarter before (a) a recession occurs in one of the next four quarters, or (b) no recessionin the next four quarters.

[10] Our procedure for calculating bootstrapped p-values and standard errors is described in Appendix B.

Figure 3 shows the fitted conditional probabilities of recession from model (1), based on only the short-term spread (red line), compared to those from model (2), the more conventional model using the long-term spread (blue line). Generally, our model exhibits somewhat steeper spikes before recessions, again indicating a sharper prediction. Notably this model provides a clearer signal before the recent "Great Recession". As of the end of the sample period in early 2019 (and the time of this writing), the near term forward spreads forecasted a substantially elevated probability of a recession.[11]

Figure 3: Estimated Recession Probabilities



Source: Federal Reserve Bank of New York; Federal Reserve Board staff estimates.

Returning to Table 1, specification (3) considers an alternative measure of the long-term spread used in the academic literature, the 10-year yield minus the 1-quarter bill rate. This measure in some sense has more overlap with the near-term spread than does the 10-year/2-year spread. Indeed, the coefficient on the spread in model (3) indicates an effect closer in magnitude to our near-term spread, while the fitted probabilities indicate that this measure discriminates almost as well as the near-term spread.

In order to gauge whether both the near-term forward spread and a long-term spreads each contribute independent information for recession prediction, specifications (4) and (5) include one of the long-term spreads together with the near-term spread in a single model. In both cases, the magnitude of the coefficient on the near-term forward spread shrinks somewhat compared to its estimate in model (1), but it remains statistically significant. In contrast, the estimated marginal effect on the probability of recession of the competing long-term spread in both cases is

[11] The near-term spread model forecasts a higher probability of recession during the effective lower bound period because it does not reflect the monetary authority's desired trajectory for policy at that time, as the fed funds rate (and 3-month Treasury bill) could be lowered no further. The near-term spread thus arguably was not reflective of expectations for macroeconomic performance in this period.

economically small and not statistically different from zero. Also, fitted probabilities do not improve relative to model (1). This result indicates that essentially all of the information in the long-term spread is subsumed by the near-term forward spread.

Perhaps a better approach to testing whether there is additional information in the term structure beyond six quarters would be to instead include the 6-40 quarter forward spread, which is the complement to (i.e. does not overlap with) our near-term forward spread. Together, they span the slope of forward rates out to 10 years. Model (6) conducts such a test, the results of which appear to again reject the hypothesis that there is additional information in the term structure beyond what is contained in our near-term forward spread.

A final question we consider is whether the near-term forward spread contains information over and above that reflected in economists' forecasts. The Survey of Professional Forecasters (SPF) asks respondents to write down (four) probabilities that the U.S economy will be in recession in each of the next four quarters. Rudebusch and Williams (2009) tested for the relative forecasting power of the 10-year/1-quarter yield spread against each of these quarterly SPF survey recession probability forecasts and found that, particularly for three and four quarters out, the yield curve was clearly more informative about recession risk than the survey.

In models (7) and (8), we examine the predictive content of these survey-based recession probabilities in a univariate specification analogous to regressions (1) – (3). In particular, we construct a single proxy to gauge the expected recession probability over next four quarters, equal to the average of the (four) SPF survey probabilities of being in recession (one for each of the subsequent four quarters).[12] The survey is conducted during the first month of each quarter. We use the survey data from the month that follows the quarter when interest spreads are measured, so that survey respondents had access to all information reflected in market interest rates when their forecasts were provided. In model (7), where the only regressor is the survey probability, its coefficient is highly significant. By itself, its prediction accuracy exceeds that of the 10-year/2-year spread, but is inferior to that of the near term forward spread.

In model (8) we examine the marginal contribution of the near-term spread when we include it together with the survey recession probability forecast. Here, we find that the coefficient on the near-term spread is only modestly smaller compared to model (1), while both the spread and the survey probability are individually statistically significant. Since model (8) performance is notably better than models (1) and (7), we conclude that the near-term spread and the survey each bring some independent information to the table. That said, it is striking that the survey-based estimates do not subsume more of the information in the near term forward spread, given that the survey-based measure were collected a month later.

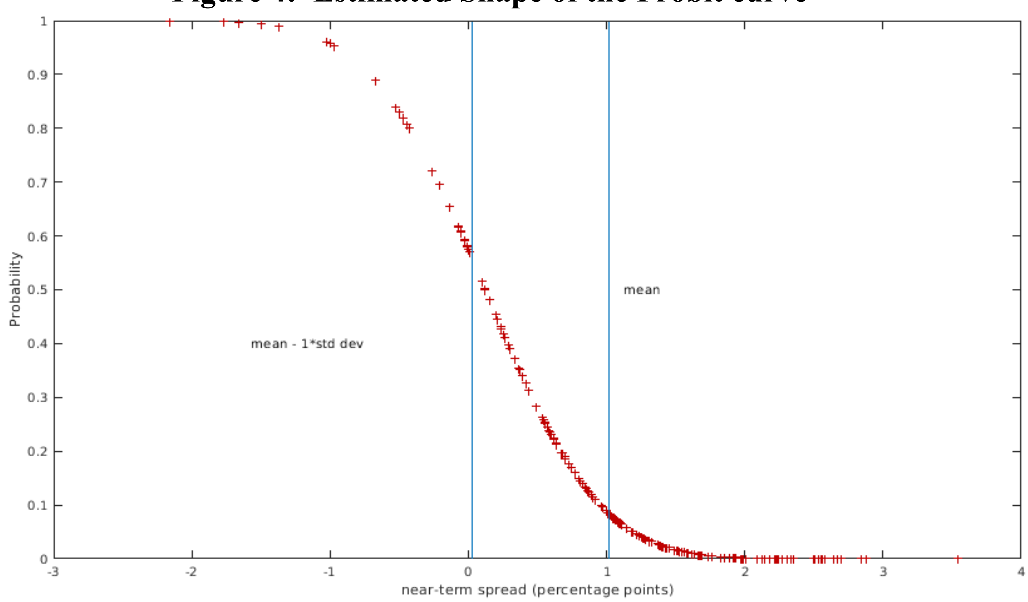
[12] We examined whether using a more flexible approach for including the four separate survey probabilities into the prediction model (rather than just their average) would indicate that these survey forecasts contained more information and found only very slight improvement to their forecasting power.

Extension to GDP Growth Prediction

Our finding of the statistical dominance of the near-term spread for forecasting economic activity might be attributable to the use of a probit framework, which accounts for only two possible outcomes—recession or no recession. If there is some additional signal embedded further out on the yield curve, perhaps it can be teased out when forecasting a more granular measure of economic activity. We consider this, focusing on predictions of GDP growth over the four quarters ahead, the same period over which the recession event was predicted. We also examine whether these spreads contain substantial information about future growth that is not reflected in SPF survey forecasts of GDP.

For predicting recession, the probit specification allowed for nonlinearity in the effect of spreads recession probability. For instance, as shown in figure 4, the probit estimation for the near-term forward spread implies that its effect is substantially larger when that spread in near zero. The result seems intuitive, given that the near-term forward spread would tend to turn negative when investors decide that the Fed is likely to soon switch from a tightening to an easing stance. Accordingly, we allow for a simple form of nonlinearity in the prediction regressions for GDP growth by including a linear term as well as a dummy equal to one when the spread is negative, thus allowing for a discontinuity in predicted GDP growth around a spread of zero.

Figure 4: Estimated Shape of the Probit curve



The first three columns (specifications (1) to (3)) of Table 2 show the results for GDP growth predictions for each of the three candidate spreads. In each case we include the spread in question plus a dummy for when that spread is negative. In model (1), using the near-term forward spread, we find that both terms contribute to predicting GDP growth. The coefficient on the spread indicates that a 1 point increase in the spread boosts expected four-quarter GDP growth by 0.69 percentage points. In addition, the coefficient on the dummy indicates that seeing a negative spread lowers expected growth by more than 2 percentage points. The adjusted R-squared indicates that this simple model explains 39 percent of the variation in realized GDP growth.

Table 2: Near-term Spread vs Far-term Spreads for Forecasting GDP Growth

Explanatory variables	1	2	3	4	5	6	7
Near-term forward spread							
level	0.69			0.80	1.28		0.31
	(0.08)			(0.09)	(0.10)		(0.47)
inversion dummy	-2.10			-2.03	-1.89		-2.30
	(0.05)			(0.05)	(0.09)		(0.03)
10y/2y yield spread							
level		0.07		-0.56			
		(0.85)		(0.29)			
inversion dummy		-2.61		-1.05			
		(0.01)		(0.29)			
10y/1q yield spread							
level			0.40		-0.59		
			(0.22)		(0.31)		
inversion dummy			-2.36		-0.73		
			(0.05)		(0.53)		
Survey expected GDP growth						1.04	0.66
						(0.01)	(0.08)
RMSE	1.75	2.04	1.91	1.70	1.71	1.95	1.63
adjusted R-squared	0.39	0.17	0.27	0.41	0.41	0.25	0.47

Notes. Data are quarterly 1972:Q1-2018:Q4 excluding the ZLB period (2009:Q1-2015:Q4). Results are for OLS regressions in which the dependent variable is real GDP growth over the subsequent four quarters, in percent. Bootstrapped significance level (under null hypothesis of no predictability) for the coefficients are reported in parentheses . The standard deviation of the dependent variable is 2.27 percent.

In comparison, using the 10-year/2-year spread in the analogous specification in column (2) explains only 17 percent of the variation in realized GDP growth. Here, only the dummy, with a coefficient of -2.6, contributes to the explanatory power. The explanatory power of the 10-year/1-quarter spread, specification (3), falls in between the first two regressions, with an R-squared of 27. The coefficient estimate on the linear term is 0.4 but is not statistically significant at conventional levels. In columns (4) and (5), we test whether knowledge of either long-term spread contributes to the GDP growth predication when the model also includes the near-term forward spread (and its indicator). When the 10-year/2-year is included in (4), neither the linear nor indicator terms are significant, whereas coefficients on the near-term spread variables are largely unaffected, and the adjusted R-squared increases only marginally from 39 percent to 41 percent. A similar result occurs when adding the 10-year/1-quarter spread. We again conclude that virtually all of the useful information for prediction is contained in the near-term term structure.

A final consideration on the information value of the near-term spread for predicting GDP growth is whether it reflects information different from the consensus survey forecasts of GDP growth.[13] Specification (6) shows the result from simply regressing four-quarter GDP growth on the consensus forecast for four-quarter GDP growth, using survey data collected toward the end of the first month of the first of the four quarters. (As with the recession probits, this gives a one-month informational advantage to SPF forecasters relative to the bond market.) As shown, the coefficient on the survey forecast is 1.04, close to unity and thus statistically unbiased and consistent with the traditional definition of forecast rationality. But the R-squared is only 26 percent, notably below the 39 for the near-term forward spread and its inversion indicator. The final specification, column (7), includes the forward spread and its inversion indicator together with the survey forecast. Here, both the inversion indicator and the survey forecast are statistically significant, though the latter with a p-value of only 8 percent. Nonetheless, the adjusted R-squared of 47 percent implies that economist forecasts contribute materially to the multivariate forecast.

In light of the strong predictive power of the near-term forward spread for economic outcomes, even relative to survey forecasts, our findings raise the question: Does this spread measure have power for predicting stock returns? And is there a beneficial portfolio strategy that would involve occasionally shifting out of equities and into cash, contingent on the near-term spread? Give that the onset of recession is generally preceded or accompanied by a substantial stock price decline, this seems a reasonable possibility; however, one might expect that much of the information driving movements in the yield curve would already be reflected in equity prices. Indeed, long-term spreads frequently have appeared in the set of conditioning variables used in the predictability literature and without much success.

Implications for Stock Market Predictability and Market Timing

In table 3, we examine the question of aggregate stock return predictability, by regressing annual (four-quarter) excess stock returns on interest spreads (each again measured as an average of the daily values over the quarter preceding the annual return period), the same periodicity of the tests for economic predictions. Of course, given the results for economic predictions, the near-term forward spread would seem to have the best shot at predicting returns. The first regression shows the results from predicting excess returns with both the level of the forward spread and the

[13] Rudebusch and Williams (2009) also touch upon this issue, but only with the limited goal of testing for "rationality" of survey forecasts, that is, whether the (10-year/3-month) term spread they use has any ability to predict GDP forecast errors. Indeed, they document some evidence against that null.

inversion dummy. The coefficient on the spread level is insignificant, but the dummy is highly significant with a coefficient of -22, and the adjusted R-squared is 0.16. Indeed, this inversion dummy exhibits impressive predictive power relative to other explanatory variables considered in the literature; the result implies a 22 percent lower average excess return when inverted, a very meaningful effect.

Table 3: Near-term Spread vs Far-term Spreads for Forecasting Excess Equity Returns

Explanatory variables	1	2	3	4	5	6
Near-term forward spread						
level	-1.53				-1.78	
	(0.59)				(0.53)	
inversion dummy	-22.69				-27.35	-23.59
	(<0.01)				(<0.01)	(<0.01)
10y/2y yield spread						
level		1.03				
		(0.73)				
inversion dummy		-5.40				
		(0.40)				
10y/1q yield spread						
level			1.76			
			(0.43)			
inversion dummy			-7.18			
			(0.40)			
Short rate				-1.85	-1.31	-1.36
				(0.15)	(0.29)	(0.26)
Earnings yield				2.69	3.28	3.28
				(0.05)	(0.02)	(0.02)
RMSE	16.22	17.70	17.35	17.22	14.90	14.96
adjusted R-squared	0.16	0.00	0.04	0.05	0.28	0.28

Notes. Data are quarterly 1972:Q1-2018:Q4, exclusing the ZLB period (2009:Q1-2015:Q4). Results for OLS regressions in which dependent variable is excess equity returnover the subsequent four quarters, in percent. Bootstrapped significance level (under null hypothesis of no predictability) for the coefficients are reported in parentheses. Standard deviation of the dependent variable is 17.96 percent.

As shown in column (2), neither the level of the 10-year/2-year yield spread nor its inversion dummy has any predictive power. Column (3) suggests some joint predictive power from the level and inversion indicator of the 10-year/1-quarter yield spread, but the adjusted R-squared of 0.04 is small relative to the near-term spread regression. Not surprisingly, adding either long-term spread measure to a regression with the near-term spread (not shown) does not boost the overall predictive power relative to model (1).

To make a connection with the broader return predictability literature, in column (4) we show results for more conventional predictors from that literature—the short interest rate and the earnings yield for the S&P 500 index (inverse of price-earnings ratio), each measured at quarter-end preceding the four-quarter prediction period. The earnings yield is fairly strongly significant and the adjusted R-squared from this regression is 0.05. In column (5) we add the near-term spread measures to the regression, and find that the inversion dummy coefficient is even larger and again highly significant. The earnings yield is also significant, with a positive coefficient, and the adjusted R-squared rises to 0.28, clearly at the high end of explanatory power found in such regressions. Finally, since the spread level and reversion indicator are presumably highly correlated, in column (6) we drop the level, which is not significant, to get a cleaner measure of the inversion dummy effect, which here gets a coefficient estimate of about -23.

Results of the return prediction tests point to a simple candidate portfolio strategy that might well dominate holding a portfolio with a constant exposure to the equity market. In particular, it suggests that a dominating strategy in any given quarter would involve being fully invested in equities if and only if the near-term spread was positive in each of the preceding four quarters. Otherwise, keep the portfolio invested in the riskless asset (Treasury bills).[14] Table 4 shows a comparison of portfolio performance statistics between always holding the equity market (row 1) versus a strategy of being out of equities if the spread has been inverted during any of the preceding four quarters (row 2). Rows 3 and 4 show the analogous strategies based on the inversion of the two long-term spreads.

Before examining the statistics, it is interesting to get a broad view of how these strategies would have performed cumulatively since 1972. Figure 5 shows the cumulative log returns to the market timing strategies versus the benchmark of simply holding the equity market. Each line plots the cumulative return, while the asterisks indicate quarters when the strategy keeps investors out of the market. As can be seen, for each spread inversion strategy, there are only five extended, and not surprisingly partially overlapping episodes when investors would have been out of the market. Most notably, while the near-term forward spread strategy emerges with only a slight advantage before 1990, after then, the strategy appears to outperform around the subsequent three recessions, racking up significant cumulative gains relative to the next best strategy.

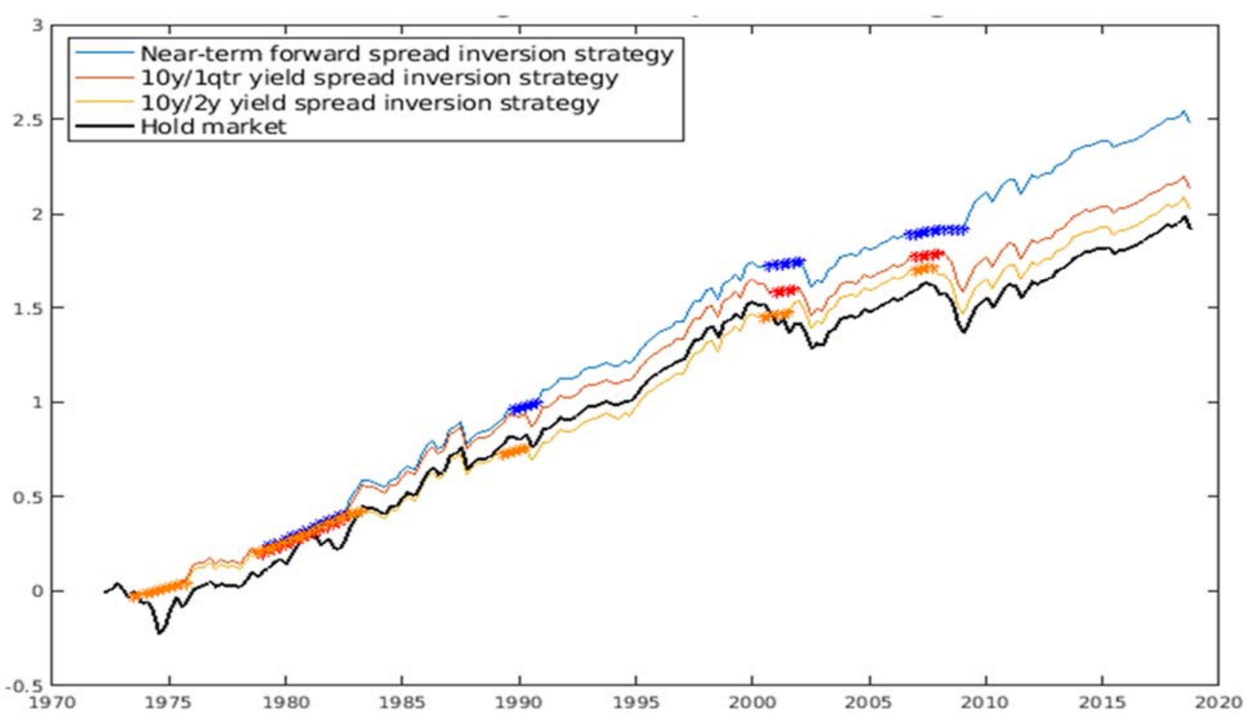
[14] In this setting, it would be easy to data mine and tweak the rule; thus, we first tried and stuck with this rule because it followed most naturally from our 4-quarter recession prediction framework.

Table 4: Performance of Market Timing Strategies

Strategy	percent "in"	mean	std	Sharpe	5th pct	95th pct	max drawdown
Market	100	1.58	8.60	0.18	-14.52	15.09	-46.33
		(0.52)	(0.73)	(0.07)	(1.83)	(1.35)	(5.61)
Near-term foward spread	77	2.13	6.65*	0.32	-9.04*	14.74	-27.85*
		(0.41)	(0.60)	(0.06)	(1.85)	(1.56)	(5.44)
10y/2y yield spread	77	1.58	7.06	0.22	-10.53	14.41	-39.69
		(0.46)	(0.72)	(0.07)	(1.99)	(1.74)	(10.46)
10y/1q yield spread	82	1.74	7.26	0.24	-11.05	14.74	-39.56
		(0.44)	(0.66)	(0.07)	(1.99)	(1.62)	(9.71)

Notes. Data are quarterly 1972:Q1-2018:Q4. Excess return statistics are reported in units of percent per quarter, except for "max drawdown," which reports the minimum observed percentage return over any four-quarter period. The row labeled "Market" refers to the strategy of always holding the market portfolio. Results for "Near-term forward spread" refers to holding the risk free asset if the near term forward spread has inverted in the current or any of the prior three quarters, and holding the market portfolio otherwise. Results for the other portfoios are defined similarly. The column "percent in" reports the fraction quarters during the sample for which the strategy was fully invested in the market portfolio. Boostrapped standard errors are shown in parentheses. Asterisks denote that the statistic differs from the corresponding statistic for the market portfoio at the 5 percent significance level.

Figure 5: Cumulative Log Returns to Market Timing Strategies



Notes. Data are quarterly 1972:Q1-2018:Q4. Cumulative log (base 10) returns are showns for various portfolio strategies. Asterisks for each line denote periods during which the portfolio strategy is invested in the risk free rate. At all other times, the strategies are invested in the market portfolio.

As shown in the first column, second row of Table 4, using the market-timing strategy based on the near-term spread results in being in equities on 77 percent of the time. Using this strategy based on near-term forward spread produces a mean quarterly return of 2.13%, which is 0.55% higher than the always-in strategy, or 2.2% at an annual rate. While perhaps material over the long run, the difference is not statistically significant relative to the high volatility of quarterly returns. On the other hand, the strategy based on forward spread inversion also results in less volatile returns, a return standard deviation of 6.65% as compared to 8.6% for always-in, a difference which is significant at the 5 percent level. Putting these results together, this strategy had a sharp ratio of 0.32 (per quarter), nearly double the 0.18 for always-in, though the gain is not statistically distinct from zero. A related benefit of the timing strategy is a trimming of the lower tail: Its 5^{th} percentile return is -9.04% compared to -14.52% for the benchmark. Similarly, the maximum four-quarter drawdown is about -27% versus -46%. Both of the latter two results are significant at the 5 percent level.

Turning to the results from strategies based on the other (long-term) spreads, it is notable that the 10-year/2-year spread strategy has the investor out of the market for the same percentage of time as the near-term forward spread, but with little of the benefit to mean returns. On the other hand, volatility shown in the third row falls in between that for the benchmark and that for the forward spread, but the improvement relative to benchmark is not statistically significant. A similar benefit also shows up, at least qualitatively, for the lower tail of returns. Lastly, the strategy based on the 10-year/1-quarter inversion has investors out of the market for a lower percentage of the quarters and produces only modest and statistically insignificant benefits for mean returns and risk.

One final (academic) point to note regarding return predictability results is the difficulty of squaring the "expected returns" implied by our regression with conventional rational asset pricing theory, that is, a world where higher conditional expected returns represent compensation for higher risk. Rather, the risk factor in our regression—the inversion dummy—appears to be a precursor of both higher risk and lower returns. Such a finding is not necessarily novel. For instance, many authors have found an insignificant or negative relationship between expected equity market returns and the conditional volatility of returns (see Glosten, Jagannathan and Runkle, 1993).In our case too, one would be hard-pressed to concoct an elegant narrative to support a rational risk pricing story. Instead, one might argue that the lower returns presaged by yield curve inversion were not truly expected, which seems plausible given the fairly small sample of episodes driving our statistical results.

Conclusion

In this paper we have documented that the near term forward spread subsumes essentially all of the information in other popular measures of term spreads when it comes to forecasting recessions and GDP growth, at least on an in-sample basis. We argued that this superior performance reflects that the near term forward spread is simply a cleaner measure of the expectations of market participants – a less distorted mirror. Moreover, the near term forward spread adds meaningfully to the accuracy of survey-based economic forecasts of real activity, particularly when using a more granular measure of activity such as GDP growth. Somewhat surprisingly, the near term forward spread also has significant information for predicting equity returns.

The strength and consistency of our findings may tempt one to treat these statistical relationships as a reliable guide for the future. But it is worth contemplating the assumptions required to extrapolate from our findings. First, it does seem reasonable to extrapolate from the narrow lesson in this exercise: Once we observe the near-term forward spread or a similar measure of the expected near-term trajectory for short-term interest rates, looking to the 10-year/2-year or other long term spreads will reveal little or no additional information about expectations for monetary policy or an economic slowdown in the year ahead.

Extrapolation may be more dicey, however, when considering the precision of the relationship between expectations for monetary policy easing and the likelihood of future recessions. The strength of that statistical relationship during the past 45 years reflects the observation of only six actual recessions, coupled with the relative paucity of false signals in which expected policy easing by market participants did not precede a recession. The most prominent false positive during our sample came with the anticipated easing triggered by the spread of the Asian financial crises in 1998, which did not result in a recession in the U.S. It is not hard to imagine that similar scenarios could generate additional false positives in the future. The near-inversion of the near term forward spread at the end of 2018 seems to have been associated with market perceptions of significant risks to the global economic outlook, including the threat of escalating trade disputes. Whether those risks manifest in a recession remains to be seen.

References

Bauer, Michael and Thomas M. Mertens (2018), “Economic Forecasts with the Yield Curve,” FRBSF Economic Letter 2018-07 (San Francisco: Federal Reserve Bank of San Francisco, March 5)

Benzoni, L., O. Chyruk and D. Kelley (2018), “Why Does the Yield-Curve Slope Predict Recessions?” Federal Reserve Bank of Chicago Working Paper 2018-15.

Estrella, Arturo and F. S. Mishkin (1998), "Predicting U.S. Recessions: Financial Variables as Leading Indicators," Review of Economics and Statistics 80(1), pp. 45-61.

Gilchrist, Simon and Egon Zakrajsek (2012), “Credit Spreads and Business Cycle Fluctuations,” American Economic Review, vol. 102 (June), pp. 1692–720.

Glosten, L.R., Jagannathan, R. and D.E. Runkle (1993), “On the Relationship Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks,” Journal of Finance, 48, pp.1779-1801.

Gurkaynak, Refet, Sack, Brian, and Jonathan Wright (2007), “The U.S. Treasury Yield Curve: 1961 to the Present,” Journal of Monetary Economics, vol. 54(8), pp. 2291-2304.

Johansson, Peter and Andrew Meldrum (2018), “Predicting Recession Probabilities Using the Slope of the Yield Curve,” FEDS Notes (Washington: Board of Governors of the Federal Reserve System, March 1).

Rudebusch, Glenn D., and John C. Williams. 2009. “Forecasting Recessions: The Puzzle of the Enduring Power of the Yield Curve.” Journal of Business and Economic Statistics 27(4), pp. 492–503

Appendix A: replicating and updating the near-term forward spread

Readers wishing to construct and update the near term forward spread for themselves can do so using the following procedure based on the methodology by Gurkaynak, Sack and Wright (2007, GSW henceforth), and data published by the Federal Reserve System.

1) Calculate the six quarter-ahead forward rate using the below formula from GSW, which specifies forward rates in terms of 6 parameters and the maturity of the forward date, *n*, in years:

$$f_t(n,0) = \beta_0 + \beta_1 \exp(-n/\tau_1) + \beta_2 (n/\tau_1) \exp(-n/\tau_1) + \beta_3 (n/\tau_2) \exp(-n/\tau_2)$$

   where $f_t(n,0)$ is the forward rate and the coefficients that determine the forward rate in each time period are $[\beta_0, \beta_1, \beta_2, \beta_3, \tau_1, \tau_2]$. The GSW formulation defines the *instantaneous* forward rate whereas we use the six quarter ahead *three-month* forward rate, which is an average of instantaneous forward rates over the interval of maturities from 1.50 to 1.75 years. To easily approximate the measure used in this paper, we advise calculating the instantaneous forward rate for the maturity at the midpoint of that interval, *n*=1.625 years. This yields a forward rate that only very rarely differs by more than one basis point from the forward rate used in this paper.[15] At the time of this writing, updates of daily values of the six parameters needed to calculate GSW forward rates are published by the Federal Reserve Board at the web address:

   https://www.federalreserve.gov/econresdata/researchdata/feds200628.xls

2) Obtain the three-month Treasury bill rate. At the time of this writing, this series can be obtained from the FRED Economic Data webpage of the Federal Reserve Bank of St. Louis:
   https://fred.stlouisfed.org/series/TB3MS

3) Calculate the near-term forward spread as the difference between the forward rate and the three month Treasury bill rate.

Appendix B: bootstrapping procedures

The statistics that we report from the probit and linear regressions in Tables 1-3 use overlapping observations. For example, for the observation of the dependent variable representing "any recession in the period from 1980:Q1 through 1980:Q4," the adjacent observation for the dependent variable represents "any recession in the period from 1980:Q2 through 1981:Q1." The overlapping nature of the data effectively reduces the number of independent observations and creates artificial serial correlation in the dependent variable. While it is valid to use standard

[15] In daily data from 1971-2018, the two measures differ by more than 1 basis point on only 0.13 percent of days.

techniques to calculate point estimates in this situation (maximum likelihood for the probits and OLS for the linear regressions, shown for instance by Hehedaard and Hodrick, 2016), the overlapping structure renders standard asymptotic inference invalid.

To produce p-values that are valid in this setup and more appropriate for our relatively short data set, we use a bootstrapping procedure. We first generate synthetic samples of the (overlapping) dependent variable by bootstrapping randomly, with replacement, from the data in blocks of 12 quarters. Bootstrapped samples each have the same length as our data sample. Independently, we block bootstrap the vector of explanatory variables. Putting together the bootstrapped dependent and explanatory variables, we create a synthetic data set in which the effects of overlapping data are present (e.g. artificial serial correlation in the dependent variable), but the model coefficients are zero in population, by construction, as is true under the null hypothesis of "no predictability." We calculate point estimates for each of 1000 such bootstrapped data samples using the same estimation procedure as for our actual data sample. Our inference is based on the frequency that the coefficients in the bootstrapped data exceed (in magnitude) those calculated in the data sample. For instance, if we find that the magnitude of a particular coefficient in our bootstrapped samples exceeds that from the actual data in only one percent of boostrapped samples, we report a p-value of 0.01.

We also calculate standard errors for the portfolio statistics in Table 4 and calculate p-values (denoted by asterisks) representing hypothesis tests for whether the statistics for the active portfolios differ significantly from those of the market portfolio. (Overlapping data is not used for portfolio statistics, except for the maximum drawdown statistic). For the standard errors of the univariate portfolio statistics, we simply bootstrap (in blocks of 12 quarters, with replacement) the portfolio returns, and calculate the standard deviations of the portfolio statistics across 1000 bootstraps. To test whether the portfolio statistics for the active portfolios are statistically different from the market portfolio, we use a different bootstrap procedure. We bootstrap, separately, two samples of the market return, and calculate the difference in portfolio statistics that occurs across the two bootstrapped samples of market returns. These portfolio statistics by construction, have zero difference in population. With 1000 pairs of bootstrapped market statistics in hand, we calculate the frequency with which the difference in the two synthetic market statistics exceeds what we observe in the actual data (the difference between the statistics for the market and for the active portfolio). For example, only if the observed sample statistic differs from the market statistic by more than 95 percent of bootstrapped differences under the null do we indicate the difference as statistically different from zero at the 5 percent level.

Reference

Hedegaard, E., and R. Hodrick, 2016, "Estimating the Risk-Return Trade-off with Overlapping Data Inference" Journal of Banking and Finance, ppg. 135-145.