Predicting Economic Performance With U.S. Treasury Yield Curve Rates

A Nonprofessional Examination of the Relation Between
Time Series, Mathematics, and Economics
July 6, 2020–Sept. 7, 2020

Edited by

Oren Merberg
Albany, NY
Alex Noonan
Rochester, NY

2020

1 Introduction

The spread between the 2-year Treasury yield 10-year Treasury yield is a commonly used macroeconomic leading indicator. Analyzing this spread as a function of time reveals its tendency to dip below zero before any major economic recession (depicted in Figure 1). The yield curve is *inverted* when the function is negative. In the past three decades, the yield curve inverted six times and each a recession followed. This is likely due to bond investors becoming pessimistic on growth in the medium- to long-term.

Predicting a recession is certainly useful, but it is not clear if the 2yr/10yr spread is able to predict other economic factors. Undoubtedly, it would be valuable for investors and traders to predict future deviations in the economy's performance. In this analysis, we set out to better understand the yield curve's behavior and its relation to economic factors.

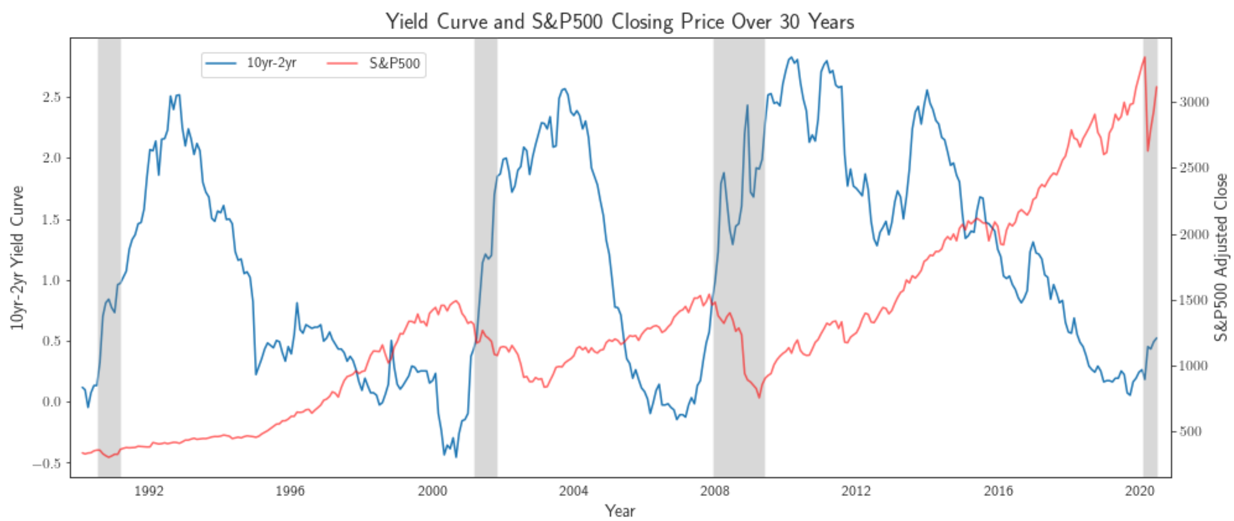

Figure 1: Comparison of yield curve and S&P 500 closing price. Shaded regions are U.S. economic recessions.

2 Data

The S&P500 [2] is an accurate representation for the performance of the economy, so we use this to test our predictions. Data for the S&P500 is freely available, ranging from 1927 to 2020. We remove observations before January 2, 1990 in order to match the Treasury's available yield curve data. The main variable of interest in the S&P500 is the `Adjusted Close` price. The data contains 7,681 observations (January 2, 1990 to June 25, 2020), which means there are 3,452 dates missing. To avoid errors, we re-index the dates to account for each day and fill missing values with the last valid date's value.

The U.S. Treasury Department provides extensive records of daily yield curve rates [1] on their website starting from January 2, 1990. To analyze the inverted curve, we take the difference between the 2yr and 10yr yield rates. The inversion start dates that we choose to analyze are 1990-02-08, 1998-04-28, 2000-01-05, 2005-11-29, 2007-04-05, and 2019-08-29.

To isolate instances of inversion, we take six (not necessarily disjoint) partitions to undergo separate analysis. Figure 2 shows both data sets for each separate time period.

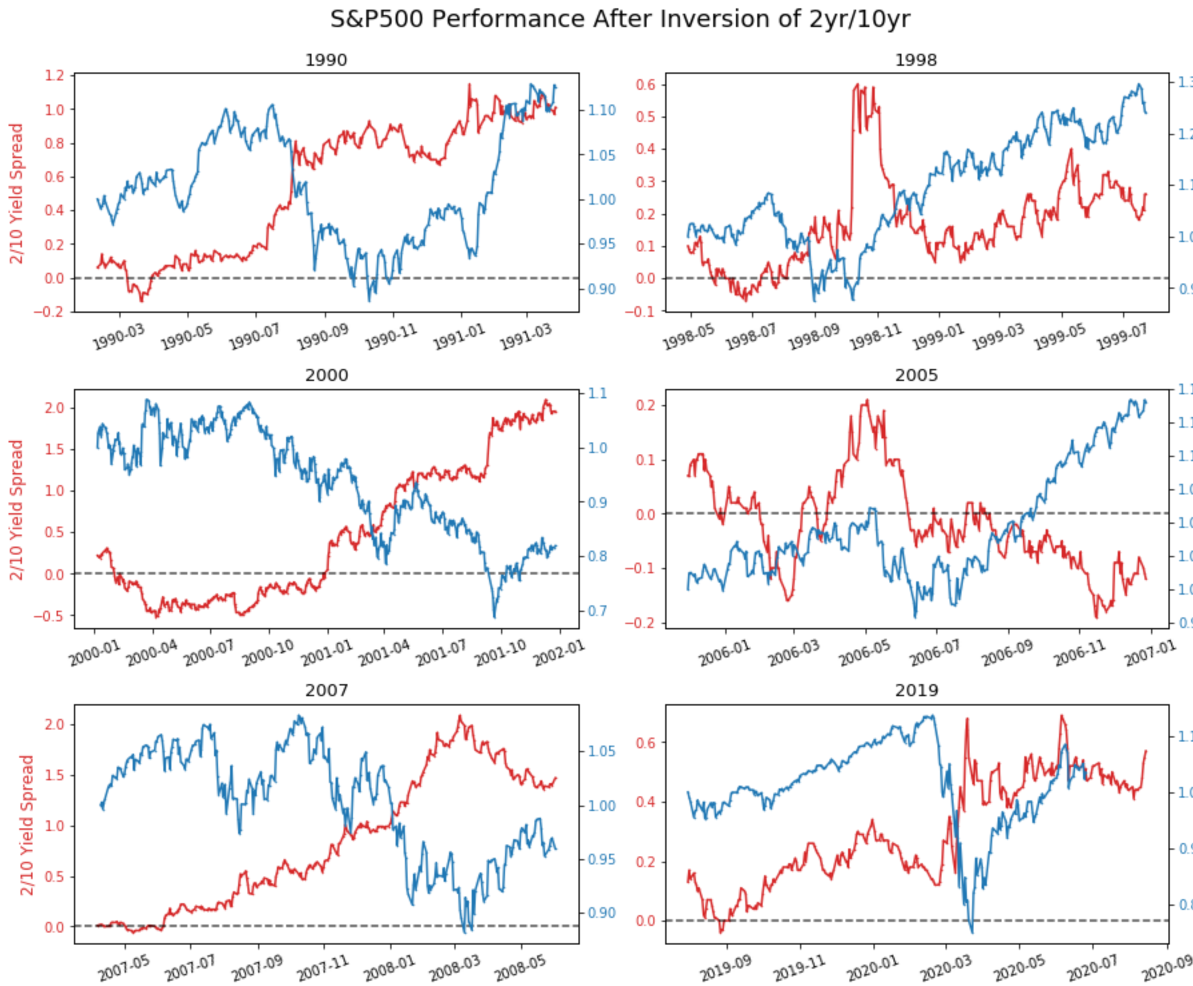

Figure 2: Comparison of yield curve and S&P500 closing price for each date range in analysis.

3 Formulation of Time Series Analysis

A time series is a collection of random variables that is indexed by time (also known as a stochastic process). We denote a time series with the following notation:

$$\{X_t : t \in T\},$$

where T is the set of time indices during the yield curve's inversion and X_t is the random variable denoting the value at each time t. For example, our first S&P500 subset would take the form $\{X_t : t = 1990/02/08, 1990/02/09, \ldots\}$ and similarly for the remaining sets of time.

It is important to also define the mean function and autocovariance function (ACF) for time series data. The formula for the mean function and ACF are

$$\mu_t = \mathrm{E}[X_t],$$
$$\gamma(t_1, t_2) = \mathrm{Cov}(X_{t_1}, X_{t_2}) = \mathrm{E}[X_{t_1} X_{t_2}] - \mu_{t_1}\mu_{t_2},$$

respectively, where $\mathrm{E}[X_t]$ is the expected average value of $\{X_t\}$ and $\mathrm{Cov}(X_{t_1}, X_{t_2})$ is the covariance of the stochastic process at two different moments in time. Notice that evaluating the ACF at the same moment in time reduces to the normal variance.

3.1 Stationarity

Unlike many distributions, the mean function and ACF are functions of time. These fundamental statistical properties are integral to predictive analysis (it would be difficult to model a Gaussian distribution with a varying mean). For this reason, most time series models assume a time-independent mean, variance, etc. This concept is called *stationarity*. In our case, we require the mean and autocovariance function to be independent of time, concisely stated in the following conditions:

$$\text{i) } \mu_t = \mathrm{E}(X_t) = \mu$$

$$\text{ii) } \gamma(h) = \mathrm{Cov}(X_{t+h}, X_t)$$

From Figure 2, it seems clear that the time series have varying means and covariances and are therefore non-stationary. However we want to be sure, so we implement the augmented Dickey-Fuller (ADF) test, described in the next section.

3.1.1 Dickey-Fuller Test for Stationarity

The ADF test belongs to a class of statistical significance tests called Unit Root Tests. If a unit root exists in a time series model, then it contains *stochastic trend* and will not converge back to its expected value. Thus, a unit root implies there is not a constant mean, and the process is non-stationary [5]. The test is set up with the following null and alternative hypothesis:

$\mathbf{H}_0$: A unit root exists.

$\mathbf{H}_1$: A unit root does not exist.

For this test, a p-value less than a significance level of $\alpha = 0.05$ means that we reject the null hypothesis and the time series is stationary. We implement the test on our six data sets and the results are listed in Table 1.

Table 1: P-values returned in the ADF test for each year on non-stationary time series.

	1990	1998	2000	2005	2007	2019
S&P500	0.735952	0.788226	0.640841	0.956666	0.482431	0.120096
Treasury	0.802514	0.119207	0.986977	0.338093	0.808354	0.602999

Indeed the ADF test outputs test-statistics that are greater than any critical value, resulting in high p-values, and we fail to reject the null hypothesis for each year.

3.1.2 Transforming Non-stationary Time Series

A standard method to make a non-stationary time series stationary is *differencing*. It is a simple and effective way to stabilize the mean function by essentially removing the time-dependent components of the stochastic process. We define the first-order difference, I(1), as the following transformation

$$\{X_t\} \to \{X_t - X_{t-1}\}$$

Figure 3 and Figure 4 show the effect that differencing has on our time series data.

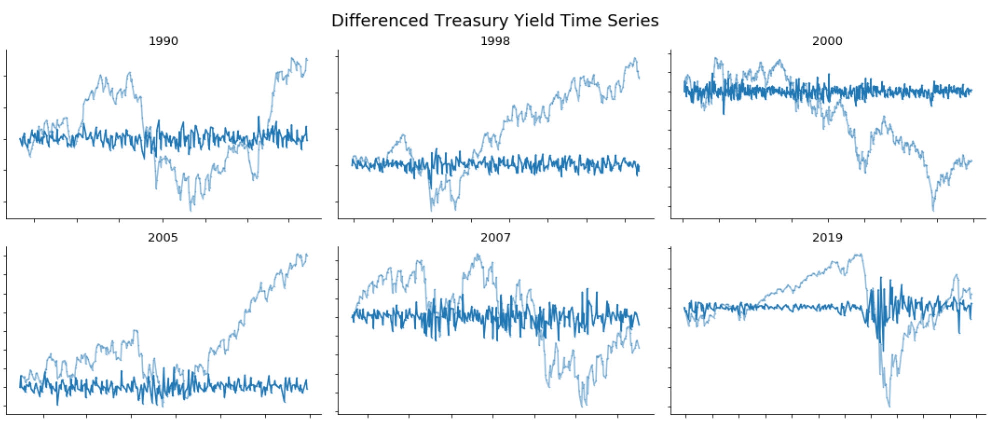

Figure 3: Original Treasury yield curves (light) overlayed with differenced yield curves (dark).

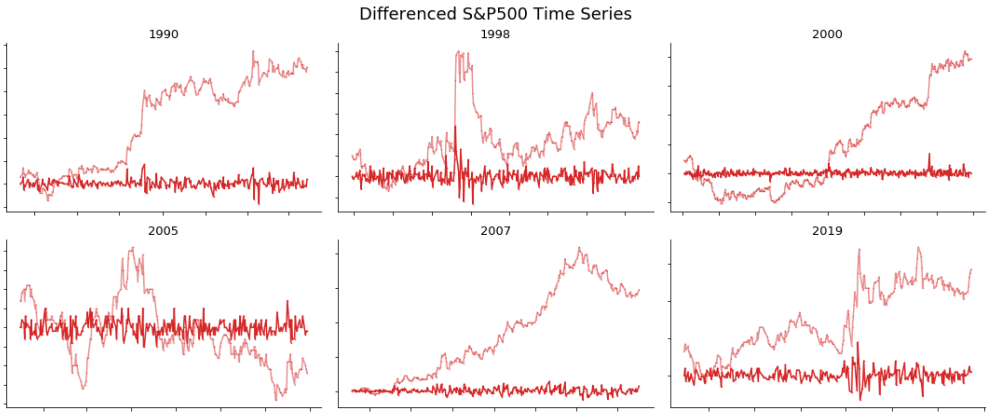

Figure 4: Original S&P500 data (light) overlayed with differenced S&P500 data (dark).

Notice the average value of each time series seems to flatten, which is consistent for a constant mean and stationarity. Again, we run the ADF test to confirm. Table 2 lists these results.

Table 2: P-values returned in the ADF test for each year on differenced time series.

	1990	1998	2000	2005	2007	2019
S&P500	$9.52 \cdot 10^{-10}$	$2.77 \cdot 10^{-30}$	$5.38 \cdot 10^{-30}$	$8.81 \cdot 10^{-26}$	0.00	$4.33 \cdot 10^{-03}$
Treasury	$9.68 \cdot 10^{-21}$	$1.06 \cdot 10^{-04}$	$2.07 \cdot 10^{-30}$	$6.17 \cdot 10^{-30}$	$6.04 \cdot 10^{-19}$	$6.60 \cdot 10^{-08}$

With the p-values much less that $\alpha = 0.05$, we can confidently reject the null hypothesis (no unit roots exist) and say that the data is stationary.

4 Granger's Causality

Determining correlation between two time series is generally complicated. A method that is often used in these situations is called Granger's Causality Test. The use of *causation* is misleading and requires some clarification in this context. Normally, if an event X causes the event Y, then X is responsible for Y happening and Y has some dependence on X occurring. For example, someone flipping the switch on a lamp causes it to light up. On the other hand, Granger-causality tests how well one time series forecasts another time series. More precisely, it determines if past (lagged) values of X can explain the variance in Y [3]. For example, someone walking into a dark room Granger-causes the lamp to turn on. The mathematics and subtleties behind the theory are beyond the scope of this paper, so we omit some details.

4.1 Implementation and Results

To test the potential strategies that could benefit from any identified relationship we measured the total return of one dollar (reinvested dividends) in the S&P 500 for a year after the inversion event as well as the monthly standard deviation of returns 3 months before and

12 months after the inversion. The hypothesis test for our data is set up like the following:

$\mathbf{H}_0$**:** X_{Yield} does not Granger-cause X_{SP500}.

$\mathbf{H}_1$**:** X_{Yield} does Granger-cause X_{SP500}.

where X_{Yield} represents the inverted Treasury yield curve and X_{SP500} represents the S&P 500 returns for each time period. Hereafter, we represent Granger-causation with an arrow (X→Y).

There are numerous adjustable parameters for this test, most importantly the number of lags to consider. We take an exhaustive approach by running the test with all lags up to 12 and taking the minimum returned p-value. Furthermore, we test the reverse case ($X_{\text{S\&P500}} \rightarrow X_{\text{Yield}}$) as it may provide deeper insight. A chi-squared test is appropriate when including many variables, however, we are dealing with small samples, so an F-Test is more appropriate. The results are listed in Table 3 with the p-values for each year.

Table 3: P-values returned in Granger's Causality F-Test for each year on differenced time series.

	1990	1998	2000	2005	2007	2019
$\mathbf{X_{Yield} \rightarrow X_{S\&P500}}$	0.0273	0.1599	0.0045	0.0265	0.0009	0.0
$\mathbf{X_{S\&P500} \rightarrow X_{Yield}}$	0.1808	0.0577	0.7976	0.0072	0.2762	0.0

4.2 Interpretation

One has to be careful when interpreting the results of the Granger's Causality Test. The results may vary depending on different factors and neglects latent confounding effects (influences on both dependent and independent variables). Another restriction is non-linear correlations as the test only considers linearity [4]. With this caveat, we proceed with our interpretation of the results.

As usual, our confidence level is $\alpha = 0.05$ and we reject the null hypothesis for any p-value below this threshold. Accordingly, we find that the inverted Treasury yield curve has some effect on the S&P500 returns for the years 1990, 2000, 2007, and 2019. In the reverse case, we find Granger-causation for the year 2007. However, with the uncertainty of this test, we cannot say with certainty that there is a direct relationship for any year. More appropriately, we can say that the inverted Treasury yield curve has some statistical significance in forecasting the S&P500.

References

[1] Daily Treasury Yield Curve Rates. https://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yieldAll.

[2] SPDR S&P 500 (SPY) Stock Historical Prices & Data. https://finance.yahoo.com/quote/SPY/history/.

[3] Stephanie Glen. Granger Causality: Definition, Running the Test. https://www.statisticshowto.com/granger-causality/, Sep 2019.

[4] Sarit Maitra. Time Series Forecasting Using Granger's Causality and Vector Auto-regressive Model. https://towardsdatascience.com/granger-causality-and-vector-auto-regressive-model-for-time-series-forecasting-3226a64889a6.

[5] Selva Prabhakaran. Augmented Dickey-Fuller (ADF) Test - Must Read Guide - ML+. https://www.machinelearningplus.com/time-series/augmented-dickey-fuller-test/.