The Term Structure and World Economic Growth: A Retrospective and 30 Years of Out-of-sample Evidence

Campbell R. Harvey

KEY FINDINGS

- The slope of the yield curve predicts economic growth and inverted yield curves precede recessions.
- The out-of-sample evidence shows that yield curve inversions preceded the four most recent recessions and there were no false signals.
- The length of the inversions closely matches the length of recessions providing further evidence of the usefulness of this leading indicator.

ABSTRACT

The inaugural issue of The Journal of Fixed Income led with an article by a junior professor at Duke University (me) who linked the slope of the yield curve—the difference between long-term and short-term yields—to future economic growth. Thirty years later, I assess the out-of-sample performance of the yield curve indicator. Four recessions have come and gone since my article was submitted to The Journal of Fixed Income in 1990. Each time the yield curve has inverted prior to the recession. I also provide some additional background regarding the genesis of my idea.

When a researcher publishes a finding, one of two outcomes often follow. The good outcome is that after further examination and testing the effect becomes weaker. The bad outcome is that the effect goes away. In Harvey (2017) I detail the many reasons that can cause a bad outcome. Often the model is overfit in the original research, which may include over-parametrization and potentially strategic data selection, such as the exclusion of influential observations.

In Harvey (2019), I explain other reasons research may not replicate. Structural changes in the economy might render the in-sample estimates unmeaningful. Also, when working with quarterly macroeconomic data, the paucity of observations leads to another possibility: the finding could just be lucky and as such will not replicate.

In “The Term Structure and World Economic Growth” in 1991, I showed that the slope of the yield curve had the ability to forecast real economic growth in a number of countries. The evidence was particularly strong in the United States. At that time, however, I included only four US recessions in my research—and it is challenging to draw inference from four observations. This article serves as an out-of-sample analysis of the original publication.
As I will show, the very simple model has fared remarkably well in 30 years of out-of-sample scrutiny. Since my paper was submitted to *The Journal of Fixed Income*, the US economy has experienced four recessions. The model provided early warning for each recession—including the global financial crisis. Remarkably, the model has not produced any false signals (yet).

**THE GENESIS OF MY YIELD CURVE IDEA**

In June 1982, I showed up for my summer internship in Toronto, the same month that *Blade Runner*, *E.T.*, and *The Wrath of Khan* were released, offering plenty of distractions for those interested in science fiction. I had two summer job opportunities, both in corporate development. One was at a prominent media company, Torstar Corporation, and the other at the largest copper miner in the world at the time, Falconbridge Ltd. I chose Falconbridge. I had no significant work experience and only the first year of a master’s program to show.

I was given a singular—but highly consequential for my career—task: build an econometric model to forecast real GDP growth. At the time, I was naive and took the task in stride as “no big deal.” From my perspective today, I find it shocking that a firm of this stature with operations in 18 countries would put a student in charge of building a forecasting model that could become a key input for their decision-making.

Having an accurate indicator of future economic growth is crucial for the mining industry and especially important for copper, whose price closely tracks the economy. Important decisions such as exploration expenditures, new mine construction, mine production levels, and mine closings depend on the future price of copper, which in turn is driven by the future state of the economy.

Many companies, such as Data Resources Inc. (DRI), Chase Econometrics, and Wharton Econometric Forecasting Associates (WEFA), made a business out of forecasting GDP (and other variables). These organizations employed hundreds of Ph.D. economists and data scientists. Corporations and organizations would subscribe to these expensive services to get forecasts they could use in their internal planning. I appeared to be a one-person substitute—and at a bargain price.

I knew the competitive landscape and realized it would be impossible for me, in the few months of my internship, to gather comprehensive macroeconomic data, assemble computing resources, and successfully specify and estimate a full-blown simultaneous-equation model of the US economy. I needed a fresh approach. Further, I was very skeptical of the macro-econometric models because 1) the data were stale and mainly reflected past economic activity; 2) the data were smoothed with the ARIMA X-11 seasonal adjustment program, which used forward information; and 3) the data were subject to revisions, and at the time, only the revised data were available to researchers, so data were not available in real time.

I had to keep it simple because a complex model would just compete with companies such as DRI and WEFA. My initial idea was to look at financial assets. I learned in my finance courses that the value of an asset today reflects expectations of its future cash flows. Perhaps I could extract useful information about expectations of future economic growth from the changes in an asset’s price. Indeed, evidence was plentiful that stock prices plummeted on bad news that held the possibility of the economy tanking.

I began my research examining whether the returns of the S&P 500 Index predicted future economic growth. Sometimes it did and sometimes it did not. The equity market was unreliable. I discovered an academic literature led by a professor at the University of Chicago, Eugene Fama (1981), that disentangled the mechanisms linking stock returns and “real activity” as he called it.
While a promising idea, the bottom line was no different in 1982 than when Paul Samuelson wrote in 1966 that “the stock market has predicted nine of the past five recessions.” But what was wrong? The theory seemed to clearly link expected future cash flows to the economy.

I thought of three reasons why the stock market might be generating so many false signals. First, stock prices reflect future expected cash flows, but these need to be discounted to present value. The discount rate reflects risk changes, and variation in the discount rate can lead to false signals. Second, dividends are uncertain and can change through time, adding another dimension of noise. Third, there is uncertainty about the time to maturity for the valuation. Each of these factors can generate false signals.

I abandoned stock returns and turned to the bond market. Two pieces of research got my attention. The first was a 1966 white paper by Reuben Kessel (1956), a Fed economist, who showed that the difference between long-term rates and short-term rates appeared to move with the business cycle. Second, I read a 1981 New York Times article that mentioned Galen G. Blomster, an economist at Northwestern National Bank in Minneapolis, who was studying AAA-bond yields and commercial paper yields. I thought these corporate bond yields were better than stocks, but suffered from the same problems as stocks because fluctuations in risk not linked to the economy could add noise.

I settled on the US Treasury term structure and ex ante thought it solved many of the problems that plagued the stock return indicator. First, the bonds were relatively risk free. Indeed, the US Treasury bill in my finance courses was referred to as the “risk-free rate.” Second, unlike dividends, coupons are fixed in size. Finally, the maturity of the bond is known. Surely, these government bonds could give us a clearer view of the future.

I also knew that nominal bond yields contain three components: expected inflation, expected real rates, and a risk premium. The risk premium was negligible for US Treasury bonds. Therefore, to get a view of the future, I could just subtract the long-term and short-term rates. The approach still was not clean, however, because the spread represented a combination of the term structure of inflation and expected real rates.

My initial estimations were very encouraging. I worked hard to put together a presentation for senior management.

Then something unexpected happened. I was laid off after five weeks of work. It wasn’t just me. The entire corporate development department was let go. As we always learn from the NBER well after the fact, we were in a recession, the second part of the so-called double-dip recession. The first dip was in 1980 and the second dip continued until the end of 1982.

At the time, I understood corporate downsizing, yet it seemed unusually harsh to sack the summer intern, especially in Canada. Again, young and naïve, I drafted a letter to Falconbridge’s board of directors arguing they had made a mistake. I later heard they were going to offer me something else within the firm for the rest of the summer—but not after the letter!

In the end, Falconbridge never got to see my model. After a number of years of economic difficulties, the company was acquired by Xstrata in 2006. Xstrata was acquired by Glencore in 2013. I joined a start-up, Business Planning and Control, with a fellow Falconbridge employee in corporate development and focused more on accounting and risk management systems. The start-up did not keep me very busy, so I continued working on my term structure idea.

I returned to academia in September 1982. Working as a research assistant for two professors, I did the estimation for one of their papers, Amoako-Adu and Ben-Zion (1983), and in doing so learned what an academic paper should look like.
I built up the courage to show them the draft of my yield curve paper. I first showed it to Professor Ben Amoako-Adu and he was impressed. Indeed, an early version of the paper was a class project coauthored with fellow student Barbara Jill Fox. He shared the paper with some of his colleagues. They encouraged me to continue to work on the paper as a special course, actually as two courses. They also, surprisingly (to me), said that I needed to apply for a Ph.D. in finance.

I did not know much about Ph.D. studies. I was the first person in my immediate family to get a bachelor-level degree. The master’s degree seemed more than adequate, but the research bug had bitten me. I liked what I was doing. I wrote several research papers in my second year, two on the yield curve and one that optimized the weights on the US index of leading indicators.

I applied to various programs and included copies of my research papers. In hindsight, I now realize I had lucked into the ideal application. I was showing evidence that I was able to pose economic questions, test hypotheses, and do empirical work before my doctoral studies.

I chose to accept the offer from the University of Chicago because of its very strong reputation and because one particular researcher, Eugene F. Fama, had ideas that overlapped with mine.

THE REFINEMENT OF THE IDEA AT THE UNIVERSITY OF CHICAGO

My first day at the University of Chicago I transferred my data and programs to their servers. The usual process is to take courses for a few years and then start doing research. I was ready to go on the research front immediately.

In real time, I had no idea how lucky I was to be at the University of Chicago in the early 1980s. My professors included Robert Lucas, Gary Becker, Lars Hansen, Eugene Fama, and Sanford Grossman. Since I left Chicago, all but one (as of writing) have been awarded the Nobel Prize. I also collaborated with another future Nobel recipient, Merton Miller, who served on my dissertation committee, which included Eugene Fama and Lars Hansen.

I was fascinated by the observation that whether I was taking a macro or micro course the theoretical framework was very similar. Indeed, I remember one particularly influential homework assignment in Sanford Grossman’s price theory course. While working on the assignment, I realized that the framework could be adopted to link the term structure of interest rates to expected consumption growth: if longer-term real rates are very low and today’s rates are higher, that means consumption growth is expected to slow.

Even such an empirical finding was not sufficient for a dissertation at the University of Chicago. Further, it was not acceptable to concoct a theory ex post that “explained” my finding. I relied on the consumption-based asset-pricing framework originally proposed by Robert Lucas (1978) and Douglas Breeden (1979). These models link expected real rates to expected consumption growth.

After two years at Chicago, I had a draft of my dissertation and successfully defended my proposal. Even though my sample included only four recessions, my committee was swayed by three factors. First, the empirical findings made sense given the underlying theoretical framework. Second, my model successfully forecast the double-dip recession in 1980 and 1981–82, whereas other commercial models did not pick up the two recessions. Third, my model was as good as, or better than, the commercial econometric services for which a subscription cost thousands of dollars. The cost of my forecast was the cost of The Wall Street Journal which, at the time, was 25 cents.
I was on the job market early—in my third year. At the time, again having very little experience in academia, this did not seem unusual to me. My chair Eugene Fama was fine with me going on the market. In the end, I was recruited to Duke University by Douglas Breeden. He had recently relocated to Duke from Stanford and his research interests intersected with mine.

I defended my dissertation in the fall of 1986 (Harvey 1986). Over the next several months, I carved out a concise version of the paper and sent it to the Journal of Financial Economics. I received a revise-and-resubmit with a very tough to-do list from the referee. My research was eventually published in 1988 (Harvey 1988).

**THE FIRST REAL-WORLD TEST OF MY MODEL**

My model was challenged by the stock market crash in October 1987. At the time, the widespread belief of economists and professional forecasters was that the United States was headed to serious recession in 1988. Newspaper headlines warned “Recession likely to follow crash.” Invited to speak at a conference in late October, I was the contrarian among the conference presenters, most of whom shared the opinion of the newspapers. I explained that the stock market was an unreliable indicator and that the bond market was far more reliable. Further, the bond market was not indicating a recession. My forecast was that real economic growth in 1988 would be a positive 4.2%, which was not just a contrarian forecast, it was an outlier. More than a year later we would learn there had been no recession; real GDP growth in 1988 was 3.8% and the growth in real personal consumption expenditures was 4.6%.

By this point, I had a number of academic projects in the works, including an unrelated paper for which I had received a revise-and-resubmit from the Journal of Financial Economics and a co-authored paper with Wayne Ferson (who was on my dissertation committee) that was heading to the Journal of Political Economy. I decided to expand the messaging of my yield curve paper to different audiences. I knew these follow-on papers would not be counted in my tenure review case, but nevertheless I wanted to communicate my ideas beyond academia and to different geographies.

In 1989, and with the encouragement of Professor Robert Whaley, I published a paper in the Financial Analysts Journal comparing the GDP-forecasting power of the yield curve to that of the stock market (Harvey 1989). While my dissertation research focused on consumption growth, I had looked at the robustness results applied to both GDP and other components such as investment expenditures. Further, personal consumption is about 70% of GDP and is highly correlated with GDP growth.

I also decided to do country studies and published papers in Canada, Germany, Italy, and France (Harvey 1991a,b; Harvey, Kaul, and Kirby 1990, and Harvey 1997). My favorite of this batch of four papers is the Canadian application. It seemed trivial to show that the US yield curve was useful in forecasting Canadian growth because Canada’s GDP is highly correlated to US GDP. My paper took a different angle. Could the difference between the Canadian and US yield curves forecast the difference between Canadian and US growth? The results suggested this was the case.

My research also caught the attention of the media and was featured in The New York Times (Silk 1989) on three additional occasions in the 1989–1991 period. Again, at the time, I did not think this was a big deal, but in hindsight I realize the media attention was very fortunate for me.

In 1990, Douglas Breeden became founding editor of a new journal, The Journal of Fixed Income. He asked me to contribute an article for the premier issue in June 1991, which became “The Term Structure and World Economic Growth” (Harvey 1991c). I am grateful he selected it as the lead article in that issue.
THE OUT-OF-SAMPLE EVIDENCE

“The Term Structure and World Economic Growth” was effectively the end of my research on the yield curve and economic growth. As mentioned earlier, the first out-of-sample test was a “no recession” call, which turned out to be correct. What about the realized recessions?

Since the publication of my dissertation, the US economy has been in recession four times. The yield curve inverted before each of these recessions.

Exhibit 1 provides some out-of-sample evidence on the predictive ability of the yield curve, both in and out of sample. The yield curve, as measured by the difference between the 10-year Treasury bond yield and the 90-day Treasury bill yield, inverted before each of the last four recessions.

Importantly, Exhibit 1 also shows there were no false signals, although there was a close call in 1998, which coincided with the LTCM event that put markets on the edge and caused a serious illiquidity event mitigated only after the Federal Reserve Bank of New York intervened. The 1998 flattening is exactly what the model would predict ex ante due to the extreme economic and market uncertainty often associated with recessions.

Also of interest are the lead times to the recessions, which are shown in Exhibit 1. Notice that the lead time is variable. Exhibit 2 explores the lead time and links the duration of the recession to the duration of the inversions.

Notice the close relation between the length of the inversion and the length of the cycle. This is particularly evident in the last four recessions, which make up the out-of-sample period. Over the entire sample, the average absolute difference between the inversion length and the recession length is only four months. In the out-of-sample period, the average difference is only two months.

EXHIBIT 1
Forecasting Record of US Treasury Yield Curve Showing Yield Curve Inversion Lead Time to Recessions

NOTE: The yield curve is measured by the difference between the constant maturity 10-year Treasury bond yield and the 90-day Treasury bill yield. The Treasury bill is converted from discount basis to yield to maturity. Source: Federal Reserve Bank of St. Louis.
The theory makes a number of predictions that include the advance warning of slowdowns as well as a match between inversion length and cycle length. The fact that the inversion lengths are close to the cycle lengths provides additional evidence that this finding is not a fluke.

THE GLOBAL FINANCIAL CRISIS RECESSION

The yield curve gave plenty of advance warning of the global financial crisis, with a lead time of 21 months. Yet, at the time, very few were interested in the yield curve signal. While the Fed had published a number of follow-on papers detailing the predictive ability of the yield curve, the Fed largely ignored the inversion. Consider the following email that I received in 2019:

“…the Fed (mainly the economists at the Federal Reserve Board) mistakenly dismissed the yield curve inversion in 2006. In the fall of 2006, I met with Tim Geithner every week to give him a briefing of the market development. Most of the meetings involved the interpretation of yield curve inversion. That was when I read your paper for the first time. I have been a fan of your paper since then, but the Board dominated the view in the Fed system at the time. The Fed dismissed the market signal about recession and kept the target rate for another whole year at 5.25%, until the fall of 2007! A clear mistake in retrospect.”

For the record, in March 2006, the yield curve inverted. At the time, the fed funds rate was 4.5%. By July 2006, the rate had been increased to 5.25%, where it remained until the end of the inversion. In August 2007, the rate was decreased by 25 basis points. When the recession began in December 2007, the rate was 4.25%.

THE COVID-19 RECESSION

In June 2019, the yield curve inverted, triggering a recession warning, but June 2019 was very different from March 2006. By 2019, the media had figured out that the yield curve had predicted seven of the last seven recessions—without a
false signal. As a result, the yield curve message was under intense scrutiny and drew considerable press attention.

The media often focused on whether the yield curve inversion was “causing” a recession. The theoretical background of the model involves no causality. The slope of the yield curve simply and efficiently reveals the expectations of real economic growth in the market. Effectively, the yield curve slope aggregates expectations. Again, this explains many episodes including the “non-recession” of 1998–1999.

The intense media attention today, however, suggests it is possible to make the case for some degree of causality coming from outside the theoretical framework. Given the yield curve’s track record, its signal is hard to ignore. After the global financial crisis recession nearly 15 years ago, CEOs could credibly tell investors they were surprised by the downturn and could rely on their unpreparedness to explain poor corporate performance or even financial distress.

By 2019, however, the media attention to the predictive ability of a yield curve inversion made it hard for any CEO to declare being blindsided by a recession in 2020 or 2021. Consequently, receiving a recession red flag naturally led to companies’ scaling back capital investment and hiring plans. These actions are best thought of as risk management, but risk management that involves decreased investment spending decreases economic growth. Hence, we have a feedback loop and a credible case for some degree of causality.

We can also look at causality another way. Financial distress is very costly and not just to shareholders. No one wants to be laid off and be forced to collect benefits, potentially for an extended period of time. Along with some causality is a benefit. If companies and consumers make sure they have adequate cash balances when they see yield curve inversions, they have a greater chance of weathering the storm. Essentially, risk management may reduce the volatility of economic outcomes.

In March 2020, we entered a recession, but this recession did not have economic roots. Whereas we can credibly blame the financial sector for taking extreme leverage in the run-up to the global financial crisis, the catalyst in 2020 was a biological crisis. The pandemic the world experienced was analogous to a natural disaster. No single economic sector was to blame, although sectors were differentially impacted by the crisis. The downturn was swift and brutal and, in the end, short.

We will never know the counterfactual, that is, if COVID-19 did not strike, would we have had a recession? In the summer of 2019, economic prospects were grim. The June 2019 Duke CFO Survey (Duke CFO Survey 2019)\(^1\) reported that 69% of US CFOs expected a recession to begin by the end of 2020 and 84% expected a recession to begin by the end of 2020 or into the first half of 2021.\(^2\)

To be clear, the slope of the yield curve obviously did not forecast a pandemic. The important point is that the yield curve did not provide a false signal. We can never know, however, if the 2019 yield curve inversion would have predicted a recession without COVID-19.

WHAT PART OF THE YIELD CURVE SHOULD WE USE?

We can measure the yield curve in many different ways. In my dissertation, I focused on the 10-year Treasury bond yield minus the 90-day Treasury bill yield.

\(^1\)The survey is now called the CFO Survey, a partnership of Duke University’s Fuqua School of Business, the Federal Reserve Bank of Richmond, and the Federal Reserve Bank of Atlanta.

I chose the 10-year because it was highly liquid. I chose the 90-day bill because it was important for the model to anchor on a short horizon and GDP is measured quarterly.

For robustness, I also presented results using the 5-year Treasury minus the 90-day bill. I found very little difference in the results when using the 5-year compared to the 10-year as the long rate. For example, the average lead time to a recession is 11 months for the 10-year spread and 12 months for the 5-year spread. The average difference between inversion duration and recession duration is four months for both spreads over the past eight recessions.

A number of researchers have pointed to the 2-year Treasury note as a superior proxy for the short-end of the yield curve. Indeed, FRED (Federal Reserve Economic Data, Federal Reserve Bank of St. Louis) has a pre-packaged graph of the 10-year bond yield minus the 2-year Treasury note yield that is widely reported in the media. The 10-year minus 2-year is sometimes called Wall Street’s favorite yield-curve indicator (see for example, Rennison 2022.)

But what is the logic of replacing the 90-day Treasury bill yield with the 2-year Treasury note yield? One reason to modify a leading indicator is if it has stopped working. This is not the case, however. The inversion measured using the 90-day bill yield has predicted eight of the past eight recessions without a false signal. Indeed, the track record of the 10-year minus 2-year is inferior in the out-of-sample period. The 10-minus-2-year yield curve inverted in June 1998, falsely predicting a recession, and failed to invert in June 2019 before the 2020 recession.

Most importantly, the theory that supports the signal suggests that anchoring with a very short-term rate is essential. Intuitively, one can think of the slope of the yield curve revealing information about forward growth. Given that GDP is measured quarterly, it makes sense to consider the short-term yield to be over one quarter, not over two years.

That said, I am not naïve. A model is just a model, and the 10-year bond yield minus 90-day bill yield might not be the most informative piece of the yield curve to examine. Indeed, in 2022, the 10-minus-2-year has been inverted, while the 10-year-minus-90-day has not. Time will tell.

LIMITATIONS OF THE YIELD CURVE MODEL

Any model is a simplification of reality. The yield curve model is remarkably simple. The framework model uses a single variable, the slope of the yield curve, to forecast GDP. The Federal Reserve plays no direct role in the model.

Three important theoretical aspects are not well appreciated.

First, the model focuses on real consumption growth, not GDP growth. Personal consumption expenditures, however, represent 68% of GDP, and as such, real consumption expenditure growth is highly correlated with real GDP growth. That said, a divergence is possible. For example, in the first two quarters of 2022, real GDP growth was negative, but in both quarters real personal consumption expenditure growth was positive.

Second, the yield curve measure should be real, not nominal. The theory links the real term-structure slope to real consumption growth. In my dissertation, I needed a model of expected inflation in order to calculate real yields (inflation-protected securities were not available). If inflation rates follow a random walk (or similar process), then when calculating the 10-year minus 90-day nominal yields, the inflation rate cancels out. This simplification can fail if the term structure of inflation expectations...
is not flat. A good example is the downward-sloping term structure of expected inflation in mid-2022. While the nominal yield spread is very flat, the real yield spread is quite high given that short-term real yields are highly negative.

Third, the model provides a prediction about growth. This means that any flattening (but not inverted) yield curve suggests slower growth. Inverted yield curves are especially bad for growth and are associated with recessions.

THE RESEARCH FOOTPRINT

Hundreds of papers have been published (and too many to cite) that examine the predictive ability of the yield curve for economic growth. An important early example was the work of researchers at the Federal Reserve, Estrella and Hardouvelis (1991). Indeed, many future innovations in this research stream originated at the Federal Reserve (for a recent example, see Engstrom and Sharpe 2022).

One important extension was the estimation of a logit-style model for forecasting recessions. The Harvey (1986, 1988) specification forecasts growth in both recessions and non-recessions. The Federal Reserve specification allows us to gauge the probability of a recession for a given yield-curve slope.

CONCLUSIONS

Inverted yield curves have preceded eight of the last eight recessions. Given the complexity of the economy, it is remarkable that this single variable has failed to provide a false signal. Yet it would be naive to think that this model will continue to have a 100% hit rate. If I were tasked with the same job I had in 1982 (to predict GDP growth), I would consider many additional indicators, including data from the CFO Survey. The slope of the yield curve would still play an important role, but it would not be the sole indicator.

Much of our work in finance focuses on expectations. Our key models show a relation between risk and expected returns. Expectations, however, are difficult to measure. As a result, we seek the drivers of and proxies for these expectations. Thirty years ago, the slope of the yield curve appeared to contain important information about expected economic growth. In the 30-plus years since I published my research findings in The Journal of Fixed Income, my initial idea has not (yet) been rejected.

REFERENCES


