An Investigation into the Probability That This is the Last Year of a U.S. Economic Expansion

by

Manfred Keil, Edward Leamer, Yao Li[1]

8 August 2022

Key Words: Business Cycles, Interest Rate Spreads, Recession Forecasting, Probit Model

JEL classification: E32, C22, R11

[1] Keil is Director, Lowe Institute of Political Economy, Robert Day School of Economics and Finance, Claremont McKenna College, 500 E Ninth Street, Claremont, CA 91711, mkeil@cmc.edu; Leamer is Professor Emeritus, UCLA, Anderson School; Li is Ph. D. Student, Kellogg School, Northwestern University. The authors thank Arturo Estrella, Campbell Harvey, Greg Mankiw, Jerry Nickelsburg, Mark Schniepp, Cameron Shelton, Donald Robertson, Andrew Newell, Lynne Pepall, Dan Richards, James Symons, Pierangelo de Pace, and an anonymous referee for useful comments. All errors are ours.

An Investigation into the Probability That This is the Last Year of a U.S. Economic Expansion

MANFRED KEIL, EDWARD LEAMER, YAO LI

ABSTRACT: The paper finds determinants for the end of U.S. economic expansions since 1960 using a limited dependent variable approach. Different from previous studies, we specify a left-hand side variable that is one during the year *prior* to NBER dated recessions rather than *during* recessions. We thereby avoid confounding the occurrence of a recession with its length. We limit the sample by excluding recession periods and the recovery months during the subsequent expansions. This eliminates contamination of the conclusions by taking out observations for which subsequent recessions are unlikely.

1. *Introduction*

Knowledge that a recession is coming soon is valuable for a variety of economic actors since these episodes can be very costly if not anticipated.[2] In many ways, the economy resembles a patient who shows signs in advance of an oncoming illness. We refer to the "in advance" phase as the *alarm period*. If we can find these signals, then, inside and outside lags in economic policy notwithstanding, medicine could be administered before the patient falls ill. The task for economic forecasters is then to find credible and reliable early symptoms for the onset of a downturn. At the same time, the limited ability of academic or professional economists to foresee recessions is well documented (see, e.g., Rudebusch and Williams, 2002; Stock and Watson, 2003a; Hamilton, 2011).

[2] We will use the terms "recession" and "contraction" interchangeably throughout the paper.

One reason why forecasters do not often make forecasts of large declines in employment and output is that they rely mostly on linear vector auto-regressions (VAR) models. This kind of specification allows the force of "regression to the mean" to dominate, which results in the forecasts of GDP growth that do not diverge much from its historical average of 3% in the U.S. (2.3% since mid-2009). Another reason is that these forecasters are reporting their mean forecast values. With real GDP growth at 3% normally and -1% in a recession, it takes a recession probability greater than 75% for the forecast average growth rate to be negative. This means that forecasters' concerns about recessions are revealed by small, not negative, forecast rates of growth of employment and output. Both these problems can be solved with a two-state model: normal growth and recession, with transition models that determine which state of nature is more likely to occur. Then instead of communicating just the mean growth of 1% it can be accompanied with a recession forecast probability of 50% (say).

The purpose of this paper is to offer a transition model built on the distinctive features of the last year of historical expansions. We will use a limited dependent variable (LIMDEP) approach to forecast recessions, or, to be more accurate, to identify the last year of expansions. We want to stress early that this is fundamentally different from the current popular LIMDEP forecast approach as we will show below in section III. The latter was the original approach by Estrella and Hardouvelis (1991) and, while refined through numerous publications, the methodology has never been questioned over the last 30 years.[3] In short, the approach we propose separates forecasting the start of the recession period from forecasting its length. The two methods

[3] The only publication that we are aware of which uses the "year before the recession" classification in a LIMDEP setup is a blog post by Domash and Summers (2022). The authors do not eliminate the recovery phase from their estimation period.

are only identical if a recession lasts exactly 12 months[4] and not a single post World War II recession has done so.

Our focus on the last year of expansions is a consequence of the question that should be asked and answered by forecasters: What feature of the data makes you think a recession will come within the next year? The answer is "something unusual about the last years of expansions." The word 'unusual' necessarily refers to a comparison with other periods. Our contribution to the literature is based on the choice of alarm periods and comparison periods. The alarm period refers to periods in which the recession alarm is designed to be rung. The recession prediction comes from contrasting the statistical properties of the alarm period with the statistical properties of the comparison period. When the alarm and comparison statistics are similar, the recession probability is small. When they are different in a statistically reliable way, the recession probability is high and an alarm is rung.

As a result of our analysis, we also suggest a new classification of business cycles from the old "expansion – contraction" classification, by adding two periods, the "recovery" and the "alarm period" to distinguish phases of the expansion. We choose the year (12-months) before contractions as the alarm period and focus the comparison between the expansions and the alarm period.[5] The traditional analysis selects an alarm period equal to recessions shifted backward 12 months through lags in the explanatory variables, and compares these with observations that include everything else, even recessions and initial recoveries. We remove the recovery phase

[4] The 1969 recession comes closest, having lasted 11 months.

[5] We use the 12-month measure following the convention of the literature, but this can be changed to any length if necessary for decision making.

from our sample period since it is highly unlikely to contain another contraction/warning period, and we do not want to bias our work to appear more accurate than warranted.

We proceed as follows: the next section gives a brief literature review of relevant previous papers that inspired and guided us. This is followed by a discussion of the difference in the sample selection process between determining the end of an expansion versus the start of a recession. Next, we use a “General to Specific” specification search to find more determinants for the end of an expansion than is usually done. The final section concludes.

II. Literature Review

There is a long “modern” literature in economics analyzing business cycles and forecasting recessions dating back to the 1930s, starting with the development of Coincident Economic Indicators (CEIs), Leading Economic Indicators (LEIs), and Lagging Economic Indicators at MIT/NBER (Mitchell and Burns, 1938).[6] Some of the more rigorous mathematical business cycle models have their foundations in Kondratieff (1925), Samuelson (1939), and Schumpeter (1939). Later Tinbergen and especially Lawrence Klein played a significant role in building econometric models of the U.S. economy. The Klein-Goldberger model (Klein and Goldberger, 1955) became the foundation of large econometric models used by the Federal Reserve and others.[7]

[6] The work by Mitchell and Burns actually followed earlier research by the Harvard University Committee on Economic Research which published widely used economic indicators as early as 1919. Under the direction of Warren Persons, the group developed the Harvard Index Chart; it failed to forecast the Great Depression.

[7] The FRB-MIT-Penn Econometric Model was one of the prominent examples in the early days (see Ando, *et al*. (1972).

Sims (1980) was one of the earliest critics of these large scale econometrics models which all had a strong Keynesian flavor. He suggested an alternative modeling strategy, Vector Autoregressions (VARs). Since these types of models can easily have numbers of parameters that overwhelm most time series data sets, some way is needed to focus the evidence on a smaller set of parameters. A common method of producing a more parsimonious parameterization is stepwise regression, sequentially omitting the variable with the lowest *t*-values in a "General-to-Specific" modeling strategy and to apply dynamic parameter restrictions when the data suggests them.[8] Another approach is to include principle components (the orthogonal linear combinations of the primary variables with the highest variances), computed from primary variables that have been normalized all to have the same variance. It is the sample variances of the linear combinations, not their statistical significance, that determines the inclusion ordering. Why do stepwise and principal components use entirely different orderings, we ask rhetorically? Which is the wiser choice?[9]

In the time series setting, the use of principal components to forecast economic activity, has been marketed with the name Dynamic Factor Models (DFMs). These models have also been used to construct CEIs and LEIs (Stock and Watson, 1989). The DFMs typically use all the data available, thus implicitly assuming that the data generation process is the same for both expansions and recessions.

[8] This is sometimes referred to the London School of Economics (LSE) methodology, associated with David Hendry and Dennis Sargan. For an early example, see Davidson *et. al* (1978). A convenient summary can be found in Verbeek (2012, 64-5).

[9] Chamberlain and Leamer (1976) justify principal components when the prior distribution has mean zero and a covariance matrix proportional to the identity matrix.

An alternative method to forecast recessions, which we will follow in this paper, is a limited dependent variable (LIMDEP) model. This strategy relies on forecasting variables predicting whether a Left-Hand Side (LHS) variable Y_t takes on the value of 1 if the economy is in a recession (as defined by the NBER) at time *t*, and 0 if the economy is in an expansion at time *t*. Hence this approach attempts to forecast two states of the business cycle: contractions and expansions.

This results in the following probit[10] model

$$\Pr(Y=1|X_1, X_2,\ldots, X_k) = \Phi(\beta_0+ \beta_1 X_1+ \beta_1 X_1+ \beta_2 X_2+\ldots+ \beta_k X_k) \qquad (1)$$

where Φ is the cumulative standard normal distribution function. The *X*s are lagged explanatory variables. The parameters are obtained through Maximum Likelihood Estimation (MLE).

Probit estimation is a standard tool in econometric analysis when working with cross-sectional data. A foundation of statistical analysis using LIMDEP in a time series framework is given by Poirier and Ruud (1988), who show that the estimator is consistent, and the authors derive its distribution. A good reference for using the appropriate standard errors for statistical inference is Chan and Kuk (1997).

The first publication of a limited dependent variable model to forecast recessions is Estrella and Hardouvelis (1991). In essence, the authors initially chose to include the slope of the yield curve lagged four quarters to forecast growth rates of a variety of real variables, including real GDP. Specifically they use the difference between a 10-Year Government Bond and the 1-

[10] Some authors have used a logit specification instead.

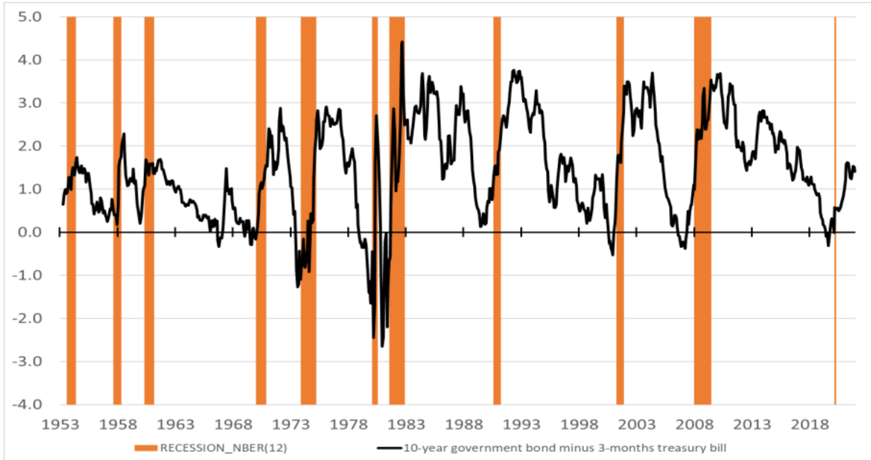
Year Treasury Bill as part of their set of leading variables.[11] The paper establishes the slope of the yield curve (the spread) as a predictor for real economic activity. Note that the spread is available at a high frequency and is not subsequently revised. Stock and Watson (1989; 383) use logit estimation to analyze the extent to which their newly constructed ILEs can forecast recessions, although this was not the focal point of their paper.

Excellent summaries of the earlier literature in forecasting recessions using this approach can be found in Stock and Watson (2003b), and Wheelock and Wohar (2009). Subsequently several Federal Reserve publications have employed the same methodology, see *inter ali*a Andolfatto and Spewak (2018), Christensen (2018), Johansson and Meldrum (2018), McCracken (2018). Bernard and Gerlach (1998) use the same model and find it to be useful in forecasting recession in eight additional countries. Camacho and Palmieri (2021), extend the analysis to primarily OECD data. Nissilä (2020) gives a comprehensive survey.

What ties most of these studies together is a binary dependent variable and the use of the interest rate spread between a long term and short term interest rate as an explanatory variable. Figure 1 plots the spread between a 10-Year Treasury Bond and a 3-Months Treasury Bill.

Figure 1: 10-Year Government Bond Minus 3-Months Treasury Bill, U.S., April 1953 - December 2019, Recessions Shaded

[11] Harvey (1988) also focuses on the relationship between the term structure and real variables (here consumption growth) but does not make use of LIMDEP estimation. The author uses stocks and bonds to forecast GDP growth.



Six of the seven recessions since 1970 have been preceded by an actual inversion and the other one by a near inversion. This can be explained by a predictive story and by a causal story. The predictive story is that an inverted yield curve amounts to a bond-market forecast that the short-term rate will soon be lowered substantially by the Fed when the recession arrives.[12] If the short-term rate were thought to be persistent, who would buy a long-term bond paying less than a short-term bond? The causal story is that banks become cautious lenders when they must pay short-term borrowers higher rates, while charging long-term borrowers lower rates. Thus, an inverted yield curve comes with a rise in lending standards which can be more important than the interest rate when marginal buyers have low credit scores.

Towards the end of our sample period, the short-term interest rate exceeded the long-term bond rate starting in March 2019. This was the result of long-term interest rates declining sharply as the outlook on the global economy deteriorated, not a rapid rise of the short-term rates. It is possible that the predictive or the causal stories are different when the inversion comes from a jacking up of short-term rates by the Fed versus when the inversion comes from a decline in the

[12] Kliesen (2022) presents an interesting analysis of how the previous episode of Federal Reserve tightening in 2015-2018 differed from the five previous episodes since 1983.

long-term interest rates dictated by global bond markets. There is nowhere near enough data to study this hypothesis (on this, see also, Cooper et *al.*, 2020). Should the Federal Reserve decide to sell its large portfolio of long-term bonds, we will get yet another scenario of long-term interest rates rising.

Equation (2) is an example of the typical probit equation used to forecast recessions:[13]

$$\widehat{\Pr(Recession = 1 \mid spread}_t = \underset{(0.19)}{\Phi(-0.40} - \underset{\substack{(0.15)\\ [0.43,1.01]}}{0.72} \times spread_{t-12}) \qquad (2)$$

t = January 1960 - December 2019, McFadden R^2= 0.263[14]

where *Recession* takes on a value of 1 for the period from the peak plus one month to the trough as dated by the NBER, and *spread* is the difference between the 10-year Government Bond and the 3-Month Treasury Bill.

The coefficient in equation (2) on the spread is only slightly smaller than the estimate of Estrella and Hardouvelis (1991) using quarterly data ending in the fourth quarter of 1988. The McFadden R^2 is also similar. Hence the relationship has remained remarkably stable over 30 years even when using higher frequency data and a slightly different lag structure (12-months lag versus

[13] Estrella and Hardouvelis (1991) used quarterly rather than monthly data. This allows them to compare the probit specification with a more standard regression that forecasts real GDP growth rates. For the probit model, *spread* is specified with four lags. Hence the forecasted probabilities of recessions are for the current quarter based on the slope of the yield curve a year earlier. Consequently, we allowed for 12 lags in the monthly specification above.

[14] "()" show HAC (Newey and West, 1987) standard errors; "[ ]" is the 95% confidence interval for the parameter of interest. Our sample ends in December 2019 since the Coronavirus recession, which started in February 2020, was impossible to forecast using economic variables. Potentially we could have extended our sample period to February 2020, but felt that there were a sufficient number of rumors regarding the virus floating around at the time that may have affected the data early in 2020.

4-quarters lag).[15] Note that this is true in the face of fundamental policy changes such as the federal funds rate basically being set to 0 for extended periods during more recent times.

For forecasting, omitted variables are not a problem if they retain their correlations with the included variables, but having more right-hand side (RHS) may create a more accurate and more resilient forecast, and may change the coefficient on the spread. In response to these concerns, Estrella and Hardouvelis (1991) add the growth rate of real GDP, the growth in the ILE, inflation, and the real federal funds rate (all lagged by a year) to the RHS. None of the new coefficients is statistically significant.[16] The coefficient on the spread variable does not change substantially and remains highly statistically significant. Estrella and Mishkin (1998b) confirm this finding. This result may explain why many authors subsequently do not use additional explanatory variables. Moreover, the relationship is remarkably stable over time (Estrella *et al*., 2003), especially when compared to models that attempt to forecast GDP growth.

III. *Contrasting the Proposed Method with the Traditional Method*

We argue that a recession forecast can be thought to come from contrasting the data during an alarm period with the data during a wisely chosen comparison period. The features of the data during the alarm period that are reliably different from the comparison period are then used to determine the "loudness" of the alarm going forward. This is a fundamentally different approach than the methodology used in the literature. We are choosing a different alarm period and a different comparison period. Our question is "Is the start of the recession no more than 12 months

---

[15] Estrella and Mishkin (1998a) use monthly data and a 12-months lag. In addition, they expand the sample to European countries.

[16] This does not mean that there are no RHS variables that can add to the explanatory power of the spread, although many authors subsequently must have felt so given the lack of experiments with additional RHS variables. Note the absence of variables potentially indicating an inventory problem in houses, automobiles, or durables.

away?" which calls for an alarm in the last year of the expansion. The traditional question is "Will we be in a recession in 12 months from now?" This creates an alarm period by shifting the recession periods 12 months backward in time.

The statistical model that we estimate contrasts the 12 months *precedin*g each recession with the previous months of the expansions excluding the (employment or output) recoveries. The logic for excluding the recoveries from our statistical model is that these periods are characterized by unusually high rates of growth of jobs and GDP, and a steep yield curve. If that were the comparison group, the rest of the expansion period could be found to be different, leading to early and inappropriate recession forecasts. For the same reason, we exclude the recessions from the comparison group because differences between the recessions and the last years of expansions cannot produce reliable recession forecasts.[17]

If the recession lasts 12 months, then the two alarm periods are identical. If the recession is less than 12 months, like the 1990/91 recession, then the Equation (2) alarm period excludes months adjacent to the recession. This means that the alarm is turned off completely in the months immediately preceding the recession, and even more worrisome, these months become part of the comparison period. If the contraction period is more than 12 months, as was the case for three post World War II contractions including the Great Recession, then the alarm period starts at the correct time but it also extends into the recession. In this case, the statistical model will look for differences between these early recession months and the comparison period. This is designing a

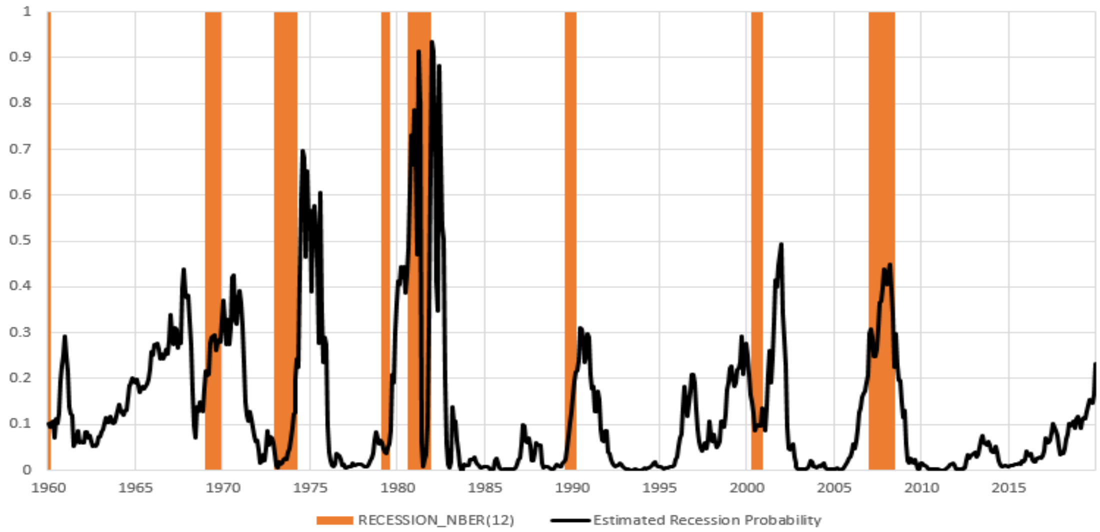
[17] One problem with our categorizations of recoveries and alarm periods is that one recession ended in July 1980 and another began in August 1981 ("double dip recession"), which leaves only 12 months between them, all of which are in the 12 months preceding the recession that began in August 1981. But these 12 months also include the recovery from the 1980 recession which took five months. We have resolved this conflict by treating the five months from August 1980 to December 1980 as part of the recovery from the recession that ended in July 1980, and therefore the 12 months before August 1981 when the next recession began are reduced to 7 months.

recession alarm that will continue to ring when the recession is already here, but this is not a substantial problem. We want to stress here that by including recession months in the alarm period, this confuses the statistical analysis by encouraging it to look for unusual behavior early in recessions.

One difficulty with equation (2), if estimated with all the data available during the sample period, is that it requires the model to predict *during* a recession whether or not the recession will still be occurring in twelve months. We find this highly problematic. If the yield curve is particularly steep in the midst of a recession, the model might find that a steep yield curve predicts that the recession will end soon and if you were not careful, you would think it had discovered that an inverted yield curve predicts a recession will come soon. In other words, the model is being used to predict recessions and *also* to predict the ends of recessions. A way to eliminate this error is to *omit* the recession periods. This changes the question from "How are the ends of expansion different from all other periods?" to "How are the ends of expansions different from the earlier periods in the expansions?"

We conclude this section with a close look at the forecasted probabilities for a recession when using equation (2). Figure 2 shows the estimated probabilities with alarm periods shaded. If this were a perfect predictor, it would be one in the alarm periods and zero everywhere else. When the probability increases substantially after the alarm periods, that is a forecast error. Many of the highest probabilities in Figure 2 occur after the alarm periods when "the house has already burned down."

Figure 2: Estimated Recession Probabilities, Probit Forecasting Equation (2), January 1960 to December 2019, Alarm Periods Shaded

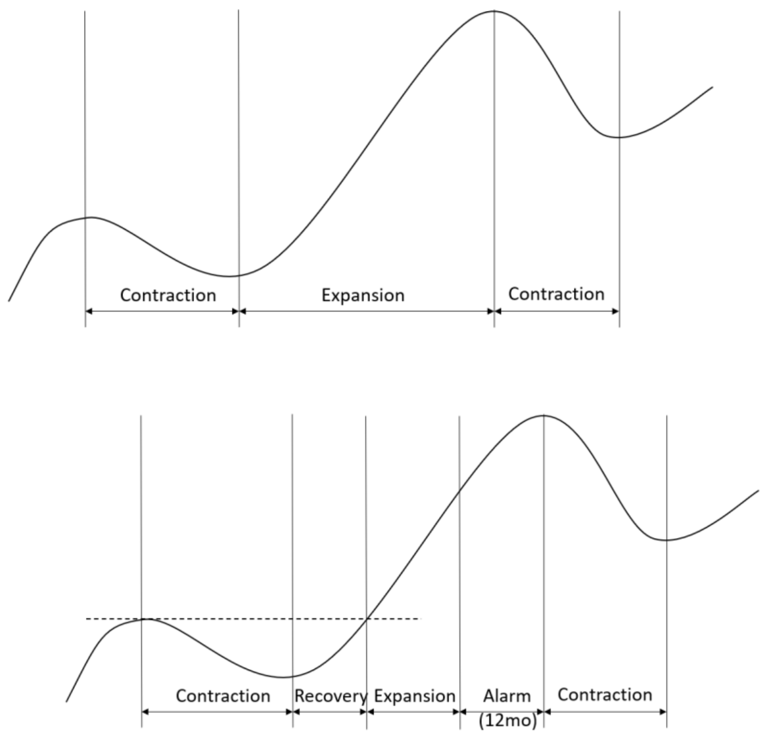


III.1. *Four States of the Economy*

The most commonly used forecasting approach divides movements in economic activity into two states: (i) an expansion state, and (ii) a contraction state, as illustrated in the upper graph of Figure 3. It is our view that we should distinguish four states instead, each of which exhibits distinct behaviors: (i) a healthy expansion state, (ii) a troubled alarm year before recessions, (iii) an unhealthy state known as contraction, and (iv) a recovery state with exceptional growth taking the economy back toward pre-contraction trend, as illustrated in the lower graph of Figure 3. Our intent is to assure that our model is focused only on the transition from healthy to contraction, not from contraction to recovery, and not from contraction to healthy economy. We want to know how the last year of an expansion is different from earlier periods in the expansion but not how it is different from the recovery. [18]

[18] Parenthetically, during the 1930s, the focus was on the transition from recession to healthy, not the focus on transition from healthy to recession.

Figure 3: Illustration of the Two-State Economic Cycle (upper) and the Four-State Economic Cycle (lower)



There is no official determination of the recovery period and we thus adopt Leamer's (2019) suggestion that the recovery ends when payrolls first exceed their previous peak. Another definition is the return to the level of GDP period prior to the contraction, or a return to a level of GDP implied by trend growth, but this raises complications especially after 2008/09 when the trend rate of growth of GDP shifted down from close to 3% to 2.3%.

III.2. *Probit Estimates*

Table 1 reports the changes to the estimated model when the specification is altered in ways discussed so far. In both cases the data set begins in January 1960. Model (a) implements Equation (2) and uses the binary recession indicator as the dependent variable, and the spread lagged 12 months as the explanatory variable with no data exclusions. This is the form of the model in all references except for Stock and Watson (1989) and Leamer (2009, 2019). Model (b) changes the dependent variable to the binary variable equal to one in the twelve months before the recession began with the recessions and recoveries omitted.

$$Year\_Before\_Recession = \begin{cases} 1\ if \text{ during 12 month prior to recession} \\ 0\ otherwise \end{cases}$$

Table 1: Probit Specification to Forecast Recessions, U.S., Monthly Data, January 1960 - December 2019 (Model (a)) or 2018 (Model (b))

Model	(a)	(b)
Dependent Variable	*Recession*	*Year before Recession*
Explanatory Variable	$spread_{t-12}$	$spread_t$
Exclusion	-	Recession, Recovery
Number of observations	718	455
Mean Dependent Variable	13%	18%
Slope Coefficient	-0.72	-1.45
|*t*-statistic|	4.80	5.80
McFadden R^2	0.263	0.404

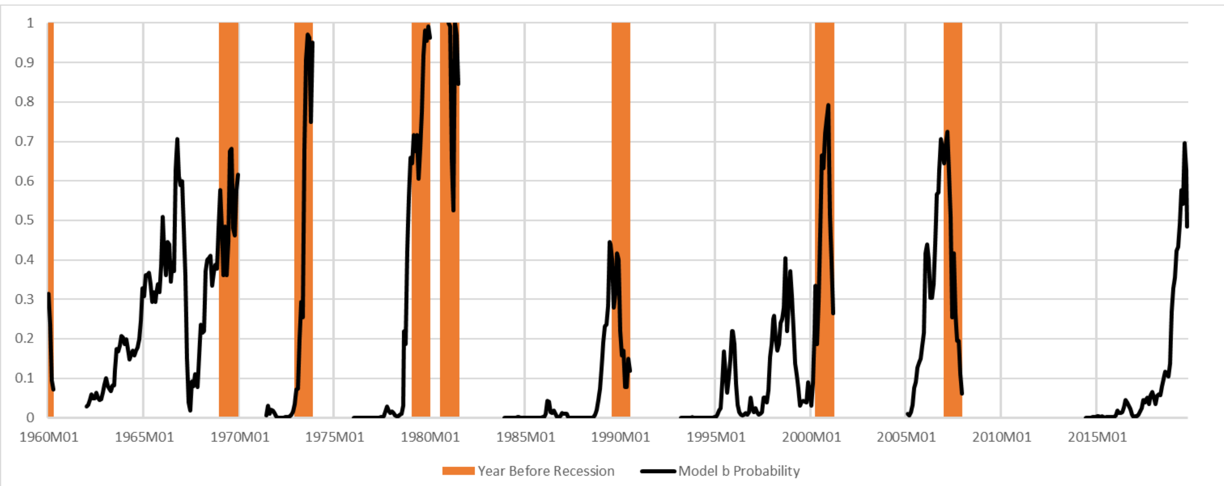
Note: Equations were estimated with Probit or GLM (Newton-Raphson/Marquardt steps). *t*-statistics are calculated using HAC standard errors. Using homoskedasticity-only standard errors results in *t*-statistics roughly twice the size)..

Note that Model (b) uses the current value of *spread* in the forecasting equation rather than its lag, since we are trying to determine if the data suggests that the current observation is in the last year of an expansion. Using this dependent variable we lose ten observations since we do not want to treat the data from March 2019 to December 2019 as part of a year before recession. Model (b) uses a different dependent variable and excludes the recession periods and the recovery periods, and this has the effect of increasing the spread coefficient, its *t*-statistic, and the

McFadden R^2. Model (b) is superior: it answers the right question and results in a better fit even with smaller numbers of observations.

Figure 4 depicts, in orange, the years before recessions and, in black, the estimated probability that a month was in the year before recession per Model (b).

Figure 4: Estimated Recession Probabilities, Probit Forecasting Model (b), January 1960 to December 2019, Alarm Periods Shaded



It is pleasing to see the improved forecast performance for the 1990 recession, which, according to Chauvet and Potter (2005) was particularly difficult to forecast with a standard static yield curve probit model. The image depicts a loud recession alarm in the middle of the 1960s which we think was offset by the expansion of federal spending for the Vietnam War. In other words, that is a valid alarm even though it did not immediately precede a recession.

IV. *Data Analysis With an Expanded Model*

Following the forecasting exercise and theoretical considerations, we believe that the Stock and Watson, and Leamer approach of forecasting the period prior to the NBER dated recession is superior in forecasting recessions, and, in addition, we will use the Leamer method of restricting the sample period in our analysis below when we add additional variables to the Model (b) specification.

In light of the previous discussion, we will use Model (b) as our base model, but will be looking for other variables that might improve its forecast ability. Given that Estrella and Hardouvelis (1991) excluded other explanatory variables on grounds that they did not contribute to forecasting recessions once the spread was included on the RHS, we need to reopen this investigation now that we have separated recession occurrence from recession length, and chosen a different sample for our estimation.

We proceed as follows in our specification search from here. Starting with Model (b), we will allow for further lags in the spread, plus include additional RHS variables and their lags that are suggested by theory or otherwise. Looking for reasonable restrictions in the data to promote parsimony, we arrive at the following specification (Model (c)):[19]

$$\overline{\Pr(1\_Year\_Before=1\,|\,spread}_t = \Phi(\underset{(0.65)}{-2.71}\underset{(1.03)}{-3.65}\times spread_t \underset{(1.60)}{-6.06}\times spread_{t-3}) + \underset{(2.99)}{5.90}\left(\frac{1}{3}\sum_{i=0}^{2} ur_{t-i} - \frac{1}{12}\sum_{i=0}^{12} ur_{t-i}\right)_t$$

$$\underset{(4.86)}{-10.18}\Delta_3(hrsManu\times shManu)_{t-9} \underset{(0.0012)}{-0.0034}\Delta_{12}housstart_t + \underset{(0.0009)}{0.0041}(housstock_t - 10,000)\times DHS10$$

$$\underset{(0.063)}{-0.184}\Delta_4(CSI)_t$$

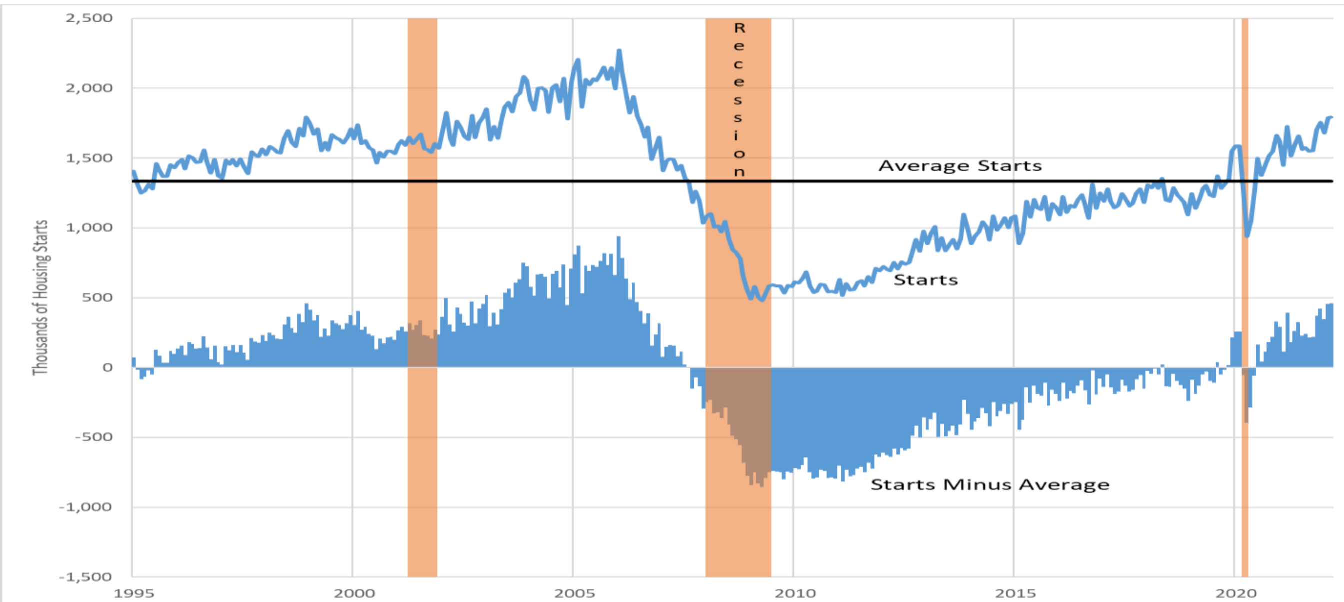
McFadden R^2= 0.886, t = 1960:M1 – 2018:M12 excluding recessions and recovery

[19] We experimented with the sample period by considering output (GDP) rather than employment returning to pre-recession levels. The coefficients remained fairly robust and are available from the authors by request.

The interpretation of the included variables and their effects on forecasting the year prior to a recession is as follows: the current value of spread and its value three months ago appear to be strong predictors. Agents will increase their estimate of an imminent recession if the increase in the spread a quarter ago (perhaps during the previous Fed meeting) is followed by keeping the spread negative three months later. Change in unemployment rates have been suggested in various publications as an alternative predictor for an oncoming recession (see, e.g., Kliessen, 2018). We found the analysis by Sahm (2019) most convincing (see also Nunn *et al.*, 2019). The author calculates the difference between the quarterly average in the unemployment rate and compares it to the average of the unemployment rate over the previous year. We experimented with other measures of the unemployment rate, such as a simple one-month change but found the Sahm measure to dominate.

Leamer (2007) stresses the central role that housing-starts play in forecasting recessions and makes the point that the Fed can influence the timing but not the total. Figure 5 illustrates housing starts from the first quarter in 1995 to the third quarter in 2022, beginning in 1995 when housing starts returned to their approximate historical normal rate of 1.5 million per year. A horizontal line depicts the mean rate of building for that whole period which is the constant rate of housing starts that would produce the same total number of houses as the actual data during this period. The difference between the actual and the mean is an indicator of overbuilding and underbuilding, which is also depicted in Figure 5. This indicates that the 2001 recession did not involve a housing correction and the overbuilding of houses stretched all the way from 1995 to near the start of the 2008 recession. It took from 2008 to 2020 for housing starts to offset the substantial amount of overbuilding that had occurred before 2008. Overbuilding has returned after 2020 but the total amount of overbuilding is not close to a dangerous level. Given that another year or two before that problem might emerge, rising interest rates may put an end to that overbuilding. That's good medicine since it prevents the overbuilding that makes housing markets fragile and ready for recession.

Figure 5: Housing Starts, Mean, U.S., Monthly, January 1995 - March 2022, SAAR



Partially as a result of the bursting housing bubble having played a central role for the economic collapse from 2008-2009, lending standards were initially significantly tighter than during the previous expansion. In addition, it took several years to work through foreclosures and forced sales. At any rate, the low level of housing starts is one of the defining characteristics of the "Not So Great Expansion" since 2009. Comparison across all post World War II recovery periods shows that the post Great Recession employment return to pre-recession levels took the longest. This does not deny the fact that unemployment rate levels eventually ended up at values below 4% - while GDP growth remained weak by historical standards (Fernald *et al.*, 2017)). However, note that housing, in itself, only makes a small contribution to GDP growth. It therefore cannot explain why the economy only grew at 2.3%, on average, during this phase rather than the 3% that seemed to have been a regularity since the 1960s (indeed the average real GDP growth rate since 1896 has been roughly 3%). Fernald *et al.* (2017) blame slower TFP growth and lower labor-force participation on the weaker GDP growth rates, although they admit that state and local governments purchase may have been weaker than in previous recoveries as a result of the housing market collapse and the financial crisis.

Leamer (2011) explains that the extreme length and depth of the last three recessions was due to the permanent loss of manufacturing jobs in contrast to the repeated V-shape for manufacturing jobs in all previous recessions. During earlier recessions, manufacturing played a more central role. However, manufacturing employment peaked in 1979 at roughly 19.5 million and by February 2020 (before the current recession) had declined to 12.8 million. While average hours of work in manufacturing may have been a good indicator of recessions in the early post-World War II period up until the Volcker recession, it seems to have lost its cyclical sensitivity as service sectors in the U.S. economy became more important. To capture this effect, we interact average hours in manufacturing with the employment share of manufacturing.

The Conference Board's Consumer Confidence Index is part of the LEI. Luzzetti et *al.* (2022) show, as reported in The Economist (2022), that the University of Michigan's Consumer Sentiment Index is superior to use for forecasting economic activity when compared to the Consumer Confidence Index. The Michigan survey asks respondents' plans for the coming year, like buying a car. The survey used by the Conference Board asks participants to forecast the usual macroeconomic variables and it is doubtful they would see in the future what is not already covered by the current RHS variables of our forecast equation. Regardless, we wanted to test if the Michigan measure contributes to the forecast of the end of expansions, and against our expectations, it did. Nissilä (2020) finds a similar result when using data for Finland. In addition, and somewhat surprisingly, the measure seems somewhat orthogonal to the other variables because it did not have any substantial impact on the previously included variables.

The final variable in our specification is the housing stock generated during the respective expansions. Here we simply accumulated housing starts from the end of each recession. Leamer (2009) stresses the role that consumer durables, housing stock, and automobiles play towards the end of each expansion, in that they accumulate to levels that make markets fragile. We allowed the effect from cumulative housing to change from negative to positive at a level of 10 million units (*DHS10* is 1 at that point and 0 otherwise). As it turns out with this variable, it is only the positive part of the slope that is statistically significant, although we were able to get a negative slope for values below 10 million units that was substantially smaller in absolute terms than the positive counterpart.

IV.1. Forecasting the Great Recession

The acid test for specification search in a forecasting setting, in our view, is the ability to forecast outside of the sample period (external validity). Here we will estimate all three models (a, b, and c) and compare their forecasting performance. For this exercise we assume macro-economic data are available only through December 2007 and no information about the 2008 period has yet arrived. Hence the methods for forecasting a recession can be estimated using data only through December 2006. Recessions forecasts for 2008 can then be obtained by using the macro data in 2007. Model (a) is used to assign to each month of 2007 the probability that 12 months later the U.S. would be in recession based on the slope of the yield curve in that month.

Model (b) and Model (c) are used to assign to each month of 2007 the probability that it is part of the last year of the expansion. Table 2 reports these results.[20]

Models (b) and (c) establish much higher recession probabilities than model (a). The highest probabilities for (a), (b) and (c) are 0.42, 0.69 and 0.99. While (b) and (c) refer to the same event, this is not the case for (a). However, the first probabilities are directly comparable because they indicate the probability that the recession begins in January 2008, either because (a) the recession occurs 12 months after January 2007, or because ((b) or (c)) January 2007 is in the last year of the expansion.

The Great Recession was notoriously hard to forecast, especially since the initial phase of the downturn was relatively mild. For the record, in December 2007, the month before the start of the Great Recession,[21] forecasts of real GDP growth for 2008 by the Blue-Chip professional forecasters ranged from 2.1% to 2.4%. The Lehman Brothers collapse of mid-September 2008 turned the initially mild decline into the then most severe post World War II recession. As late as May 2008, Federal Reserve Chair Bernanke said, in a speech at the Federal Reserve Boston, that he was more concerned with inflation than with unemployment. Still, there were clear warning signals from as early as March 2008, when Bear Stearns had to be bailed out that the downturn resulting from the financial crisis could become severe.

[20] Alternatively, we could have advanced the sample period one month at a time for the estimation period and forecasted the next month (pseudo out of sample forecasts or *poos*). This type of forecast is desirable if different models contain lag dependent variables, which is not the case here. At any rate, the coefficients do not change by a sufficiently large amount to change the impression we get from the results in Table 2.

[21] The NBER (2019) dated the peak of the previous expansion for December 2007 and states that the recession begins in the same month as the peak occurs. For coding purposes, we assume that the recession started in January 2008 and ended in June 2009, resulting in an 18 months downturn. The subsequent 128 months expansion then began in July 2009. We have made similar adjustments for previous business cycles.

So, what did the forecasters say just prior to the Lehman Brothers event? The Federal Reserve Bank of Philadelphia conducts a monthly survey of professional forecasters.[22] The median response from 47 professional forecasters during the third quarter of 2008 about real GDP growth during the fourth quarter of 2008 was -0.7%, clearly missing the large decline of -8.5%.[23] Our preferred specification suggests an end of the expansion within a year of above 95% from February to May 2007. Early warning signals indeed.

During the year prior to the recession, Model (a) has elevated probabilities of around 35% between July 2007 and July 2008, always being around 35%. This is not the case for Model (c). The probability of the expansion ending within a year is less than 2% from January to October 2006. It rises to 16% in January 2007 and is above 90% from February to June 2007 before falling to 88% in July (forecasting an early 2008 recession), thereby ringing a serious alarm bell for policy makers to wake up and adjust economic policies. The reason for the low probability in January of 2007 is that the spread was only mildly negative (-0.08) from August and September 2006, while doubling its value in October (-0.19) and November of 2006 (-0.34); Model (c) picks up this increase immediately, similar to Models (a) and (b), but reacts more so three months later in February 2007. The probabilities hardly decline after that, and always remain above those forecasted by Models (a) and (b).

The conclusion from the forecasting exercise must be that Model (c) substantially outperforms the traditional NBER recession date forecasting model Model (a). We strongly feel

[22] This is the oldest survey of forecasts in the U.S. It was started in 1968 by the NBER and the American Statistical Association as a quarterly survey. It has been with the Federal Reserve Bank of Philadelphia since 1990 (see, e.g., Federal Reserve Bank of Philadelphia, 2019). The Livingston Survey, which is the oldest survey of economic forecasts and was started in 1946, shows that forecasters missed both the 1981/2 and 1990/1 recessions.

[23] Forecasts of the unemployment rate were similarly optimistic: the median forecast in November was for the unemployment rate to peak at slightly above 7% a year later.

that the "real time" forecasting exercise confirmed empirically the superiority of the model established early on theoretical grounds.

Table 2: Forecasted Probabilities for Recession, Great Recession, Models (a), (b), and (c), Estimation Period January 1960 - December 2006, Forecast Period: January to July 2008.

Forecasted Month (i)	2008 Jan	Feb	March	April	May	June	July
Pr (Y=1) Model (a)	0.38	0.40	0.42	0.37	0.31	0.20	0.27
Pr (Y=1) Model (b)	0.61	0.65	0.69	0.58	0.47	0.23	0.38
Pr (Y=1) Model (c)	0.16	0.99	0.99	0.93	0.99	0.97	0.88

IV. Conclusion

In this paper we select the final year of economic expansions as the alarm period and build statistical methods for a recession alarm equal to the probability that a month is in the alarm period. Our work differs from previous papers on the subject that use a limited dependent variable analysis to forecast recessions in several significant ways. Unlike many studies that associate a binary recession indicator with lagged explanatory variables, we choose a binary last-year-of-expansion indicator with concurrent explanatory variables, omitting the recession and recovery data. This is done to ensure that we are asking the right question.

If all recessions lasted exactly one year, and if the method adopted in the standard literature uses explanatory variables with 12-month lags, then the two approaches are close, but not quite the same because our approach excludes both contractions and recoveries. Since in real life contractions are of variable lengths, then two approaches perform quite differently.

After adjusting the dependent variable and sample period, we present a forecasting equation that takes into account a variety of variables other than the interest rate spread to forecast an economic slowdown. Forecasting the Great Recession by limiting our end of sample period to December 2006, our specification clearly outperforms both the traditional approach of forecasting NBER recession dates and the modified model that forecasts the end of the expansion year using the spread only.

There are some limitations to our analysis which we hope to address in future research. To conduct a forecasting exercise in "real time" we used the Great Recession, because it was notoriously difficult to predict, especially if you focused on the period December 2007 to August 2008, that is prior to the Lehman Brothers event. Subsequent work could do a similar exercise with other recessions that were difficult to forecast such as the dot-com downturn in 2001.

Focusing on U.S. business cycles is somewhat limiting, although we view it as a convenient starting point given the methodology chosen by the NBER dating committee in setting peak and trough dates. In the future, we intend to expand the analysis to include other OECD countries and to compare our analysis with the literature in the field. Finally, we have designated a 12-months period as our alarm period. We did so since this is in line with the lag structure of the existing literature. However, there is room to experiment with other alarm periods, such as 6-months or even 24-months. However, having a half-year planning horizon does not allow for much time to turn the U.S. economy, especially given the inside and outside lags in policy making – such a short warning period would give doctors barely time for prescribed medicine to have an effect on the outbreak of a disease. Looking for warning signs two years ahead may simply show us that there are few if any symptoms.

Bibliography

Ando, A., Modigliani, F., and R. Rasche (1972). "Equations and Definitions of Variables for the FRB-MIT-Penn Econometric Model, November 1969." In *Econometric Models of Cyclical Behavior*, vol I edited by B. Hickman NBER Studies in Income and Wealth, No. 36, New York: Columbia University Press, 543-600.

Andolfatto, D. and A. Spewak (2018). "Does the Yield Curve Really Forecast Recessions?" *Economic Synopsis*, Federal Reserve Bank of St. Louis, November.

Bernard, H. and S. Gerlach (1998). "Does the Term Structure Predict Recessions? The International Evidence." *International Journal of Finance and Economics*, 195-215.

Camacho, M. and G. Palmieri (2021). "Evaluating the OECD's Main Economic Indicators at Anticipating Recessions." *Journal of Forecasting*, 80-93.

Chan, J. and A. Kuk (1997). "Maximum Likelihood Estimation for Probit-Linear Mixed Models with Correlated Random Effects." *Biometrics 53*, 86-97.

Chamberlain, G. and E. Leamer (1976), "Matrix Weighted Averages and Posterior Bounds," *J. Royal Statistical Society*, B 38, 73-84.

Chauvet, M. and S. Potter (2005). "Forecasting Recessions Using the Yield Curve." *Journal of Forecasting*, 77-103.

Christensen, J. (2018). "The Slope of the Yield Curve and the Near-Term Outlook." Federal Reserve Bank of San Francisco, *Economic Letters*.

Cooper, D., Fuhrer, J. and G. Olivei (2020). "Predicting Recessions Using the Yield Curve: The Role of the STanceof Monetary Policy." *Current Policy Perspectives,* Federal Reserve Bank of Boston, February.

The Economist (2022). "Surveying The Coming Wreckage." April 29, Finance & Economics.

Davidson, J., Hendry, D., Srba, F., and S. Yeo (1978). "Econometric Modeling of the United Kingdom." *Economic Journal*, 661-92

Domash A. and L. Summers (2022). "History Suggests a High Chance of Recession over the Next 24 Months." *Blog Post*, Kennedy School, Harvard University.

Estrella, A. and G. Hardoulin (1991). "The Term Structure as a Predictor of Real Economic Activity." *Journal of Finance*, 555-76.

Estrella, A. and F. Mishkin (1998a). "The Predictive Power of the Term Structure of Interest Rates in Europe and the United States: Implications for the European Central Bank." *European Economic Review*, 1375-201.

Estrella, A. and F. Mishkin (1998b). “Predicting U.S. Recessions: Financial Variables as Leading Economic Indicators.” The *Review of Economics and Statistics,* 45-61.

Estrella, A., Rodrigues, A. and S. Schich (2003). “How Stable is the Predictive Power of the Yield Curve? Evidence from Germany and the United States.” *Review of Economics and Statistics*, 629-44.

Federal Reserve Bank of Philadelphia (2019). “Annualized Percentage Change in Mean Responses – Real GDP Growth.” *Survey of Professional Forecasters*.

Fernald, F., Hall, R., Stock, J., and M. Watson (2017). “The Disappointing Recovery in U.S. Output after 2009.” *Brookings Papers on Economic Activity*, Spring, 1-58.

Hamilton, J. (2011). “Calling Recessions in Real Time.” International Journal of Forecasting, 1006-26

Harvey, C. (1989). “Forecasts of economic growth from the Bond and Stock Markets”, *Financial Analysts Journal*, vol. 45, n. 5, pp. 38-45.

Harvey, C. (1988). “The Real Term Structure and Consumption Growth." *Journal of Financial Economics* 22, 305-33.

Johansson, P. and A. 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, 2018.

Klein, L. and A. Goldberger (1955). *An Econometric Model of the United States*, 1929-1952. Amsterdam: North Holland.

Kliesen, K. (2022). “A Comparison of Fed ‘Tightening’ Episodes since the 1980s.” *Working Paper* 2020-003, Federal Reserve Bank of St Louis, February.

Kliesen, K. (2018). “Recession Signals: The Yield Curve versus Unemployment Rate Troughs.” *Economic Synopsis*, Federal Reserve Bank of St. Louis, June,

Kondratieff, N. D. (1925). “Die Langen Wellen der Konjunktur.” *Archiv fur Sozialwissenschaft und Sozialpolitik*, 573-609.

Luzzetti, M., Ryan, B., Weidner, J. and A. Yang (2022). “Consumer Sentiment Gap is Proving Prescient for Curve Shape.” *US Economic Notes,* Deutsche Bank Research.

Mitchell, W. and A. Burns (1938). “Statistical Indicators of Cyclical Revivals,” chapter in Mitchell and Burns (eds.) *Statistical Indicators of Cyclical Revivals*, New York: National Bureau of Economic Research.

Leamer, E. (2019). “Recession Forecasts Based on Monthly Data.” *Mimeo*, UCLA Anderson School, Anderson Forecast.

Leamer, E. (2011). “Why Are Some Recoveries Short and Others Long?” July 2021 - NBER Working paper 28982

Leamer, E. (2009). *Macroeconomic Patterns and Stories*. Berlin: Springer.

Leamer, E. (2007). “Housing *is* the Business Cycle.” *Proceedings* - Economic Policy Symposium - Jackson Hole, Federal Reserve Bank of Kansas City, pp. 149-233.

McCracken, M. (2018). “What Are Indicators Saying about a Potential Recession?” *On The Economy: Economic Blog*, Federal Reserve Bank of St. Louis.

Nissilä, W. (2020). “Probit Based Time Series Models in Recession Forecasting - A Survey with an Empirical Illustration for Finland.” *Bank of Finland Economics Review*, No 7/2020, Bank of Finland, Helsinki.

National Bureau of Economic Research (2019). “U.S. Business Cycle Expansions and Contractions.” http://www.nber.org/cycles/cyclesmain.html

Newey, W. and K. West (1987). “A Simple Positive Semi-Definite, Heteroskedastic and Autocorrelation Consistent Covariance Matrix.” *Econometrica*, 703-8.

Nunn, R., Parsons, J and J. Shambaugh (2019). “How Will We Know When a Recession is Coming?” *Report*, Hamilton Project, June 6, Washington: Brookings Institute.

Porier, D. and P. Ruud (1988). “Probit with Dependent Observations.” *Review of Economic Studies*, 593-614.

Rudebusch, G. and J. Williams (2002). “Forecasting Recessions: The Puzzle of the Enduring Power of the Yield Curve.” *Journal of Business and Economic Statistics*, 492-503.

Sahm, C. (2019). “Direct Stimulus Payments to Individuals.” *Report*, Produced by the Hamilton Project, Washington: Brookings Institution.

Samuelson, P. (1939). "Interactions Between the Multiplier Analysis and the Principle of Acceleration". *Review of Economic Statistics*. 21 (2): 75–78.

Schumpeter J. (1939). *Business Cycles: A Theoretical, Historical and Statistical Analysis of the Capitalist Process*. New York and London: McGraw Hill.

Sims, C. (1980). “Macroeconomics and Reality.” *Econometrica*, 1-48.

Stock, J. and M. Watson (2019). *Introduction to Econometrics*. 4th edition, New York: Pearson.

Stock, J. and M. Watson (2003a). “How Did Leading Indicator Forecasts Perform during the 2001 Recession?” FRB Richmond Economic Quarterly 89: 71-90.

Stock, J., and M. Watson (2003b). "Forecasting Output and Inflation: The Role of Asset Prices." *Journal of Economic Literature*, 41 (3): 788-829.

Stock, J. and M. Watson (1989). “New Indexes of Coincident and Leading Economic Indicators.” in Blanchard and Fischer (eds.), *NBER Macroeconomics Annual*, Cambridge: MIT Press.

Verbeek, M. (2012). *A Guide to Modern Econometrics*, 4th edition. Chichester: Wiley.

Wheelock, D. and M. Wohar (2009). “Can the Term Spread Predict Output Growth and Recessions? A Survey of the Literature.” Federal Reserve Bank of St. Louis *Review*, September-October 2009, 419-40.

Appendix A: Variables and Sources

The following variables and their sources were used throughout the paper:

spread	difference between the 10-year Government Bond (Fred) and the 3-Month Treasury Bond (FRED)
ur	U.S. unemployment rate, Seasonally Adjusted (FRED)
recession	1 if in recession, 0 otherwise; based on NBER recession dates
Year_Before_Recession	1 if during 12 month prior to recession, 0 otherwise
age	the length of the expansion: 1 is assigned to the first period after the end of the previous recession
hrsManu	Average Weekly Hours of Production and Nonsupervisory Employees, Manufacturing (FRED)
shManu	manufacturing employment share; calculated as manufacturing employment/Total NonFarm Employment (FRED)
housstart	Housing Starts: Total: New Privately Owned Housing Units Started (FRED)
housstock	cumulative housing starts in each expansion. Calculated as the sum of all housing starts from the first period after the end of the previous recession to the current period
CSI	monthly University of Michigan Consumer Sentiment Index since 1978; before 1978, the data is interpolated based on non-monthly observations

FRED is the Federal Reserve Economic Database maintained at the St. Louis Federal Reserve.

The sample period is January 1960 to December 2018. Recession months, as dated by the NBER, are omitted as are months of the initial recovery (until employment returned to pre-recession levels. As a result, the observations we used were as follows: 1960:M1 - 1960:M4, 1962:M1 - 1969:M12, 1971:M6 - 1973:M11, 1976:M1 - 1980:M1, 1981:M1 - 1981:M7, 1983:M12 - 1990:M7, 1993:M3 - 2001:M3, 2005:M2 - 2007:M12, 2014:M5 - 2018:M12. As mentioned in the main text, we re-estimated our preferred specification using a slightly longer sample period, which is the result of replacing employment recovery with real GDP recovery to pre-recession levels.