Forecasting Economic Crises
An Empirical Approach

Submitted by:
Felix Edelthalhammer
Ismail Ismail
Jonas Volkenand

As part of:
PaRE1To Project: „Research in Empirical Economic Topics“

Supervisors:
Julian Dieler
Prof. Dr. Gebhard Flaig

March 2016
# Directory

1. Introduction ........................................................................................................... 1  
2. Theory ....................................................................................................................... 2  
   2.1 Conventional Methodology ............................................................................... 2  
   2.2 Binary Model ...................................................................................................... 3  
   2.3 Our Model .......................................................................................................... 4  
3. Application and Outcomes .................................................................................... 7  
   3.1 Model Quality Evaluation .............................................................................. 7  
   3.2 Comparison ........................................................................................................ 8  
   3.3 Empirical Analysis ............................................................................................ 9  
4. Conclusion ................................................................................................................ 20  
Sources ....................................................................................................................... 22  
Appendix ...................................................................................................................... I
1. Introduction

The past several crises, most recently the so-called financial crises around 2008, have shown two things: First of all there exists room to improve economic forecasting indicators and secondly, a well-working early warning system for economic crises could be an extremely helpful tool to prevent great distortions and to help central banks, public institutions and private companies to react and take countermeasures soon enough and avoid higher losses.

In previous decades economists have developed scientific, objective methods for a number of different topics in forecasting. One of the most enforced topics in today’s economic research is surely the forecasting of the GDP.

This paper will focus on an often less noticed but, when it comes to practicability, not less important area of forecasting: Trying to explore the possibilities and capabilities of leading indicators in predicting economic crises with a sole focus on promptly available and barely subsequently revised data. Thereby we use recessions as a proxy for crises. The exact definitions will be mentioned later in the paper.

We start our endeavor by analyzing past and current economic forecasting efforts. Without question, a fair piece of success has been achieved, as far as Short-, Mid- and Long-term forecasts of real production and existing leading indicators are concerned. But most of the work shows some kind of common and fundamental issues such as significant stability problems in forecasting continuous growth rates over several periods as Estrella, Rodrigues and Schich mentioned in their paper 2003. Furthermore, there seems to be a strive in research towards meeting the requirement of a “general purpose indicator” working well for all countries for all periods of time. We strictly refuse to meet this claim. Like Stock and Watson showed in 2003 it does not seem possible to create such a universally applicable predictor. We therefore chose to accept differences between countries and coherences of their markets and focus on finding the right predictors or a combination of them separately for each of our observed countries, namely Germany and the United States of America. To do so, we extend an existing leading indicator model and create ways to pool the results in order to find the best fitting model for both countries observed. In section 2 we will discuss the actual research and theory and introduce our model. Section 3 gives an empirical analysis and section 4 concludes.
2. Theory

2.1 Conventional Methodology
The existing literature on economic forecasting - especially GDP growth rates or the absolute level of real output - contains a variety of different approaches towards distinct problems. For example, short term predictions deal with the problem of a deferred date of availability in macroeconomic data. Widely recognized methods in short and mid-term forecasting are Autoregressive Moving Average Models (ARMA) which primarily exploit information on autocorrelation of a time series to predict its future values. This basic model has been enhanced by various steps and fitted to a number of application areas using not only the particular variables past but also information of different variables with explanatory power to map the development of the initial dependent variable. Beside ARMA there are many other approaches towards the construction of leading indicators to forecast e.g. the GDPs future progress.

Leading indicators are based on a few assumptions. The most relevant are forward looking markets and individual actors being able to anticipate future developments in their current behavior which is then reflected in time series data, for instance stock market prices. This popular example describes stock market prices as the expected, discounted future earnings of companies. Therefore a stock market index can be expected to be a leading indicator for the future condition of firms and the economy in its entirety. Other examples, which we will come across more often in the further course, are interest rate spreads in government bonds. Researchers discovered a forerunning correlation of these spreads with future economic growth. One way to interpret this finding is that forward-looking actors anticipate expected future economic conditions. Whenever market players expect great economic progress, they start selling long term (e.g. 10 year) bonds because they fear rising inflation rates going along with a rise in interest rates (Fisher Equation). As the nominal yield on ordinary Treasury bonds is fixed, selling them in order to buy similar assets in the future of higher interest rates is the answer. The escalating supply causes a fall in prices and leads to increasing real returns, which results in a steeper yield curve and an increasing bond spread (Bernard Baumohl, 2005).

To forecast long term horizons, simulation studies can be used, however those suffer from partly restrictive assumptions and can only help to draw a vague idea of what will happen 30 years from now, which brings us straight to problems that go along with the usage of the aforementioned concepts. The most important issue of conventional models is instability, meaning that it is hardly possible to maintain a satisfying forecast performance over time and countries:
“In short, there appear to be no subset of countries, predictors, horizons, or variables being forecast that are immune to this instability.” (Stock & Watson, 2003, p. 812)

In the next section, we will introduce a contribution that helps to tackle the just mentioned problems.

2.2 Binary Model

General model specifications:

<table>
<thead>
<tr>
<th>t: Date of regression</th>
<th>n: Forecasting horizon</th>
<th>x: The independent variable</th>
<th>Y: The dependent variable (For example in a binary model Y equals 1 for a recession in a period, otherwise it equals 0.)</th>
</tr>
</thead>
</table>

\( k \): The lag parameter \( k \) (\( \geq n \)) describes the magnitude of temporal delay between the independent and the dependent variable.

A graphical illustration will yield the following:

With a lag of three periods, the delay between the independent variable \( x \) and the dependent variable \( Y \) is illustrated as follows:

\[ k = 3, t = 216 \]

While a lag of 36 periods is illustrated in this graph:

\[ k = 36, t = 216 \]

Correlating \( Y \) and \( x \) with different lags shows up existing leads or lags between the dependent and the independent variable. The effects of \( x \) variables in past periods on future periods \( Y \) values are measured by correlation and the temporal delay is determined by the lag parameter \( k \).

Table 1: General model specifications

Regarding the above mentioned stability issues Estrella, Rodrigues and Schich (2003) describe how these problems can be opposed by the usage of a binary probability model to forecast the likelihood of a recession in a specific period. With very limited manifestation possibilities of the dependent variable such models prove to be generally more stable. Estrella et al. therefore use a Probit regression similar to equation (1), in which the dependent variable takes the value 1 for the existence of a recession in a period \( t + n \) and 0 otherwise.

\[
P(Y_{t+n} = 1) = \Phi[\beta_0 + \beta_1 x_t]
\]

The Probit regression – using the cumulative normal distribution function as response function – delivers estimated P values indicating the probability of a preexisting recession in period \( t + n \), estimated with data available in period \( t \). Estrella, Rodrigues and Schich (2003) and Estrella and Trubin (2006) restrain their model to single indicators and a fixed forecast horizon.
Nevertheless the basic approach of such Probit models can be enhanced and used for a variety of indicators and lags. We will draw on this simple Probit regression as a basis for our own model with extensions to the procedure in two steps.

2.3 Our Model

Our basic intention is to construct leading crises indicators that work well, each for a specific country. We therefore take up the previously introduced model by Estrella and Trubin (2006) and extend it in two steps. The basic characteristics of the resulting “Optimal Lag Model” are: A fixed forecasting horizon, optimal selection of the lag \( k \) and rolling regression windows which are explained hereafter.

First step: For our model we use a fixed forecast horizon \( n \) in order to predict the time period \( t + n \). This enables us to evaluate the performances of alternative model setups. A fixed time horizon is essential for comparison but more importantly a reasonable convention, not only in economic forecasting. We therefore regress the effect of the independent variable on the dependent variable over the complete sample of observations with equation (2) and test for 6 or 10 (depending on \( n \)) different lags \( k \).

\[
(2) \quad P(Y_{t+n} = 1) = \Phi[\beta_0 + \beta_1 x_{t+n-k}]
\]

An optimal lag \( k_{opt} \) for the entire sample is then determined using the Akaike Information Criterion (AIC). The minimal AIC and hence its corresponding optimal lag \( k_{opt} \) and \( \beta \) parameters are then used to conduct a projection for \( t + n \). Accordingly, \( Y \) at \( t + n \) is explained by the independent variables value at \( t' = (t + n) - k_{opt} \). This method ensures a projection of \( t + n \) in the time period \( t \) with indicator data from the period \( t' \leq t \).

To ease the understanding of the model and be able to deliver its core statement we will describe it using the ten to one year spread of the federal government bond yield curve. We label the interest rate on ten year bonds \( i_{t}^{10} \) and accordingly, the one year rate \( i_{t}^{1} \). The spread \( \Delta i_t = (i_{t}^{10} - i_{t}^{1}) \) is our \( x \) variable. In our proposed model, the estimated value of the independent variable \( Y \) (its expected value) is the predicted probability of the occurrence of a recession in \( t + n \). The exact definition of recession is going to be mentioned later in our paper. In our example, \( x \) is the interest rate spread; in general \( x \) could be any data correlated to the future recessional behavior of the economy. Examples include fluctuating commodity prices such as the price of gold, silver or oil, stock markets or house price indices and many more.
If we closely regard the interest rate on ten year bonds $i_t^{10}$ the variable $t$ stands for the date of the “Probit” regression. We use a pseudo out of sample\(^1\) setup in order to evaluate the prognosis quality of our model.

We run a regression of the recession probability on the 10 to 1 year bond spread testing for 10 different lag-lengths (6-15 months) for every $t$.

Using the bond spread example with a selected $n = 6$:

1. $\text{Reg. } k=6 \Rightarrow P(Y_{t+6} = 1) = \Phi[\beta_0 + \beta_1(i_{t+6-6}^{10} - i_{t+6-6}^{1})]$

   ...

10. $\text{Reg. } k=15 \Rightarrow P(Y_{t+6} = 1) = \Phi[\beta_0 + \beta_1(i_{t+6-15}^{10} - i_{t+6-15}^{1})]$

Assuming that the minimal AIC belongs to $\text{Reg. } 5$, $k^{opt}$ equals 10. We then put all parameters from $\text{Reg. } 5$ and the spread value at $t' = (t + 6) - 10$ into equation (2). As a result we obtain the estimated probability of a recession in $t + 6$.

The selection of the optimal lag in the optimal lag model is an objective criterion to separate the lead/lag relationship between dependent and independent variables from the a priori given forecasting horizon. Furthermore it allows the operator to easily adapt this method to any possible dataset without restrictions on forecasting horizon or additional research on correlations.

**Second step:** Here we add rolling regression windows to our model. This means, that the Probit regressions are not conducted over the entire sample but over the particular length of a window. It seems like the coherences between different market variables vary over time and can therefore bias later regression results. With rolling windows we make sure to just catch up the influential impacts of the last periods. Moreover our setup lets the resulting crises probabilities be consistent over time and enables us to use thresholds more easily and with more certainty regarding the fact that a threshold value will uphold its relevance in the future as well. For the US we use windows of 180 observations (15 years; monthly data), for Germany we use windows of 40 observations (10 years, quarterly data) as they exposed the best performance.

---

\(^{1}\) A Definition can be found in the next section.
**Crises definition and Data:** To exert our model, we use data from two different countries on two different continents, Germany and the United States of America. Consequently, the definition of recession varies between both countries:

**For the United States of America** we use the National Bureau of Economic Research (NBER) method in order to define recession. The NBER maintains a chronology of the US business cycle in order to identify the dates of peaks and troughs that are the basis for identifying expansions and recessions. The period from a peak of the business cycle to the trough of the business cycle is a recession and the period from a trough to a peak is an expansion. A recession is defined by the NBER as a significant decline in economic activity spread across the economy and lasting longer than a few months. In choosing the dates of those turning points, the exact dates for the start and the end of a recession or an expansion are set (NBER).

**For Germany** we use a more basic procedure to identify a recession, the so called “Press” method, as it is widely used in news reports and in the public media. It relies on the idea that a recession occurs if the economic growth is negative in two consecutive quarters and the end of a recession (or the start of an economic expansion) is initiated by two consecutive quarters of positive economic growth rates (Schirwitz, Seiler and Wohlrabe, 2009).

The usage of different definitions of recession by two closely related economies is not necessarily of negative impact on the information and data analysis. Since the two observed economies are treated differently with country specific variables, which is a result of the basic idea not to find one universal optimal recession indicator for all countries, but rather a country specific optimal indicator that reflects the information and behavior of each individual economy.

**General data specifications:** As independent variables for our model we preferably use easy accessible, few revised and shortly published data that seems likely to be correlated with economic crises. Accordingly the majority of the data have a financial background, like interest rate spreads, commodity prices or stock market indices but also convenient national account system data and more. This setup enables us to supply the latest results at any given moment, it makes revisions unnecessary and therefore eases the process of back testing the indicators. A great part of our indicators are interest rate term spreads. This means that we use the differences in interest rates between bonds with diverse durations. Since we are interested in the variance of these rates caused by different runtimes and not by the higher risks of longer bond durations we use Treasury bill rates. Government bonds are widely recognized as risk free and should therefore
show a minimal bias in term spreads through risk markups. To catch up issuer risks we test data like BAA-AAA rating spreads.

For Germany, we use quarterly data going back until 1970, while for the US - where generally a lot more detailed data are available - we use monthly time series going back to 1962.

Our main source for US data is the Federal Reserve Bank of St. Louis (FRED), for Germany we mainly referred on information by the “Deutsche-Bundesbank” and “Statistisches Bundesamt”.

3. Application and Outcomes

3.1 Model Quality Evaluation

RMSE: In order to objectively measure our models prognosis quality and to compare it to alternative models, we draw on one of the most common empirical methods: The Root Mean Squared (forecasting) Error (RMSE). It is defined as displayed in (3):

\[
(3) \text{RMSE} = \sqrt{\frac{\sum_{t=0}^{T} (y_t - \hat{y}_t)^2}{T}}
\]

Thus the Squared Error equals the squared difference between the realization of the dependent variable \((y)\) and its earlier estimated value \((\hat{y})\) for each \(t\) in our sample. Adding up the “SE” over every predicted value from \(t = 0\) to \(T\), calculating its mean and take its root, results in the RMSE. Here \(t = 0\) is the first time period for which an estimated value \(\hat{y}\) exists and \(T\) the last period with already realized \(Y\).

Although the RMSE does not seem to be the optimal tool to evaluate the forecasting performance of a binary model, we decided to stay with it for a first model comparison but do not base our results entirely on it.

Pseudo out of sample: Since in sample analysis is not a proper way to examine the actual prognosis quality of time series models, out of sample approaches are used.

A “regular” out of sample analysis would be to estimate the model based on all the data up to, and including, \(y_t\) to forecast \(\hat{y}_{t+n}\), then wait \(n\) periods until \(y_{t+n}\) has realized and calculate the prognosis error. For obvious reasons, it is advantageous to just simulate the out of sample process from a retrospective point of view. This means that an earlier date than today is used as

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]

\[\]
The “regular” out of sample analysis method is then used on this previous starting point. Afterwards the resulting forecasting error term is used like in the “regular” out of sample analysis, to estimate the model’s forecasting abilities in all following periods (Stock and Watson 2012).

To check the forecasting performance of the different indicators, different prediction horizons and the variety of setups for our Optimal Lag Model, we apply the pseudo out of sample process.

### 3.2 Comparison

First of all, we have to answer whether our model can add quality to the existing and acknowledged process. As our model is based on the approach by Estrella and Trubin (2006) we test our setup against theirs using pseudo out of sample. We deploy the RMSE to measure the forecasting error.³

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE; 12 months forecast</th>
<th>RMSE; 6 months forecast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ordinary Model (binary)</td>
<td>0.3161</td>
<td>0.3153</td>
</tr>
<tr>
<td>Optimal Lag Model</td>
<td>0.3122</td>
<td>0.3147</td>
</tr>
</tbody>
</table>

Table 2: RMSE comparison

As Table (2) shows, we can consider our model as weakly preferable. The Optimal Lag Model has a smaller RMSE. More importantly it features a more objective selection of the used lag, not choosing it a priori. Also the lag selection process is automated, therefore less laborious, more universally and simply to apply. These results in mind, we now go on presenting our findings.⁴

The two graphics stated below display the effect of regression windows. When they are in use, the calculated Probability values are much more stable over time as mentioned before. This enables us to set threshold values as shown later.

---

³ For comparison we recreated Estrella and Trubin’s setup since they do not publish out of sample data.
⁴ We used the calculated term spread by Estrella and Trubin (2006) for comparison.
3.3 Empirical Analysis
In this section we give a brief overview on the structure we set up to unveil our results and start to present our findings starting with US data. Results for Germany are discussed afterwards.

As mentioned earlier, a fixed prognosis horizon is an absolutely sensible convention not only in economic forecasting but also in e.g. weather prognosis. This is due to the fact that we want to know whether it will be rainy or sunny tomorrow and the day after. So we captured this demand by presenting our results classified by the examined time horizons 3, 6 and 12 months. We tested a 9 months interval as well. Corresponding results can be found in the appendix as they do not differ from the other findings significantly. A list of all tested indicators can be found in the appendix, too.

We further structured the following presentation by starting to display our findings, computed with optimal lag selection and rolling regression windows engaged, by arranging them in the categories “single indicator”, “pooling of forecasts” - both with and without a “threshold” engaged - for each time horizon. Since not every tested dataset delivered good results, the appendix contains the here omitted graphs as they deliver interesting insights in the performance of all tested indicators.
12 Months Forecasting Horizon (US):

**Single indicator** (One independent variable)

In this section the outcomes of our Optimal Lag Model using rolling regression windows and a 12 months forecasting horizon are discussed. Figures 3-6 show the probability of a recession forecasted using S&P closing data in figure 3, gold prices in figure 4, risk spread BAA-AAA in figure 5 and 10 to 1 year term spreads in figure 6 for the United States economy. It is called “single indicator”, as one independent variable is consulted as an indicator.

The green graphs present the recessions that occurred in the US over the last 50 years. The oil crises in 1973, two recessions in 1980 and in 1981-1982, the recession of the early 1990’s, the dotcom bubble in 2001 and lastly the recession in 2007, well known as the “financial crises”.

The black graphs represent our findings using our model for each dataset observed. Figure 3, examining the S&P closing data, (nearly) perfectly described the recession of 1973 which was caused by the oil crises in that year due to the quadrupling of oil prices by OPEC countries. It also gave an idea of the recession of the early 80’s. On the other hand, S&P closing data failed to project the dotcom bubble of 2001, while the Great Recession of 2007 was nearly perfectly projected by our model. Interestingly, relying on the S&P 500 the US had to face a recession until 2015, which did not actually occur. Figure 6 takes a relatively similar path mostly picturing
recessions that are related to the dataset used, here the 10-1 term spread. Gold price data and the bond risk spread (Moodies BAA-AAA) do not deliver well working early indicators as their graphs mostly swing out too late, if they do at all.

**Threshold (single indicator)**

For the probability of recession given our model, we set a certain threshold value in order to receive discrete values of recession probability, either 0 for values below the threshold, which equals the assumption “No recession”, or 1 for values above the threshold which equals the assumption “Recession” and thereby try to obtain additional prognosis quality. If we set a threshold of 40% for instance, which is an a priori selected value, all the values higher than 40% will be set to one, and consequently all values lower than 40% will be set to zero. This process is demonstrated here. Figure 7 shows the 10-1 spread in bond rates indicator and the mentioned 40% line. Figure 8 displays the results. Due to the regression windows in use, we are confident to set an a priori threshold at a certain value because, as seen in figures 1 and 2, windows ensure constant probability swings of indicators over time.

![Diagram](image1)

*Figure 7: Term spread 10-1; 12 months; US; threshold=0.4 displayed*

![Diagram](image2)

*Figure 8: Threshold (0.4); 10-1 term spread; 12 months; US*
Pooling of forecasts and Threshold: (Results from more than one single indicator averaged.)

This approach takes the forecasting results of individual indicators into consideration and averages them in a certain way to assemble a new indicator. By pooling forecasts we try to improve single-indicator prognosis results. The process of averaging can vary across time horizons because we tried to find the best way of pooling each.

Figures 9 and 10 take the threshold method a step further and use it on pooled forecasts with 12 months forecast horizons. In figure 9, we create a new graph using the same requirements as before (12 months forecast horizons, rolling regression windows) and pool 4 different sets of data into one graph in order to strengthen the predictive power of the model. We see that the graph of the pooled forecast horizons gives a better projection than each single indicator on its own. Using the threshold method on the pooled 12 months forecast, we yield the red graph in figure 10; we see an improvement to our model’s predictive power.

---

5 We used different methods to find the best pooling possible. If not labeled otherwise we used a priori chosen fixed weights for the subcomponent indicators.

6 Establishing a threshold with the data available until today is not a (pseudo) out of sample method. Nevertheless we are confident that the chosen threshold values will hold true for the future because of the regression windows setup.
6 Months Forecasting Horizon (US):

Single indicator

The same methods we deployed to project the 12 months fixed forecast horizon are now used in order to project the 6 months forecast horizons. We see some differences in the results of the graphs as the forecast horizon is a far shorter time period; however the purpose of the graph analysis stays the same. For the 6 months horizon we here show the findings of both the 5 and 10 to 1 term spreads as well as Building Permits and the Investment share in the GDP of the US. As one can see, some of the graphs manage to identify a number of past recessions, namely the 5 and 10 to 1 term spreads. Building permits and the investment share of GDP also contribute to the projection of recessions, bearing in mind that using a threshold value will further change the
outcome. Figure 12 shows our finding that 1\textsuperscript{st} differences unlike presumably thought, do not have a strong explanatory power on recessions in our model.

**Threshold (single indicator)**

We use a threshold value here as well in order to receive discrete values of recession probability. Figure 17 shows the threshold method applied to the 10-1 term spread with a 6 months horizon, as demonstrated in figures 7/8 before. In comparison to figure 8, it delivers slightly more precise results as the time interval between date of prognosis and realization has decreased and therefore more information are available.

![Figure 17: Threshold (0.5) single indicator; term spread 10-1; 6months; US](image)

**Pooling of forecasts and Threshold:**

Figures 18 and 20 are the pooled graphs for the 6 months fixed forecast horizon. These graphs use the same requirements as before while pooling two respectively seven datasets into one graph to further strengthen the projection power of the resulting indicator. Interestingly the 10-1 spread combined with the 1 year-funds rate spread has an almost equal performance as the pooling setup extended by five other independent explanatory variables. The pooled 6 months fixed horizon graph used with a threshold can be seen in figure 19 and 21.
Here the precision, especially in the early nineties, leaves room for further improvement. But both graphs manage to project every recession in the given timeframe, making this setup a viable projection source.

**3 Months Forecasting Horizon (US):**

**Single indicator**
Due to the short nature of the 3 months forecast, the indicators tend to project the recession later than the 3 months forecast horizons which means they tell us about a potential recession after it has already begun. Because this phenomenon only appears infrequently we decided to show 6 graphs for the fixed 3 months horizon as well, including the term spreads used before, building permits and the Purchasing Manager’s Index. The last named indicator was tested on the other horizons too, but it had not shown predictive capability there.

**Threshold (single indicator)**

Again, we use a threshold value here. Figure 28 displays the 10-1 term spread with the threshold (40%) engaged. Due to the short nature of the forecast horizon used here, we can see that the 3 months fixed forecast horizon describes all the recessions depicted above better than 6 or 12 months horizons. This result becomes obvious analyzing figures 29-32 in comparison to figures 18-21 or especially 9 and 10.
Pooling of forecasts and Threshold:

Figures 29 and 31 show the pooled graphs for the 3 months fixed forecast datasets. We pooled four respectively five single indicators into one graph in order to further strengthen the projection power of the resulting indicator. Just like the 12 and 6 months forecasts, pooling here gives us a better projection than each single indicator on its own, making the further analysis of the graph (as well as the usage of thresholds) easier and objectively more fitting. Interestingly, quite complicated weighting-methods for pooling, like weights based on the former variance of each indicator (figure 31), add no obvious superior forecasting power in comparison to the simple fixed weights in figure 29. This finding is consistent with actual literature (e.g. Stock and Watson, 2003).
The pooled 3 months fixed horizon graph used with a threshold can be seen in figures 30 and 32. The pooled indicators with a threshold of 20% (Fig. 30) and 35% (Fig. 32) are valuable additions to the common forecasting methods used. At first sight, the graphs project every recession in the viewed timeframe, making the model a viable projection source, which it partly is. Sadly though, the resulting indicators swing out too late in some cases.

The results for Germany - with its quarterly data basis and unlike market coherences - differ from the US results and are discussed in the following:

**4 Quarters Forecasting Horizon (GER):**

![Figure 33: 10-1 term spread; 4 quarters; GER](image)

![Figure 34: 5-1 term spread; 4 quarters; GER](image)

**2 Quarters Forecasting Horizon (GER):**

![Figure 35: 10-1 term spread; 2 quarters; GER](image)

![Figure 36: 24-6 term spread; 2 quarters; GER](image)
The strong forecasting relation does not quite hold for Germany like it does for the United States. While some economic crises can be forecasted precisely, the model fails for others. Surprisingly an automatic improvement of the performance with shorter horizons cannot be observed. Instead some indicators seem to work better for longer and some for shorter predicting periods. However a pooling of four indicators in one quarter predictions accomplishes to detect all crises, except the dotcom bubble in 2001, with an acceptable forerun. Furthermore the results approve our assumption, that the optimal lag model can easily be adapted to other countries and indicators because of its resistance against breaks in market coherences and stability issues and its automatic combination of variables and their optimal lags.

**Application suggestion**

Various indicators for different forecasting horizons show a good performance in detecting economic crises. To create one well working specification of the model for the United States we consider using a six months pooling of either the two spreads 10-1 and 1 year-funds rate or the
more complex pooling of five indicators mentioned above. This simple pooling with fixed a
priori chosen weights even with only two indicators is able to detect all US recessions since 1972
with sufficient forerun making it a valuable tool to timely inform about future economic crashes.
Obviously the model does not hit all recession timeframes exactly. Indeed our basic aim is not to
do so but to get early hints of arising economic crises in the next months. This aim is perfectly
matched by the specification of the model in the United States. For Germany a combination of
indicators is at least capable to foresee most of the past recessions with an acceptable forerun. A
first attempt can be to use a simple 24-6 months term spread on a two quarter forecasting basis,
which detected most of the past recessions. More data has to be analyzed in future research.

4. Conclusion
Bearing in mind the past efforts of leading economists, we have created a crises indicator that is
based on the efforts of Estrella and Trubin (2006) and further adjusted it in order to reach a better,
more general model. Our aim was to construct a country specific indicator to obtain an early
warning system for economic crises. Fixed forecasting horizons are considered as user friendly
and enable us to compare results more easily.

The first step we took to establish our model was adding a method to determine the lag with the
highest explanatory power for any indicator dataset at any point of time, instead of using fixed
lead-lag relations. This of course is what makes our model “optimal lag” and enables us to easily
apply it to a lot of possible variables. Moreover this process is an - in the actual literature often
missing - objective criteria for selecting the right lags.

The second step was to use rolling regression windows. This gives our model a decisive
advantage. It makes us only catch up the relevant impacts of the last periods and furthermore
adding rolling regression windows lets the resulting probabilities be constant over time which
allows us to use thresholds.

The newly created model is able to project the past recessions including the very rarely predicted
recession in the US in 2007/2008 with at least 6 months forerun by using the pooling of forecasts
and the threshold value we thoroughly explained in our paper. We furthermore achieve stable
results over several time periods with easily accessible data. The strong explanatory power of our
US indicators gains its strength from the vast availability of data, a criterion that lacks for
countries like Germany. Despite not reaching an optimal level yet, we can find evidence for the
predictive potential of our model which has indicated at least some of the worst economic fluctuations, even for Germany.

In conclusion our model grants an objective way to analyze the coherences between the economic development and different market indicators and is able to detect and show up future crises at an early point. Due to its binary design and rolling windows it is furthermore able to hold the predicting relations stable over time and it is less fragile to systematic changes. Nonetheless we see room for improvement and further research in economic forecasting in general and our approach as part of this in specific.
Sources

Literature:


Data:


German GDP Data: Statistisches Bundesamt: https://www.destatis.de/DE/Publikationen/Themen/ VolkswirtschaftlicheGesamtrechnungen/Inlandsprodukt/InlandsproduktberechnungLangeReihen.html

(1) Federal Reserve Bank of St. Louis (Fred): https://research.stlouisfed.org/fred2
(2) University of Michigan: http://www.sca.isr.umich.edu/tables.html
(3) Yahoo Finance: http://finance.yahoo.com/q/hp?s=^GSPC+Historical+Prices
(4) UNstat: http://unctadstat.unctad.org/wds/TableViewer/tableView.aspx?ReportId=30727
(5) finanzen.net: http://www.finanzen.net/rohstoffe/goldpreis/Chart
(6) statista.de: http://de.statista.com/
(7) FED of New York / Arturo Estrella: (https://www.newyorkfed.org/research/capital_markets/ycafaq.html
(10) multpl.com: http://www.multpl.com/table?
Appendix

<table>
<thead>
<tr>
<th>Source</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data USA</td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>10 year - 1 year U.S. Treasury bonds rate</td>
</tr>
<tr>
<td>(7)</td>
<td>10-year – 3 months bond rate basis data from Estrella</td>
</tr>
<tr>
<td>(1)</td>
<td>5 year- 1 year U.S. Treasury bonds rate</td>
</tr>
<tr>
<td>(1)</td>
<td>Risk Spread Moody’s BAA – AAA bonds</td>
</tr>
<tr>
<td>(1)</td>
<td>Risk Spread Moody’s BAA – 10 year U.S. Treasury bonds</td>
</tr>
<tr>
<td>(3)</td>
<td>Standard &amp; Poors 500 Index close values (monthly mean)</td>
</tr>
<tr>
<td>(3)</td>
<td>Volume of S&amp;P 500 Index</td>
</tr>
<tr>
<td>(1)</td>
<td>1 year Treasury bond – funds rate</td>
</tr>
<tr>
<td>(10)</td>
<td>Price Earnings Ratio of S&amp;P 500</td>
</tr>
<tr>
<td>(5)</td>
<td>Gold price</td>
</tr>
<tr>
<td>(1)</td>
<td>Monthly Building permits USA</td>
</tr>
<tr>
<td>(1)</td>
<td>Oil price (WTI)</td>
</tr>
<tr>
<td>(9)</td>
<td>Oil Price (Brent)</td>
</tr>
<tr>
<td>(4)</td>
<td>Median value of Texas and Brent Oil</td>
</tr>
<tr>
<td>(4)</td>
<td>Oilg; Crude petroleum, average of UK Brent (light), Dubai (medium) and Texas (heavy), equally weighted ($/barrel)</td>
</tr>
<tr>
<td>(2)</td>
<td>Consumer Sentiment Index</td>
</tr>
<tr>
<td>(1)</td>
<td>Purchasing Manager Index</td>
</tr>
<tr>
<td>(1)</td>
<td>Personal Savings Seasonal Adjusted</td>
</tr>
<tr>
<td>(1)</td>
<td>Motor Vehicle Retail Sales, Units, Seasonally Adjusted</td>
</tr>
<tr>
<td>(1)</td>
<td>Wilshire US Real Estate Investment Trust Total Market Index</td>
</tr>
<tr>
<td>(1)</td>
<td>Real Retail and Food Services Sales, Seasonally Adjusted</td>
</tr>
<tr>
<td>(1)</td>
<td>Invests - Investments as share of GDP</td>
</tr>
<tr>
<td>Data Germany</td>
<td></td>
</tr>
<tr>
<td>(8)</td>
<td>10 years – 1 year German Government Bond rates</td>
</tr>
<tr>
<td>(8)</td>
<td>1 year – 0.5 years German Government Bond Rates</td>
</tr>
<tr>
<td>(8)</td>
<td>2 years – 0.5 years German Government Bond rates</td>
</tr>
<tr>
<td>(8)</td>
<td>2 years – 1 year German Government Bond rates</td>
</tr>
<tr>
<td>(8)</td>
<td>5 years – 1 year German Government Bond rates</td>
</tr>
<tr>
<td>(8)</td>
<td>20 years – 10 years German Government Bond rates</td>
</tr>
<tr>
<td>(6)</td>
<td>Ifo Index (Geschäftsklima)</td>
</tr>
<tr>
<td>(6)</td>
<td>ZEW Sentiment</td>
</tr>
</tbody>
</table>

The following graphs are the results of the data we tested over the various forecasting time horizons. They did not make it into the main part of the paper, mainly because they do not, or only to a small extent have predictive power but still they deliver interesting insights.
Indicator Data USA:

**PMI 12-months**

**CSI 12-months**

**Brent Tex oil median 12 months**

**SP Volume 12-months**

**Brent Oil 12-months**

**Oilg 12-months**

**Price Earnings Ratio S&P 12-months**

**Texas Oil 12-months**

**Spread Estrella 12-months**

**Spread Baa-10y gov treasury 12-months**
In this section lists of the calculated p-values and selected lag lengths can be viewed for the 10-1 termspread example over a 12, 6 and 3 months forecasting horizon:
Indicator Data Germany:

- Spread 2y-1 Germany 4-quarters
- Spread 1y-6m Germany 4-quarters
- ZEW Situation Index Germany 4-quarters
- ZEW Sentiment Index Germany 4-quarters
- Ifo Germany 4-quarters
- Spread 2y-6m Germany 4-quarters
- Spread 20-10 Germany 4-quarters
- Spread 1y-6m Germany 2-quarters
- Ifo Germany 2-quarters
- Spread 1y-6m Germany 1-quarter