Towson University
Department of Economics
**Working Paper Series**

Working Paper No. 2023-05

# Could Diffusion Indexes Have Forecasted the Great Depression?

By Gabriel Mathy and Yongchen Zhao

September 2023

# COULD DIFFUSION INDEXES HAVE FORECASTED THE GREAT DEPRESSION? *

Gabriel Mathy†
*American University*

Yongchen Zhao‡
*Towson University*

**ABSTRACT:** Was the Depression forecastable? In this paper, we test how effective diffusion indexes are in forecasting the deepest recession in U.S. history: the Great Depression. Moore (1961) considered the effectiveness of diffusion indexes historically, including for the Great Depression, though he only did so retrospectively and did not forecast out-of-sample. We reconstruct Moore's diffusion indexes for this historical period and make our own comparable indexes for out-of-sample predictions. We find that diffusion indexes, including the horizon-specific ones we produce, can nowcast turning points fairly well. Forecasting remains difficult, but our results suggest that the initial downturn in 1929 may be forecastable months before the Great Crash. This is a novel result, as previous authors had generally found the Depression was not forecastable.

**KEYWORDS:** Diffusion Index; Great Depression; Forecasting

**JEL CODES:** N12 · C53 · E32 · E37

---

* This paper is dedicated to the memory of Herman Stekler, who suggested this research topic before his passing and inspired our work on this topic. We thank audiences at the Federal Forecasters Conference and the H.O. Stekler Research Program on Forecasting Seminar at George Washington University for helpful comments. We would like to thank the editor and two referees for useful comments as well.
† mathy@american.edu. Department of Economics, American University, 4400 Massachusetts Avenue NW, Washington, DC 20016, United States.
‡ yzhao@towson.edu. Department of Economics, Towson University, 8000 York Road, Towson, MD 21252. United States.

# COULD DIFFUSION INDEXES HAVE FORECASTED THE GREAT DEPRESSION?

## 1. INTRODUCTION

The Great Depression was the deepest downturn in US history and was more severe than any economic crisis before or since. Being able to forecast the turning points of such a deep downturn would be extremely valuable for anyone, but particularly for policymakers, asset holders, and business firms. Previous studies have examined the degree to which the Depression could be forecasted. Dominguez et al. (1988) found that the Depression was not forecastable, even using state-of-the-art time-series techniques (which, at the time, were based on vector autoregressions). They also found that Harvard and Yale based forecasting groups could not forecast the Great Depression. Other authors have also found that the path of output and the price level in the Depression were not forecastable.[4] The dispersion of forecasts widened during the Depression too, consistent with increased difficulty in forecasting (Romer 1993). Cortes et al. (2021) had some success forecasting the initial downturn in 1929 using a bank stock index they developed, though we find that bank stock prices do not provide good forecasting ability when the entire 1930s is considered.

Which series will be useful in forecasting can also inform our view of the origins of the Depression. Friedman and Schwartz famously argued that the origins of the Depression in the United States were monetary, originating in tight monetary policy by the Fed in the late 1920s (Friedman and Schwartz 1963). Temin (1976) argued that the Depression saw low interest rates and little tightening by the Fed and that Keynesian explanations were superior, though he did not provide a specific causal mechanism. Eichengreen (1996) provided an explanation for the Depression centered on the flawed interwar gold standard which generated economic collapse for countries that stayed on gold and which allowed for recoveries in countries that abandoned the barbarous relic. Kindleberger (1986) provides an international perspective focused on the flawed

[4] See Goldfarb et al. (2005), Hamilton (1992), Klug et al. (2005), and Mathy and Stekler (2018), *inter alia*.

interwar gold standard as well, arguing that the United States could not manage the global system as the British had before World War I. While monetary explanations are central to many explanations of the Depression, monetary variables do not figure prominently in our analysis. We find that real quantities of production are most important for forecasting depressions using our techniques, though this does not imply that monetary forces were not important indirect determinants of the Depression of the 1930s.

Diffusion indexes are a useful forecasting method which are less popular than they were in the past. A seminal example of this tool comes from Moore (1961). He applied it to historical data at the time, including the Great Depression period of the 1930s. A diffusion index is an aggregate of individual series. Its value equals the percentage of its component series that are expanding. The more series that are expanding, the less likely a recession is underway; and the more series that are contracting, the more likely a recession is in progress. Diffusion indexes have been successfully used in the modern period for forecasting,[5] so we apply this technique to the Great Depression to see how effective diffusion indexes could be in forecasting this severe downturn.

Although diffusion indexes can be used to forecast different aspects of the business cycles, here we mainly focus on the turning points.[6] They are notoriously difficult to identify in advance, but they are important for economic actors to forecast (Stekler 1972). NBER recession dates are used: The business cycle peak was reached in August 1929, two months before the Great Crash on Wall Street, which is often seen by the laypersons as heralding the start of the Depression. The trough occurred in March 1933. The next expansion lasted until May 1937, followed by a recession until June 1938. The US economy then would not see a recession until after World War 2. Using data up to December 1940, we focus solely on the pre-war period. Of primary interest is the ability of diffusion indexes to forecast the Great Contraction period from 1929-1933 when real GDP fell by a third. We also show the performance of the diffusion indexes in the 1920s and consider later business cycles in the 1930s as well.

We begin with Moore's analysis as a baseline, as he applied a diffusion index to the Great Depression already. We update this analysis with some additional series and see if we can improve on these methods. Moore only used a crude diffusion index based on whether more series were

[5] See Zhao (2020) and references therein.

[6] We also examine the ability to forecast the severity of the business cycle in the Depression in an online appendix.

expanding or contracting. We propose improvements to his diffusion index methodology and attempt to construct horizon-specific indexes. In the process, we explore the potential limitation of the traditional diffusion index imposing equal weights on its components. We also evaluate the performance of the diffusion indexes alongside forecasts made using more modern econometric techniques such as factor-based models, which are nowadays intuitively interpreted as "diffusion index" but require much more sophisticated calculations to construct. In our exercises, we use real-time data when it provides more information than historical forecasts.[7] We find that, despite their simplicity, our diffusion indexes have solid performance as forecasting methods, showing properties similar to or better than alternative techniques.

The rest of the paper is organized as follows: Our data set is introduced in the next section. We present the design of our empirical exercises in Section 3, including how we construct various diffusion indexes. Section 4 contains our main results on the forecasting performance of the diffusion indexes. The next two sections consider extensions and robustness checks: We augment our data set with real-time data vintages and repeat the main exercise in Section 5. We consider making forecasts at a quarterly frequency in Section 6. The last section contains some concluding remarks. An appendix, to be located online, is added to the end of the manuscript.

## 2. DATA

We first reconstruct the diffusion index used by Moore (1961) using the latest data. The series he used are listed in Table 1. We are able to locate the series he used or close substitutes in all cases. The Standard Statistical Company Industrial Output Index is only available in the publications of the Standard Statistical Company, but this publication ran for a few years only and is hard to find today. Instead, we substitute the "Index of Industrial Production and Trade for United States" compiled by Barron's Magazine. As this is a similar index compiled by a private sector organization, it should contain similar information. After reconstructing Moore's index, we then see if we can improve on Moore's diffusion index by adding other variables that we thought might

[7] For many series, any data revisions are small enough to make little difference.

be useful for forecasting. In general, as can be seen below, Moore did a pretty good job, and our additional variables don't improve forecasting performance too much.[8]

[Table 1 here]

We begin with monthly variables. We add bank debits for both New York and 140 cities outside of New York. These data are transactions that correspond to expenditures, and they correspond closely to economic activity in this period. We include bank rates as well, since they are also related to the business cycle in this period. To get additional disaggregated information on production, we include a diffusion index of different components of industrial production, as well as a diffusion index of eight leading indicators.[9] The index of American Business Activity by the Cleveland Trust Company, created by a respected contemporary forecaster, Col. A. P. Ayres, is also included as an additional business cycle indicator. Furthermore, we include a few manufacturing variables that are highly cyclical, including factory employment in automobiles, the gross hiring rate in manufacturing, and the layoff rate in manufacturing. We also include security issuance by corporations as well as state and local governments, as these are correlated with the business cycle. Retail sales are included as well. They co-move closely with the overall business cycle in this period. In addition, we include the percentage of steel ingot production capacity utilization – another highly cyclical industrial indicator. Finally, we include the wholesale price index for all commodities excluding farm products and food. This can be seen as an analogue to the modern "core" price index, excluding changes in food prices that may fluctuate in ways uncorrelated with business cycles.

[8] A potentially useful indicator of the financial market is discussed in Cortes et al. (2021), who showed that their hand-collected bank stock index could explain a significant portion of the variations in industrial production during the Great Depression. Their index is proprietary. We hand-collected data from the same source and constructed a similar index ourselves. While we find that our version of the index peaked before the 1929 business cycle peak, we were unable to acquire data to construct a long enough time series for our purpose. Also considering that Moore's data set already contains the Dow-Jones Industrial Stock Price Index, we thus leave this bank stock index out of our subsequent empirical exercises.

[9] These series are: (Inverted) Liabilities of Business Failures; Industrial Stock Prices; New Orders for Durable Goods Manufacturers; Residential Construction Contracts; Commercial and Industrial Construction Contracts; Average Workweek in Manufacturing; Number of New Business Incorporations; and Basic Prices (Moore 1961).

Although our main empirical exercise focuses on constructing monthly diffusion indexes, we also consider additional quarterly data in an extension. We add an index of the profits of major manufacturing corporations as well as an index of freight car loadings for 19 commodity groups. We also add an index of railroad freight tons originated. In addition, we use some early estimates of Gross National Product from Barger (1942) and three measures of investment: manufacturing inventory investment, new capital formation of plant, and new capital formation of equipment. The full list of series used is shown in Table 1. Data sources include the NBER Macrohistory database, the Saint Louis Fed's FRED database, and the Bureau of Economic Analysis's Survey of Current Business.

In the exercises below, we use historical data as our baseline. Real-time data, when available, are used to check the robustness of our main conclusions. During our period of interest, data revisions were less common than they are today, though some did occur. Industrial Production is updated and revised frequently by the Federal Reserve Board, and we use the real-time data vintages from ALFRED by the Saint Louis Federal Reserve Bank. Among the rest, variables reported as indexes are the most frequently revised. For those, we use data from the appendixes of the Survey of Current Business as this publication contained an extensive appendix of real-time data. Specifically, we have real-time data for Pig Iron Production, Steel Ingots Production, and Automobile Tire Inner Tubes Production from December 1926 to December 1940 and for Steel Ingot Production Capacity from January 1927 to December 1940.

## 3. DIFFUSION INDEXES FOR PREDICTING THE GREAT DEPRESSION

The research program to measure business cycles saw numerous advances in the interwar period, with the best example being the seminal work of Burns and Mitchell (1946). Rhodes (1937) advocated principal component analysis, arguing that the first principal component of a series was a useful measure, which was later expanded upon by Stock and Watson (2002) to create a series of monthly indicators for the U.S. economy. Mitchell et al. (2012) also used techniques that somewhat resemble diffusion indexes to construct monthly indicator for the interwar British economy. For more on business forecasting in the interwar period, see Lenel (2018, 2021). Haney (1931) provided a list of indicators that can be used to forecast turning points based on previous experience in forecasting during the 1920s, though this was found to not be helpful in forecasting turning points in Mathy and Stekler (2017).[10]

The term “diffusion index” as widely known in all but recent literature refers to the analytical tools developed by Geoffrey Moore at the NBER in the 1950s (see Moore (1954)). NBER researchers built on and expanded the literature on diffusion indexes, and the methodology has since been employed by research institutions and government statistical agencies, e.g., Getz and Ulmer (1990). Early evolutions of the index are summarized by Broida (1955). Although Stock and Watson later popularized the use of this term to refer to “estimates of the unobserved factors in a dynamic factor model” (Stock and Watson 1998, 2002), the diffusion indexes and their variants based on Moore’s initial contributions remain widely and regularly used today by statistical agencies.[11] While we do not dispute the interpretation of Stock and Watson, to avoid any confusion, we do not refer to estimated factors in a factor model as diffusion indexes in this work.

Our main empirical exercise involves constructing three predictors and evaluating their ability to forecast the Great Depression. Of the three, we are primarily interested in the first – a diffusion index in the spirit of Moore (1954). This exercise is designed for two primary purposes. First, we determine whether the peak in 1929 and the trough in 1933 are foreseeable, and if they are, how far in advance. Second, we test whether the variables used by Moore, as well as the additional ones we select, contain information useful for this forecasting task – and in particular, whether using the additional variables allows us to achieve improved forecasting performance.

While Moore constructed his index with the benefit of hindsight, we restrict our use of future information as much as is practical to simulate a contemporary forecast. Specifically, we construct our index in a manner similar to how a typical recursive out-of-sample forecast is made: For making a forecast for time period $t$ with a horizon $h$, the estimation sample includes observations up to $t - h - 1$. Taking a nowcast (i.e., $h = 0$) for period $t$ as an example, this nowcast would be made using a training sample that contains data up to $t - 1$, since the data *for* period $t$ would not be available *during* period $t$. Generally, as the training sample expands, the forecasting model is re-estimated, so the nowcast for $t + 1$ is made using information up to and

[10] A more in-depth discussion of Haney’s forecasting methods can be found in Mathy and Stekler (2017), Appendix C.

[10] Examples include the US Bureau of Labor Statistics, e.g., Lepoutre (2022), as well as institutions such as the Conference Board and the Institution for Supply Management.

including that from period $t$. In addition to limiting the data we use to make each forecast, we impose a few other restrictions detailed below. Despite our best effort, it is unavoidable that some aspects of our exercises, such as what variables we use to augment Moore's data set, benefit from knowledge unavailable to forecasters in the 1930s.

The first of the three indexes we construct largely resembles Moore's work. The index itself is in essence a standard diffusion index that shows the proportion of its components expanding. Let $c_j$ be the $j$th component of a diffusion index $D$ with $j = 1, 2, \dots, N$, we have

$$D_t = \frac{1}{N}\sum_{j=1}^{N} \mathrm{I}\left[T_j\left(c_{j,t}\right) > 0\right], \tag{1}$$

where $\mathrm{I}(\cdot)$ is the indicator function and the function $T_j(\cdot)$ transforms the data appropriately. The index $D_t$ is bound between 0 and 1, where a value of 1 means all the components of the index are expanding. Thus, the lower the value of the index, the more likely the economy is in a recession. Moore's work showed that this index, despite its simple construction, performed well in signaling the approach of a possible turning point with its peak precedes that of the business cycle.[12] When constructing this index, we follow Moore's construction in spirit and depart from his specific procedure where necessary to make our forecasts "out-of-sample."

Two issues we face warrant a discussion. The first arises out of the difficulty associated with seasonal adjustment. Although Moore (1961) made no specific suggestion regarding the appropriate procedure for seasonal adjustment for the purpose of index construction, the author made it very clear that seasonal variations in $c_j$ must be removed before any further analysis is carried out.[13] While most of our variables do not contain a visible seasonal component, some exhibit seasonality in a way that vary significantly during our sample period. This makes reliable seasonal adjustment in recursive out-of-sample forecasting very difficult – we either run into the

[12] In Moore's work, the value of $T_j(\cdot)$ is set, based on visual inspection of the data, to zero if $c_j$ is below its previous peak and one otherwise. Moore also reported the index as a percentage instead of a proportion. Obviously, the peak of $c_j$ can only be identified ex post. Given this, Moore's observation cannot be simply interpreted as the index having strong ex-ante *forecasting* power.

[13] In a list of the desirable characteristics of index components, Moore included "the smaller and more regular the seasonal variations that have to be 'eliminated' before the specific cycles can be studied" (p. 204).

issue of significant residual seasonality or excessive revisions to the seasonally adjusted values. The limited length of our time series, plus our limited toolset (which must not include recently developed methods, e.g., X13) further complicates this task.[14] The second issue concerns data transformation $T_j(\cdot)$, or more specifically, the definition of "expansion." Moore defined expansion as the period between the previous trough and the subsequent peak. We are unable to do so since it is impossible to determine the "subsequent peak" without using data that would not have been available, not to mention that dating the turning points would introduce a significant delay in the sense that a turning point can only be identified well after the fact. Therefore, we instead follow the common practice today and consider "increase" as "expansion" so that our diffusion index shows the proportion of components increasing.

To address the two issues, we need to specify the transformation function $T_j(\cdot)$ for each component $j$ so that, to a reasonable extent, the transformed series $T_j(c_j)$ does not exhibit a significant amount of seasonal variation and is likely to assume positive values during periods of expansion. To avoid unnecessary data distortions and to keep the procedure straightforward (and thus easy to interpret), we use differences and transformations such as percent change from a year ago where appropriate, based on our visual inspection of the data.[15] The transformation that calculates the percent change from a year ago is especially appropriate for our purpose since it not only removes seasonal variations but also helps to smooth the resulting index.[16] The transformation applied to each component is specified in Table 1. Figure 1 shows a few examples of the original data and the transformed data. In the top and the middle plot, the series before transformation showed notable seasonal variations whose patterns changed significantly over time

[14] Falkner (1924) and Spurr (1937) provide descriptions of popular seasonal adjustment procedures during that time. Not all methods are applicable in real time (such as those relying on centered moving averages). Most methods work reliably only if the sample size is large and the seasonal patterns do not vary significantly over time. See Bell and Hillmer (1984) and Wright (2013) for a discussion of some issues with the introduction of noise and distortion through seasonal adjustment. Seasonal adjustment methods were a rich research area in the 1920s and 1930s and were implemented by Mitchell (1927) and Burns and Mitchell (1946) in their business cycle analysis.

[15] Although we inspected data from 1920 to 1940 when making this decision, we would have come to the same conclusions based solely on the data prior to the Great Depression.

[16] A smoother index makes it easier to identify its turning point, thus easier to use to forecast. Zhao (2020) discussed the need for having a smooth diffusion index and explored methods (such as the HP filter) that can be used to achieve this objective in cases where such a need exists. We do not find our indexes volatile enough to require smoothing.

(especially during the 1920s). The bottom plot shows a series without significant seasonal movements in the 1920s but with a clear seasonal component in the 1930s. As shown in Figure 1, in all three cases, our transformations ensure that the transformed data do not exhibit visible seasonal variations and are such that their increases (decreases) largely correspond to periods of business cycle expansion (contraction).

[Figure 1 here]

A standard diffusion index like the one specified in equation (1) has two important limitations. First, it is not targeted to any particular forecast horizon. The index is constructed using the latest data of all of its components. Even if these components all lead the business cycle, the amount of lead time may be uneven and time-varying. Second, by construction, the index assigns equal weights to all its components. This may be undesirable when, for example, the components include measures of different sectors of the economy that are of different sizes. As a result, when the index does not perform satisfactorily, it is difficult to ascertain the cause, as both the selection of the index components and the methodology of the index construction affect its performance. We therefore construct a few additional diffusion indexes and make two additional sets of forecasts using alternative methodologies. We hope these additional results help us better understand the effectiveness of our diffusion indexes and data set.

The additional indexes are constructed to target particular forecast horizons. Instead of simply using the latest observations of all components, these horizon-specific indexes use the data that are best suited to forecast at the chosen horizon whenever possible. The index that is constructed to be an $h$-period-ahead forecast for time $t$ is given as

$$D_{t,h} = \frac{1}{N}\sum_{j=1}^{N} \mathrm{I}\left[T_j\left(c_{j,\min(t-h-1,t-l_j)}\right) > 0\right], \tag{2}$$

where $l_j$ indicates the lead time of component $c_j$. More specifically, before we construct the index by calculating the proportion of its components that are expanding, we identify the lead time of each component ($l_j$) and determine the observation that we should use based on the component's lead time and the forecast horizon $h$. Consider $D_{t+2,2}$ for example. It is our forecast for time period $t+2$ that is made in period $t$ (e.g., $h=2$) with the latest observations of the index components dated $t-1$. Suppose a component $c_j$ of the index has a lead time of $l_j = 4$ periods. Given this lead time, the observation of this component that is best suited for forecasting period $t + 2$ is that of the period $t - 2$. So, we use the value of the component from period $t - 2$ to construct the diffusion index, even though we have the value from period $t - 1$ in our information set. For the components with a lead time of 2 periods or less, the value best suited for forecasting period $t + 2$ is not known at period $t$, so we simply use the latest known value, i.e., the value from period $t - 1$.[17] We subsequently refer to the standard diffusion index given in equation (1) as the diffusion index with an *indefinite* horizon ($h = i$). The additional indexes we just introduced in equation (2) are identified by their specific forecast horizon.

When selecting the variables to use for his index, Moore attempted to identify the lead time of each candidate by visually inspecting the data, pinpointing the specific dates of the variable's peaks and troughs, and comparing these dates with the dates of the business cycle turning points. Since our exercise aims to produce out-of-sample forecasts, we must avoid dating the cycles of each component. Thus, a different strategy is used instead: Before any estimation, we apply the aforementioned transformation to each component to remove seasonal variations and align positive (negative) values with expansions (contractions). Then, for each $j$, we run a set of thirteen probit regressions of recession (a binary variable taking the value of 1 during recessions) on a specific lag of the component $T_j(c_{j,t-r})$ where $r = 0,1,2, \dots, 12$.[18] Since our data set only contains the indicators that we expect to lead the business cycle, we impose as an assumption $l_j > 0 \ \forall \ j$ and do not consider the possibility of them lagging behind the cycle in setting up these regressions. We then evaluate each model's prediction using the ROC analysis and calculate the corresponding AUC score for each $j$ and $r$. The AUC score, let it be denoted by $A_{j,r}$ is a measure of how well a continuous indicator $T_j(c_{j,t-r})$ classifies a binary event, i.e., recession, where higher values mean better forecasts. The lead time of a component $j$ is then estimated to be $l_j = \operatorname{argmax}_{0 \leq r \leq 12} A_{j,r}$. For these estimations, we consider two training samples: Sample (a) goes back to July 1919 and

[17] One could argue that we should drop the index without a sufficiently long lead time in such cases. We experimented with this approach. Given our relatively small data set, dropping the index without a sufficiently long lead time simply leaves too few components at horizons longer than six months, and the resulting index did not perform well even at some shorter horizons (where more components are left).

[18] A lag of the component on the right-hand side of the regression means that the variable leads the business cycle.

sample (b) January 1925.[19] Both samples end in December 1928. The first training sample contains three recessions and the second only one. Table 2 lists the lead time of each variable in our data set estimated using the two training samples.

[Table 2 here]

The horizon specific diffusion indexes given in equation (2) address the limitation of the standard diffusion index not built with a clear forecast horizon in mind. However, the limitation remains that the index assigns equal weights to all its components $T_j(c_j)$. To explore the implications of this limitation, we make two additional sets of forecasts and see if they perform better systematically. The first simply uses a probit model to produce a forecast of recession probability: The dependent variable is the binary recession indicator, and the independent variables are the $T_j\left(c_{j,\min(t-h-1,t-l_j)}\right)$'s, i.e., the same as the components of the diffusion index. The forecasts are thus

$$P_{t,h} = \Phi\left[\sum_{j=1}^{N} \hat{\beta}_j T_j\left(c_{j,\min(t-h-1,t-l_j)}\right)\right], \tag{3}$$

where $\Phi$ is the cumulative distribution function of the standard normal distribution and the $\hat{\beta}_j$s are the coefficient estimates from the probit model. Compared with the equal weights used by the diffusion index, the probit model effectively allows different components to have different weights. The differing weights can be seen as coming from two sources. The first is immediately clear: the coefficients of the variables, i.e., the $\hat{\beta}_j$s, act directly as weights. The second source of differing weights has to do with the data: the levels of the transformed variables $T_j(c_j)$ are used in the regression, while only their signs $\mathrm{I}\left[T_j(c_j) > 0\right]$ are used by the diffusion indexes in equation (1) and (2). Thus, the differing magnitudes of the variables naturally act as weights – for example, a sector in steep decline would be "weighted" more heavily. The second set of forecasts we make in addition to the standard diffusion index is based on the dynamic factor model (DFM) used in Zhao (2020).[20] The data set is again identical as before. From the data, we extract two common factors,

[19] The data for some variables start afterwards (see Table 1). For these variables, the training sample starts when the data become available.

[20] The setup of the model is the same as in Zhao (2020), i.e., the system of equations contains two factors, two shocks, and the two factors follow a first order vector autoregression. Given our focus on the diffusion index in the sense of

which are then used in a probit model to produce a forecast of recession probability. In other words, instead of running the probit model directly on the dozens of variables in the data set over a limited time span, we run the model using only two common factors. Compared with the probit regression using directly the individual variables, this approach offers two benefits. First, it lessens our concern about the degrees of freedom due to the limited length of our time series. Second, by extracting and using only the common factors, we hope to "filter out" potential noise in the data. Of course, the DFM forecasts retain all the benefits of the simple probit model when compared to the standard diffusion index, e.g., the model does not simply assign equal weights to all the variables.

Corresponding to the empirical strategies discussed above, we report three sets of forecasts in the following sections: (A) Proportion expanding, i.e., the standard diffusion index as in equation (1) for $h = i$, and (2) for other horizons; (B) Predicted probabilities from the simple probit model, as in equation (3); and (C) DFM forecasts, i.e., from the factor model. All three vary from 0 to 1. Index (A) assumes low values during recessions, while (B) and (C), being recession probabilities, assume high values during recessions. The parameters of the models used by (B) and (C) are estimated using all available historical data up until the observation used by (A).[21] For each of the three, we make forecasts at the following horizons: $i$, 0, 1, 3, 6, and 9 months, where a horizon of 0 month means a nowcast and a horizon of $i$ means "indefinite" as discussed above.

## 4. FORECAST PERFORMANCE OF THE INDEXES

### 4.1 Can we use the diffusion index to forecast the Depression? If so, how far ahead?

Before we attempt to use the diffusion indexes to forecast the Great Depression, we briefly examine the dynamics of the indexes in the years leading up to the crucial turning point. Since the horizon specific indexes require a training sample for its construction, we focus only on the indexes with an indefinite horizon. Figure 2 presents a comparison of three sets of predictions from the

---

Moore's work instead of the DFM, we leave the full specification of the model in Online Appendix A, which is also available upon request from the authors.

[21] As explained above, for any variable, the observation used by (A) depends on both the forecast horizon and the variable's lead time. The only exception is the diffusion index with an indefinite horizon, which always use the latest observation of each variable.

start of available data (1920) to the end of 1928. The three series depicted are: (A) the standard diffusion index, (B) one minus the predictions from a probit model, and (C) one minus the predictions from a dynamic factor model. The latter two are reported in this way for ease of interpretation – during recessions, all three series, as shown in the figure, should have low values rather than high values. It is worth noting that the values depicted in the figure are in-sample predictions instead of true out-of-sample forecasts.

As shown in Figure 2, the in-sample fit of the probit model that produced predictions (B) was almost perfect, which supports the notion that our data set as a whole contain information that allows the identification of recession periods.[22] Meanwhile, both the diffusion index (A) and the factor model prediction (C) performed well in capturing the uncertain nature of the US economy from 1925 to 1928, which partly reflect concerns about events across the Atlantic. A noteworthy aspect of these two series is their volatility. Taking the diffusion index as an example, we observe a number of sharp increases followed by equally rapid decreases; and the index itself may lead or lag the actual turning points by different amounts in different years: The index failed to identify the onset of the 1920 recession until late in that year, and its rebound lagged the trough of this cycle by about six months. However, the peak and trough of the index aligned very well with the turning points of the 1923 to 1924 recession, despite an arguably false prediction around Feb 1924. Thus, researchers in the 1920s using the index as a forecasting tool for subsequent years would be wise to consider a number of factors jointly when making the decision of whether to interpret a certain value/change in the index as a signal for recession – the level of the index and its direction of change, the persistence of its signals, as well as how smooth the index is or has been in recent time.

[Figure 2 here]

With the above considerations in mind, we turn to the results from our evaluation sample. The first question we hope to answer using our results is whether the business cycle peak in 1929 and the subsequent trough in 1933 are predictable using diffusion indexes. We are interested in the information content of the data available at the time: was there enough information that would have allowed forecasters to predict the turning points? If, given the information contained in the

[22] This, of course, does not simply mean that forecasts constructed using this data set can be perfectly accurate.

data set, the turning points should have been predictable, we would also be interested in knowing how far ahead the prediction could have been made. This requires us to look closely at the dynamics of the various indexes around these two turning points. Since we are mainly concerned with two particular points in time, when comparing the performance of our indexes, we rely on visual inspections. Each plot in Figure 3 is composed in a way similar to Figure 2, comparing the three indexes. The top plot compares the forecasts with an indefinite horizon. The middle and bottom plot compare the forecasts with a horizon of one and three months, respectively. The horizontal axis shows the date of the target, not the forecast – for example, the three-month-ahead forecast corresponding to 1930m1 is made for this month using information available three months ago.[23] In the figure, we also include a reference line at 0.5. Note that this is purely a visual reference. The value 0.5 need not be the threshold one uses to convert the probability forecasts or the index values into a binary recession forecast.

[Figure 3 here]

As shown in Figure 3, the diffusion indexes perform well when it comes to predicting the peak in 1929 – the indexes send strong signals of the upcoming peak months before it can be officially dated: even with modern data and techniques, the NBER's Business Cycle Dating Committee decides on the dates of peaks and troughs many months after the fact. We can draw similar conclusions regarding the other two sets of forecasts. Let us focus our attention for now on the diffusion index: The top plot shows that one can identify the peak of the index (1929m5) no later than 1929m7, after two consecutive declines that brought the index to its new low. This, along with the further declines in subsequent months, sends a strong signal that the business cycle peak is likely very near, although the specific month of the peak would remain unknown, as this index has an indefinite horizon.[24] As shown in the middle plot, the one-month-ahead prediction is past its peak once the value for 1929m11 is known in 1929m10. Similarly, we can see in the bottom plot that the 1929m11 value is clearly below the previous peak – and this value is known three months ahead in 1929m8. When it comes to the prediction of the trough in 1933, the diffusion

[23] The forecasts with an indefinite horizon is evaluated like a nowcast (i.e., $h = 0$). The same applies to all other figures.

[24] Recall that the index for indefinite horizon is constructed using the latest observation of all its components. The procedure does not rely on us determining the lead time of each component, and thus is not specific to either training sample (a) or (b).

index with an indefinite horizon starts to trend up in 1933m5, while the one- and three-month ahead index start to trend up no later than 1933m6 and 1933m7, respectively. Here, the diffusion indexes again send a strong signal of the trough almost simultaneously as the actual turning point.

Although the diffusion index does not utilize information on the magnitude of its components nor does it attempt to assign different weights to them, the performance of the index remains robust. This can be observed by comparing the factor model forecasts and the diffusion index in Figure 3. Both series of forecasts have almost identical peaks and troughs, so that neither has a clear advantage in terms of how far ahead the turning point can be identified. This implies that, for the purpose of predicting the two business cycle turning points, there is little information loss from using a standard/naïve diffusion index with equal weights relative to using an index that weights its components. Moreover, unlike the diffusion indexes, the one- and three-month-ahead forecasts of the factor model showed a significant increase around late 1931, sending a misleading signal of a recovery while the trough was more than a year away. In addition, we note that the predicted probabilities from the simple probit model do not as perform well. Partly because of the length of our time series, the model tends to fit extremely well *in sample*, thus giving undue weights to a few variables that may turn out to be poor predictors out-of-sample. The resulting forecasts thus tend to be as volatile as they are extreme, e.g., they can change from 0 to 1 then back to 0 in a three-month period as shown in the top plot. Most importantly, the probit model forecasts do not allow us to identify the turning points any earlier than the other forecasts do. Therefore, in subsequent analyses, we mainly focus on the diffusion index.

In our observations above regarding Figure 3, our timing of the turning points in the forecasts is conservative. Depending on how uncertain forecasters feel, they may be inclined towards sounding the alarm for the incoming peak earlier than we choose to. To illustrate this point, we compare the diffusion indexes for forecasting three-, six-, and nine-month ahead in Figure 4. At all three horizons, the forecasts peak well before 1930m1. This means that forecasters who rely on these longer horizon indexes could identify signals of the business cycle peak as early as mid-1929, six to nine months ahead of time.

[Figure 4 here]

To further explore the implications of constructing a diffusion index with a specific forecast horizon versus an indefinite forecast horizon, we make a comparison as presented in Figure 5. Specifically, we plot, side by side over the entire forecast period from 1920 to 1934, the diffusion index with an indefinite horizon (i.e., always using the latest observation of its components) and the index with a horizon of zero. As Figure 5 shows, while the index with an indefinite horizon clearly leads the business cycle more than the index with a horizon of zero does, the amount of lead time varies. The index with a horizon of zero, i.e., for nowcasts, is better aligned with the recessions, especially after 1926. This observation should serve to reinforce our earlier conclusion that the business cycle peak and trough around the Depression are identifiable using a diffusion index.

[Figure 5 here]

### 4.2 The usefulness of additional data

So far, we have exclusively examined the indexes constructed using Moore's 21 variables. Next, we turn to the question of whether the additional monthly and quarterly variables we selected provide further improvement in the indexes' forecast performance. Since we are still primarily interested in the prediction of the turning points, we continue to rely on visual inspections of the forecasts.

[Figure 6 here]

In Figure 5, we compare the diffusion indexes with an indefinite horizon constructed using three progressively larger data sets. The first is based on the 21 series used by Moore. The second is based on the first index with additional monthly indicators of our choosing. The last data set adds to the second index some indicators available at a quarterly frequency. Our indexes remain monthly in all three cases. The quarterly series are converted to a monthly frequency using linear interpolation. Figure 6 shows that the additional data we introduced helps with the forecasting of turning points in subtle, yet important ways. First, with the additional data, the indexes decline more smoothly throughout the initial months of the recession, potentially eliminating doubts regarding the state of the business cycle that one might have had in 1929m10 if one were to rely solely on the 21 series. Second, as the economy moves out of the recession, the indexes made with additional data, especially the quarterly series, do not have the same sharp decline in 1933m3 and 1933m4.

### 4.3 The effect of using different training samples to determine the components' lead time

Recall that when constructing diffusion indexes with specific horizons, we relied on a training sample to determine the lead time of each of the components. We explored two different training samples, where training sample (a) covers the three recessions before 1929 and training sample (b) covers only the recession immediately preceding that of 1929. Figure 7 compares the indexes constructed based on these training samples. Both indexes use all the data we have available, including the quarterly series. The indexes for nowcasting are in the top plot and the ones with a horizon of three months are in the bottom plot.

[Figure 7 here]

Figure 7 shows that a shorter training sample, i.e., training sample (b), helps to identify the peak in 1929 more accurately in the case of the nowcasts – the index constructed using this training sample shows an earlier and steadier decline around the onset of the recession in the second half of 1929. In addition, the indexes based on a shorter training sample have a more rapid decline during the first year of the recession and a slightly more pronounced rebound immediately out of the recession in early 1933. However, the differences between the indexes constructed using the two training samples become almost negligible for the three-month-ahead forecasts (and those at longer horizons, which are omitted from the figure). That said, although we believe a shorter training sample helps here, we caution against generalizing the conclusion beyond the context of forecasting the 1929 peak.

### 4.4 Overall classification accuracy of the indexes

Finally, we turn to the overall classification accuracy of the indexes in predicting the binary state of the economy, i.e., recession vs. non-recession. This is measured using the AUC statistic. Specifically, for each forecast series, we construct a nonparametric ROC curve and calculate the area under the curve. For the forecasts made using data sets augmented with the variables we selected, we take their AUC statistic and test the null hypothesis that it is the same as the AUC statistic of the forecasts made using Moore's series alone. In other words, we test whether the additional variables we selected help to improve forecast accuracy. For this purpose, we choose the nonparametric test of correlated ROC curves developed by DeLong *et al*. (1988) since all the forecasts are targeting the same set of outcomes. The results are reported in Table 3 using an evaluation sample of 1929 to 1934.[25]

[Table 3 here]

Most of our previous observations are confirmed by these results. As the horizon increases, the accuracy of all forecasts deteriorates. Considering its simplicity, the diffusion indexes perform very well, especially at short horizons, where the AUC statistics reach above 0.9. Augmenting Moore's data set with the additional variables, especially the quarterly variables, helps to further improve the indexes' forecast performance at nearly all horizons. For the purpose of nowcasting, using the standard diffusion index with an indefinite horizon provides better overall accuracy than using the horizon-specific diffusion index. For the purpose of forecasting the 1929-1933 recession using horizon specific indexes, the shorter training sample, i.e., training sample (b) prove to be more helpful.

It is worth noting that the AUC statistics do not reflect the forecast's ability to identify the turning points. Since the two turning points represent only 2 out of 60 observations in the evaluation sample, forecasts that are better used for predicting the turning points do not necessarily exhibit better overall accuracy. For example, the forecasts from the probit model (B) based on Moore's data have consistently higher AUC across all horizons than those of the diffusion index, despite the latter being the one better suited for predicting turning points as shown in our analyses above.

## 5. REAL-TIME DATA VINTAGES

We carried out our forecasting exercises discussed above by making out-of-sample forecasts using revised data. In this section, we repeat this process using appropriate real-time data vintages. Note that despite the use of real-time data vintages, the forecasts we report in this section are not true

[25] The test of DeLong *et al*. (1988) is not designed specifically for time series data with a high level of serial dependence. The results should thus be interpreted cautiously, especially in cases where it suggests a statistically significant difference between two ROC curves. The authors are unaware of mature tests of equality of correlated ROC curves that are better choices given our context and application. In the discussions below, we primarily focus on the AUC statistics themselves, which remain valid.

real-time forecasts. A few difficulties make it impractical for us to carry out a true real-time forecasting exercise. The first concerns the publication lags of the variables in our data set. We are unable to ascertain the lag structure for all of our variables. This issue becomes more complex if we were to attempt to account for possible lags in data transmission in this historical context. The second difficulty stems from our limited access to real-time data vintages. As discussed in Section 2, we have obtained real-time data for a number of variables that we believe are important "drivers" of our results. Many variables were not revised, but in some cases, we simply did not have access to real-time data for all our variables. In addition, to the best of our knowledge, real-time data for the NBER recession dates do not exist for our target period.[26] Finally, a few variables in our data set are themselves (diffusion) indexes that are constructed after the 1929-1933 recession. As long as we use these series, we cannot avoid using the hindsight of those who developed the indexes. Therefore, we do not attempt to create forecasts in real time. Instead, we simply use the real-time data available to us as a robustness check to see if the substitution of our real-time data vintages for historical data would alter our baseline results in any fundamental way.

[Figure 8 here]

In Figure 8, we repeat some of the earlier exercises using forecasts constructed with appropriate real-time data vintages. The top plot repeats the same exercise shown in the top plot of Figure 3. The middle plot repeats the same exercise shown in Figure 4. The bottom plot repeats the same exercise shown in Figure 5. Generally speaking, the diffusion indexes that are constructed using the sign of the components are little changed. Since the transformations applied to most variables rely on values from a year ago in order to mitigate seasonality, minor data revisions do not alter the sign of the transformed values. Of course, the changes in the magnitude of these transformed values lead to slight differences in the forecasts made using the probit model and the factor model. However, we do not observe anything that contradicts the conclusions reported earlier. These results based on real-time data vintages only serve to reinforce our main conclusions from the previous section.

[26] Burns and Mitchell provide some business cycle indicators for the NBER in the 1920s and 1930s, but the recession dates were made after the fact and were not revised in this period (Romer and Romer 2020).

## 6. QUARTERLY FORECASTS USING DIFFUSION INDEXES AND MIXED FREQUENCY DATA

In this section, we briefly consider whether we can better forecast the depression using the standard diffusion indexes if we focus on forecasting at a lower frequency, i.e., quarterly. The primary motivations for this choice are (1) the index at a lower frequency may be smoother; and (2) forecasting at a lower frequency allows us to utilize data series that are unavailable at the higher, monthly frequency. As discussed in Zhao (2020), the smoothness of the index is important when it comes to extracting binary predictions, as a smoother series may allow us to better identify turning points and thus reduce the uncertainty of the forecasts. Although the index is to be constructed at the quarterly frequency, we do not want to throw away information contained in the higher frequency monthly variables. We thus construct a mixed frequency data set by allowing each monthly variable to enter the data set three times, with the values for the three months in a quarter. This strategy also increases the number of candidate series for the diffusion index, which helps to further smooth the index. In Figure 9, we plot the quarterly diffusion indexes. The top plot compares the three indexes with indefinite horizon that are constructed using the three data sets. The middle plot shows the comparison for the indexes constructed for nowcast (i.e., $h = 0$). The bottom plot shows the comparison for $h = 1$. Note that for these quarterly forecasts, the horizons are stated in quarters, and given the reduced number of observations, training sample (b) is not applicable.

[Figure 9 here]

As expected, Figure 9 shows that the quarterly indexes are smoother, which makes it easy to identify the turning points of the indexes. For example, the top plot shows that the diffusion index with an indefinite horizon clearly started to decline in 1929q4. If we also consider the indexes with specific horizons shown in the lower two plots, this decline can be identified during this quarter, though not one quarter ahead of time. The quarterly indexes appear to be able to "forecast" the 1933 trough only since 1933q3, which is one quarter after the actual trough. Comparing the indexes made with different data sets, we observe some differences between the Moore's series and the augmented data sets, but not between the two augmented data sets. This is understandable since the number of additional monthly variables exceeds that of the additional quarterly variables, not to mention the design that creates three quarterly series from each monthly series. The added data provide some benefits especially during the first year of the recession. However, they appear to do little to help with the forecasting of the 1933 trough. While temporal aggregation could smooth out some of the noise in higher-frequency data, we do not find that this effect is large enough to offset the resulting loss in information and in the timeliness of the forecasts.

## 7. CONCLUDING REMARKS

This paper studied whether diffusion indexes provide a good forecast of the US Great Depression, as measured by the nowcasts and forecasts of turning points and the forecast ability of diffusion indexes for the Great Contraction period of 1929-1933. We have found that Moore's original diffusion index has decent performance, and that models based on the addition of other series performs at least as well. We are able to forecast the initial turning point when recession begins in 1929 in advance. This forecast is accurate several months ahead of the iconic stock market crash of October 1929. We also have some success in forecasting other turning points in advance, though the performance degrades in the late 1930s. This contrasts with previous forecasters that viewed the Depression as unforecastable, even using modern techniques and with the benefit of hindsight. Our results demonstrate the value of diffusion indexes in forecasting, specifically in an important historical episode. This suggests that diffusion indexes should be used more widely to provide useful forecasts, especially historically.

## REFERENCES


Barger, Harold. "Outlay and Income in the United States, 1921–1938." *NBER Books* (1942).

Bell, William R., and Steven C. Hillmer. "Issues involved with the seasonal adjustment of economic time series." *Journal of Business & Economic Statistics* 2, no. 4 (1984): 291-320.

Broida, Arthur L. "Diffusion indexes." *The American Statistician* 9, no. 3 (1955): 7-16.

Burns, Arthur F., and Wesley C. Mitchell. *Measuring business cycles*. No. burn46-1. National bureau of economic research, 1946.

Cortes, Gustavo S., Bryan Taylor, and Marc D. Weidenmier. "Financial factors and the propagation of the Great Depression." *Journal of Financial Economics* 145, no. 2 (2022): 577-594.

DeLong, E. R., D. M. DeLong, and D. L. Clarke-Pearson. "Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach." *Biometrics*, 44 (1988): 837–845.

Dominguez, Kathryn M., Ray C. Fair, and Matthew D. Shapiro. "Forecasting the depression: Harvard versus Yale." *The American Economic Review* (1988): 595-612.

Eichengreen, Barry J. *Golden fetters: the gold standard and the Great Depression, 1919-1939*. NBER series on long-term factors in economic development, 1996.

Falkner, Helen D. “The Measurement of Seasonal Variation.” *Journal of the American Statistical Association* 19, no. 146

Friedman, Milton, and Anna J. Schwartz. "A Monetary History of the United States, 1867–1960." *NBER Books* (1963).

Giannone D, Reichlin L, Small D. "Nowcasting: the real-time informational content of macroeconomic data." *Journal of Monetary Economics* 55 (2008): 665–676.

Getz, Patricia M., and Mark G. Ulmer. "Diffusion indexes: a barometer of the economy." *Monthly Labor Review* 113 (1990): 13.

Goldfarb, Robert S., Herman O. Stekler, and Joel David. "Methodological issues in forecasting: Insights from the egregious business forecast errors of late 1930." *Journal of Economic Methodology* 12, no. 4 (2005): 517-542.

Haney, L.H. , 1931. Business Forecasting. Ginn and Company.

Hamilton, James D. "Was the deflation during the Great Depression anticipated? Evidence from the commodity futures market." *The American Economic Review* (1992): 157-178.

Kindleberger, Charles Poor. The world in depression, 1929-1939. Univ of California Press, 1986.

Klug, Adam, John S. Landon-Lane, and Eugene N. White. "How could everyone have been so wrong? Forecasting the Great Depression with the railroads." *Explorations in Economic History* 42, no. 1 (2005): 27-55.

Lahiri, K., Monokroussos, G., and Zhao, Y. "Forecasting Consumption: the Role of Consumer Confidence in Real Time with many Predictors." *Journal of Applied Econometrics* 31 (2016): 1254– 1275.

Lenel, Laetitia. "Mapping the Future." *Jahrbuch für Wirtschaftsgeschichte/Economic History Yearbook* 59, no. 2 (2018): 377-413.

Lenel, Laetitia. "Searching for a Tide Table for Business: Interwar Conceptions of Statistical Inference in Business Forecasting." *History of Political Economy* 53, no. S1 (2021): 139-174.

Lepoutre, TJ. “Diffusion Indexes of State and Metropolitan Area Employment Changes.” *Monthly Labor Review*, 2022, 1–13.

Mathy, Gabriel, and Herman Stekler. "Expectations and forecasting during the Great Depression: Real-time evidence from the business press." *Journal of Macroeconomics* 53 (2017): 1-15.

Mathy, Gabriel, and Herman Stekler. "Was the deflation of the depression anticipated? An inference using real-time data." *Journal of Economic Methodology* 25, 2 (2018): 117-125.

Mitchell, Wesley Clair. "Business cycles: The problem and its setting." *NBER Books* (1927).

Mitchell, J., S. Solomou, and M. Weale (2012). Monthly GDP estimates for inter-war Britain. Explorations in Economic History 49, 543-556.

Moore, Geoffrey Hoyt. *Business cycle indicators*. Vol. 1. Princeton, NJ: Princeton University Press, 1961.

Moore, Geoffrey Hoyt. "Analyzing business cycles." *The American Statistician* 8, no. 2 (1954): 13-19.

Rhodes, E.C., 1937. The construction of an index of business activity. Journal of the Royal Statistical Society 100, 18–39.

Romer, Christina D. "The nation in depression." *Journal of Economic Perspectives* 7, no. 2 (1993): 19-39.

Romer, C. D., & Romer, D. H. (2020). “NBER business cycle dating: Retrospect and prospect.” *NBER and the Evolution of Economic Research, 1920–2020.* The ASSA Annual Meeting, San Diego, California, January 2020.

Spurr, William A. “A Graphic Method of Measuring Seasonal Variation.” *Journal of the American Statistical Association* 32, no. 198 (1937): 281–89.

Stekler, Herman O. "An analysis of turning point forecasts." *The American Economic Review* 62, no. 4 (1972): 724-729.

Stock, James H., and Mark W. Watson. "Diffusion indexes." *NBER Working Paper 6702* (1998): 1-67.

Stock, James H., and Mark W. Watson. "Macroeconomic forecasting using diffusion indexes." *Journal of Business & Economic Statistics* 20, no. 2 (2002): 147-162.

Stock, James H., and Mark Watson. "Dynamic factor models." *Oxford Handbooks Online* (2011).

Temin, Peter. *Did monetary forces cause the Great Depression?* Norton, 1976.

Temin, Peter, and Barrie A. Wigmore. "The end of one big deflation." *Explorations in Economic history* 27, no. 4 (1990): 483-502.

Wright, Jonathan H. "Unseasonal seasonals?" *Brookings Papers on Economic Activity* 2013, no. 2 (2013): 65-126.

Zhao, Yongchen. "Predicting US Business Cycle Turning Points Using Real-Time Diffusion Indexes Based on a Large Data Set." *Journal of Business Cycle Research* 16, no. 2 (2020): 77-97.

## Table 1. List of variables and data availability

This table lists the variables in our data set, whether they are used by Moore, and the amount of data we have on each variable. A star * after a variable name indicates that we have real-time data for the variable. Superscripts after a variable name indicates the transformation. "a" means the first difference; "b" means the variable is transformed by subtracting 50 from its levels. The absence of any superscript after a variable means it is transformed to the percent change from a year ago.

| Description | Data Start | Data End | Number of Observations |
|---|---|---|---|
| **Variables used by Moore** | | | |
| Automobile Production, Passenger Cars, Factory Production | 1920m1 | 1940m12 | 252 |
| Automobile Production, Trucks | 1920m1 | 1940m12 | 252 |
| Automobile Tire Inner Tube Production* | 1922m1 | 1940m12 | 228 |
| Average Hours of Work Per Week, Manufacturing Industries | 1921m6 | 1940m12 | 235 |
| Bank Clearings Outside New York City | 1920m1 | 1940m12 | 252 |
| Contracts for Industrial Buildings (sq. ft.) | 1920m1 | 1940m12 | 252 |
| Dow-Jones Industrial Stock Price Index | 1920m1 | 1940m12 | 252 |
| Index of Department Store Sales | 1920m1 | 1940m12 | 252 |
| Index of Factory Employment, Machinery | 1920m1 | 1940m12 | 252 |
| Index of General Business Activity (AT&T) | 1920m1 | 1937m12 | 216 |
| Index of Industrial Production and Trade | 1920m1 | 1940m12 | 252 |
| Index of Value of Residential Construction Contracts | 1920m1 | 1940m12 | 252 |
| Index of Wholesale Prices | 1920m1 | 1940m12 | 252 |
| Industrial Production Index* | 1920m1 | 1940m12 | 252 |
| Liabilities of Business Failures | 1920m1 | 1933m1 | 157 |
| Paper Production, All Grades | 1920m1 | 1940m12 | 252 |
| Pig Iron Production* | 1920m1 | 1940m12 | 252 |
| Production Worker Employment, Manufacturing | 1920m1 | 1940m12 | 252 |
| Railroad Operating Income | 1920m1 | 1940m12 | 252 |
| Revenue and Non-Revenue Net Ton-Miles of Freight Hauled | 1920m1 | 1940m12 | 252 |
| Steel Ingot Production* | 1920m1 | 1939m12 | 240 |
| **Additional monthly variables** | | | |
| Bank Debits for New York | 1920m1 | 1940m12 | 252 |
| Bank Debits, 140 Centers Outside of New York City | 1920m1 | 1940m12 | 252 |
| Bank Rates on Customer Loans, Leading Cities | 1920m1 | 1939m2 | 230 |
| Cum. Net Diff Ind, Industrial Production[a] | 1919m3 | 1940m12 | 262 |
| Diff Ind of Eight Leading Indicators[b] | 1919m4 | 1940m12 | 261 |
| Index of American Business Activity, Cleveland Trust | 1920m1 | 1940m12 | 252 |
| Index of Factory Employment, Automobiles | 1920m1 | 1940m12 | 252 |
| Labor Turnover, Gross Accession Rate, Manufacturing | 1920m1 | 1930m12 | 132 |
| Labor Turnover, Layoff Rate, Manufacturing | 1920m1 | 1930m12 | 127 |
| Productive security issues, corporate | 1922m1 | 1940m7 | 217 |
| Productive security issues, municipal, state, etc. | 1922m1 | 1940m7 | 223 |
| Retail Index US | 1920m1 | 1939m12 | 240 |
| Steel Ingot Production Capacity Utilization* | 1927m1 | 1940m12 | 168 |
| WPI, Commodities ex Farm Products and Foods | 1920m1 | 1940m12 | 252 |
| **Additional quarterly variables** | | | |
| Cum. Net Diff Ind, Profits of 17-700 Manuf. Corps. [b] | 1920m7 | 1940m10 | 244 |
| Diff Ind, Freight Car loadings, 19 Commodity Groups[a] | 1927m1 | 1939m10 | 154 |
| Gross National Product | 1922m1 | 1940m10 | 226 |
| Manufacturing Inventory Investment[b] | 1921m4 | 1938m10 | 211 |
| New Capital, Equipment[a] | 1919m4 | 1940m10 | 259 |
| New Capital, Plant[a] | 1919m4 | 1940m10 | 259 |
| Railroad Freight Tons Originated, Carload | 1921m1 | 1940m10 | 238 |

## Table 2. Estimated lead time of our variables

The table shows the estimated amount of lead time, in months, of each variable in our data set. Two sets of estimates are provided as the estimation is carried out using two subsamples. Subsample (a), which starts in July 1919, contains the three recessions before that of 1929 to 1933, while subsample (b), which starts in Jan 1925, covers only the recession immediately before it. Only subsample (a) applies to quarterly variables.

| Variable | Subsample (a) | Subsample (b) | Variable | Subsample (a) | Subsample (b) |
|---|---|---|---|---|---|
| Automobile Production, Passenger Cars, Factory Production | 12 | 0 | Index of Value of Res. Cons. Contr. | 3 | 8 |
| Automobile Production, Trucks | 12 | 0 | Index of Wholesale Prices | 12 | 3 |
| Automobile Tire Inner Tubes Production | 1 | 0 | Industrial Production Index | 12 | 0 |
| Average Hours of Work Per Week, Manufacturing Industries | 0 | 1 | Labor Turnover, Gross Accession Rate, Manufacturing | 0 | 0 |
| Bank Clearings Outside New York City | 12 | 0 | Labor Turnover, Layoff Rate, Manufacturing | 0 | 4 |
| Bank Debits for New York | 1 | 0 | Liabilities of Business Failures | 0 | 0 |
| Bank Debits, 140 Centers Outside of New York City | 12 | 0 | Manufacturing Inventory Investment | 12 | |
| Bank Rates on Customer Loans, Leading Cities | 3 | 5 | New Capital, Equipment | 0 | |
| Contracts for Industrial Buildings | 12 | 5 | New Capital, Plant | 0 | |
| Cum. Net Diff Ind, Industrial Production | 0 | 8 | Paper Production, All Grades | 12 | 10 |
| Cum. Net Diff Ind, Profits of 17-700 Manuf. Corps. | 0 | | Pig Iron Production, 1000s Gross Tons, M, NSA | 12 | 0 |
| Diff Ind of Eight Leading Indicators | 3 | 4 | Production Worker Employment, Manufacturing | 12 | 9 |
| Diff Ind, Freight Car loadings, 19 Commodity Groups* | 0 | | Productive security issues, corporate | 3 | 9 |
| Dow-Jones Industrial Stock Price Index | 0 | 12 | Productive security issues, municipal, state, etc. | 10 | 0 |
| Gross National Product | 12 | | Railroad Freight Tons Originated, Carload | 9 | |
| Index of American Business Activity | 12 | 12 | Railroad Operating Income | 0 | 0 |
| Index of Department Store Sales | 7 | 0 | Retail Index US | 9 | 1 |
| Index of Factory Employment, Automobiles | 12 | 0 | Revenue and Non-Revenue Net Ton-Miles of Freight Hauled | 8 | 1 |
| Index of Factory Employment, Machinery | 11 | 1 | Steel Ingot Production | 12 | 8 |
| Index of General Business Activity | 12 | 12 | WPI, Commodities ex Farm Products and Foods | 12 | 2 |
| Index of Industrial Production and Trade | 12 | 0 | | | |

* The length of the time series is too short for estimation. We assume the variable is coincident.

## Table 3. Forecasting performance of various indexes measured by AUC statistics – 1929 to 1933

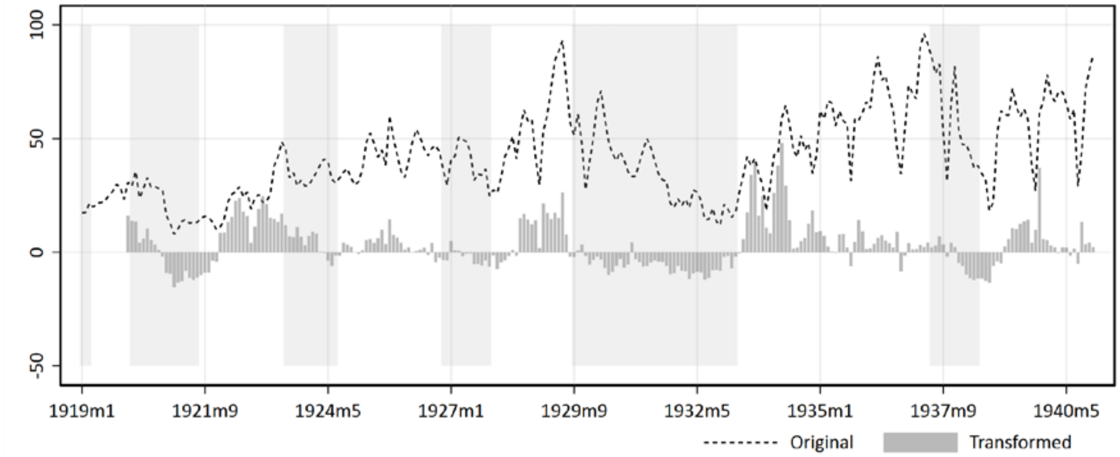
This table shows the area under the nonparametric ROC curve (AUC statistic) for each set of forecasts. The evaluation sample is from Jan 1929 to Dec 1933. Each column corresponds to a horizon. Horizon "i" means "indefinite" – these forecasts are evaluated like nowcasts and they do not have a corresponding training sample as the training sample refers to the one used to determine the lead time of the index components. For the forecasts made using data sets with the monthly and quarterly variables we selected, a test of the null hypothesis that the AUC is the same as that of the forecasts made using Moore's series alone is carried out. When this null hypothesis is rejected at the 10% level, the AUC statistic is set in bold.

| Index and Data Set | | Training sample (a) | | | | | Training sample (b) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $h = i$ | $h = 0$ | $h = 1$ | $h = 3$ | $h = 6$ | $h = 9$ | $h = 0$ | $h = 1$ | $h = 3$ | $h = 6$ | $h = 9$ |
| **A. % Expanding** | | | | | | | | | | | |
| 1. Moore series | 0.95 | 0.69 | 0.66 | 0.68 | 0.62 | 0.58 | 0.92 | 0.88 | 0.84 | 0.71 | 0.63 |
| 2. Additional monthly variables | 0.96 | **0.73** | **0.70** | **0.68** | **0.67** | 0.58 | 0.92 | 0.88 | 0.84 | 0.74 | 0.65 |
| 3. Additional quarterly variables | **0.98** | **0.79** | **0.78** | **0.75** | **0.72** | **0.62** | **0.95** | **0.93** | **0.88** | **0.79** | 0.65 |
| **B. Predicted Probability** | | | | | | | | | | | |
| 1. Moore series | 0.93 | 0.92 | 0.92 | 0.85 | 0.70 | 0.62 | 0.82 | 0.79 | 0.82 | 0.74 | 0.59 |
| 2. Additional monthly variables | **0.84** | 0.82 | 0.83 | **0.67** | **0.66** | 0.69 | **0.85** | **0.73** | **0.74** | 0.59 | 0.61 |
| 3. Additional quarterly variables | 0.73 | 0.78 | 0.75 | **0.67** | 0.51 | 0.68 | 0.82 | 0.67 | 0.80 | 0.62 | 0.59 |
| **C. DFM Forecast** | | | | | | | | | | | |
| 1. Moore series | 0.98 | 0.99 | 0.97 | 0.94 | 0.81 | 0.61 | 0.93 | 0.95 | 0.92 | 0.90 | 0.68 |
| 2. Additional monthly variables | **0.99** | **0.99** | **0.98** | **0.95** | 0.74 | 0.61 | 0.97 | **0.95** | **0.92** | **0.92** | **0.69** |
| 3. Additional quarterly variables | **0.99** | **1.00** | **0.99** | **0.98** | **0.89** | 0.67 | **0.97** | **0.95** | **0.93** | **0.94** | **0.77** |

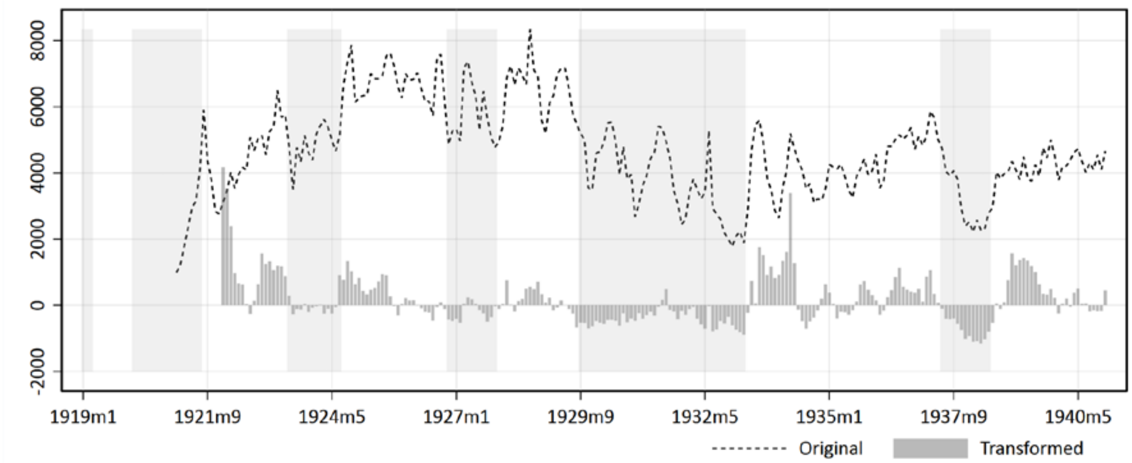
## Figure 1. Selected indicators and their transformed values

The three plots below present the raw data (dashed line) and the transformed values (bars) for three variables related to automobile production. They show that the transformed data do not exhibit visible seasonal variations and are such that positive values are typically associated with business cycle expansions. Shaded areas correspond to NBER dated recession periods.

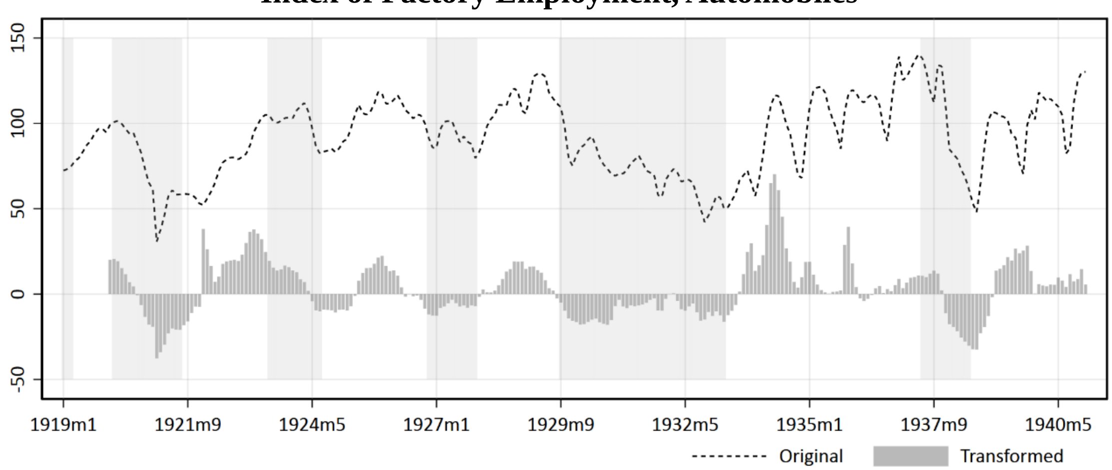
**Automobile Production, Trucks, 1000s**




**Automobile Tire Inner Tubes, 1000s**

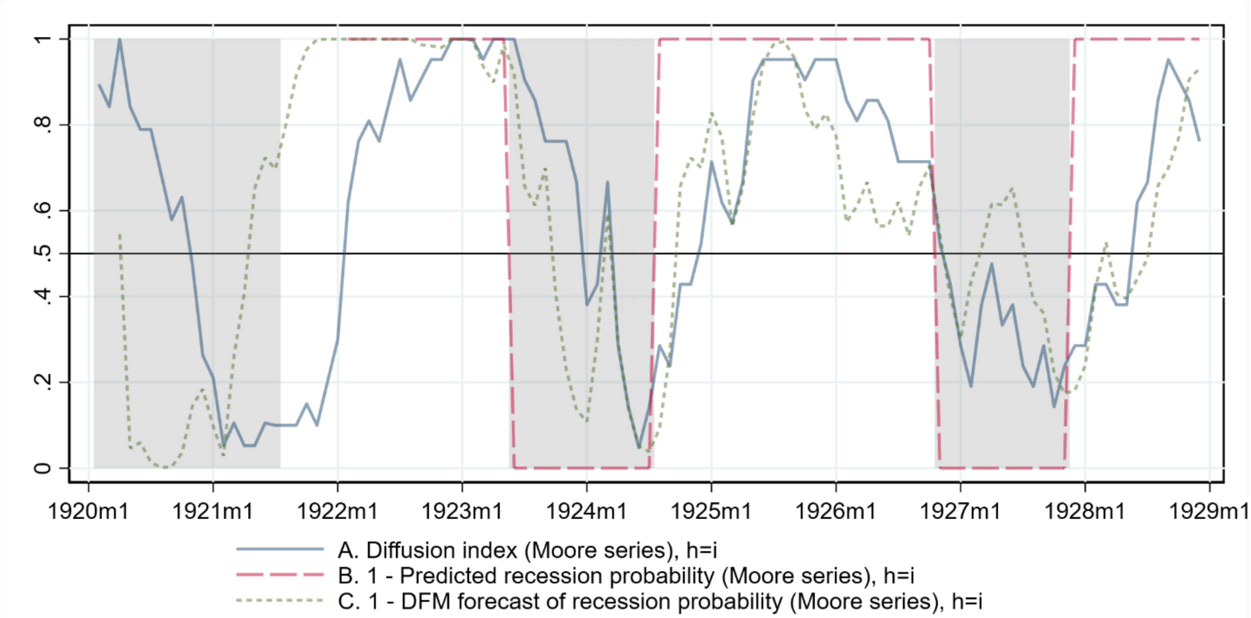



**Index of Factory Employment, Automobiles**



## Figure 2. Indexes with indefinite horizon, 1920 to 1928

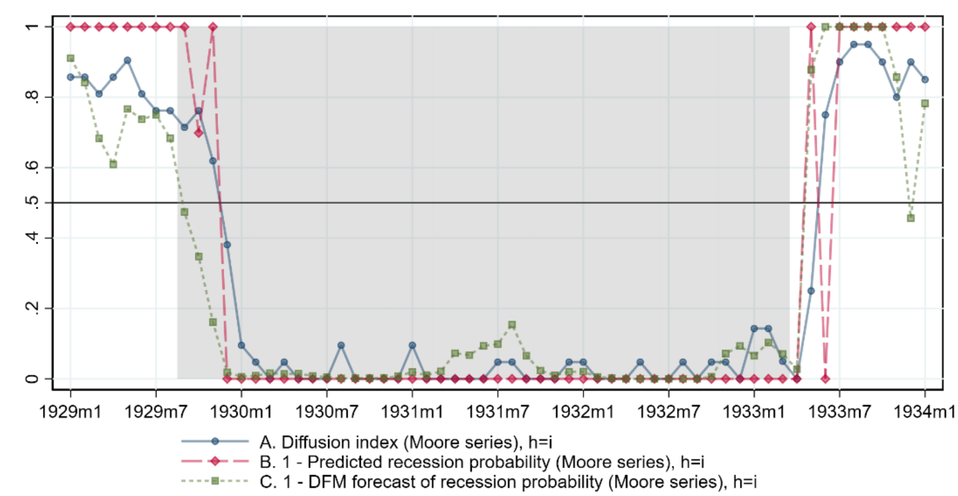
For the years 1920 to 1928, i.e., before our evaluation sample starts, this figure presents a comparison of three sets of predictions, the standard diffusion index, the predictions from a probit model, and the predictions from a dynamic factor model. Where applicable, the predictions are made in sample. The black horizontal grid line corresponds to the value 0.5. Note that, for ease of interpretation, for indexes B and C, we report one minus the predicted recession probability so that low values in the figure correspond to recession periods. Shaded areas correspond to NBER dated recession periods.

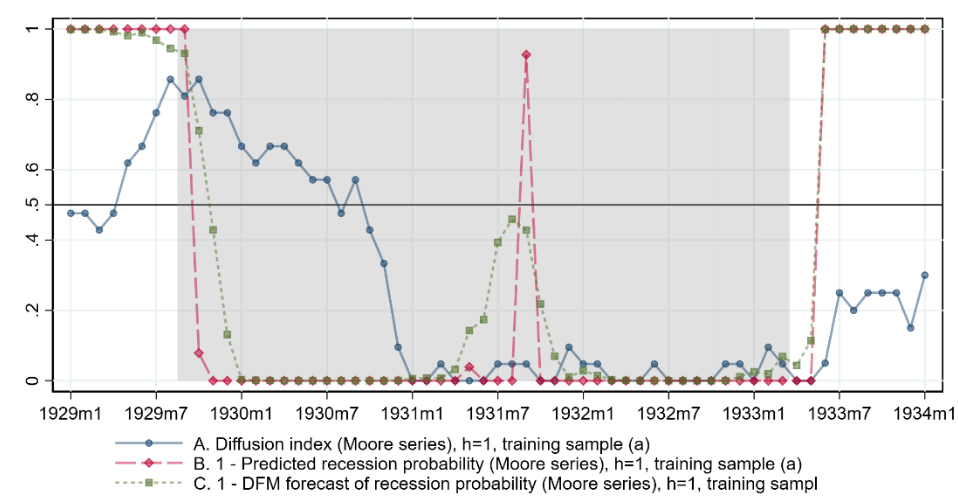


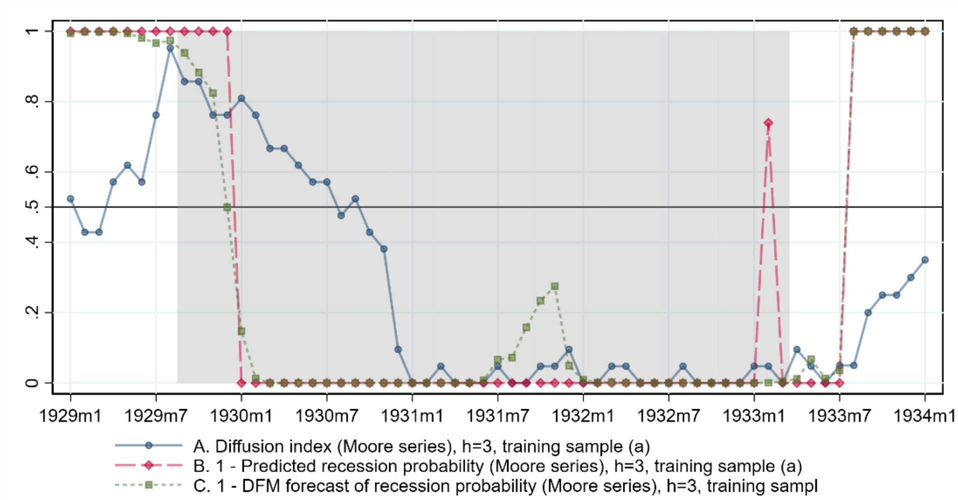

Note: Due to limited data availability of the auto tire inner tube production variable, index B is missing for the initial months.

## Figure 3. Comparing the indexes around the business cycle turning points of 1929 and 1933

Each plot presents a comparison of three sets of predictions, the standard diffusion index, the predictions from a probit model, and the predictions from a dynamic factor model, around the business cycle turning points of 1929 and 1933. The top plot compares the forecasts with an indefinite horizon. The middle plot compares the forecasts with a horizon of one month. The bottom plot compares the forecasts with a horizon of three months. The black horizontal grid line corresponds to the value 0.5. Note that, for ease of interpretation, for indexes B and C, we report one minus the predicted recession probability so that low values in the figure correspond to recession periods. Shaded areas correspond to NBER dated recession periods.

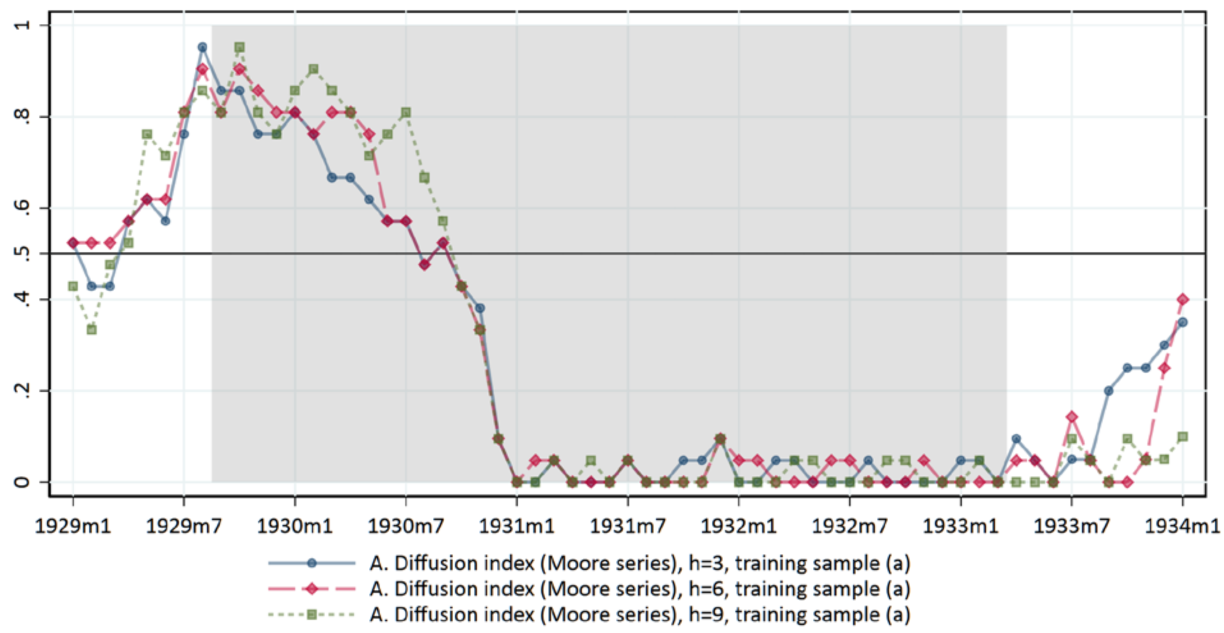


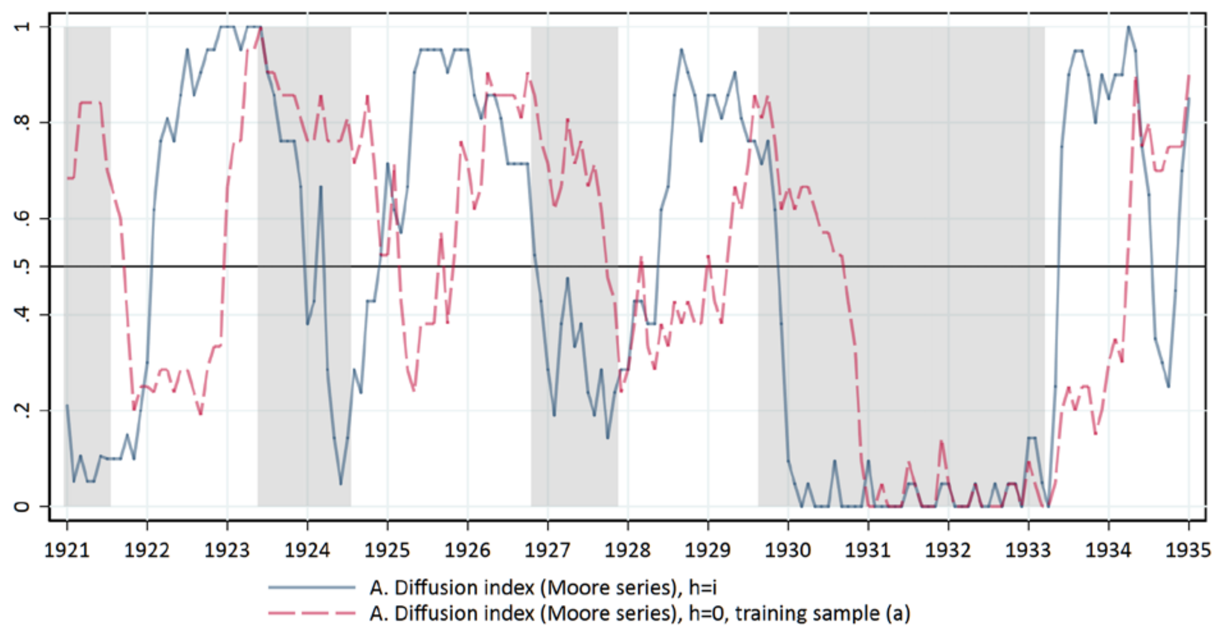
## Figure 4. Comparing the diffusion index at horizons three, six, and nine months

This figure compares the diffusion indexes with horizons of three, six, and nine months. Note that the horizontal axis shows the date of the target, not the forecast – for example, the three-month-ahead forecast corresponding to 1930m1 is made for this month using information available three months ago. Shaded areas correspond to NBER recession periods.



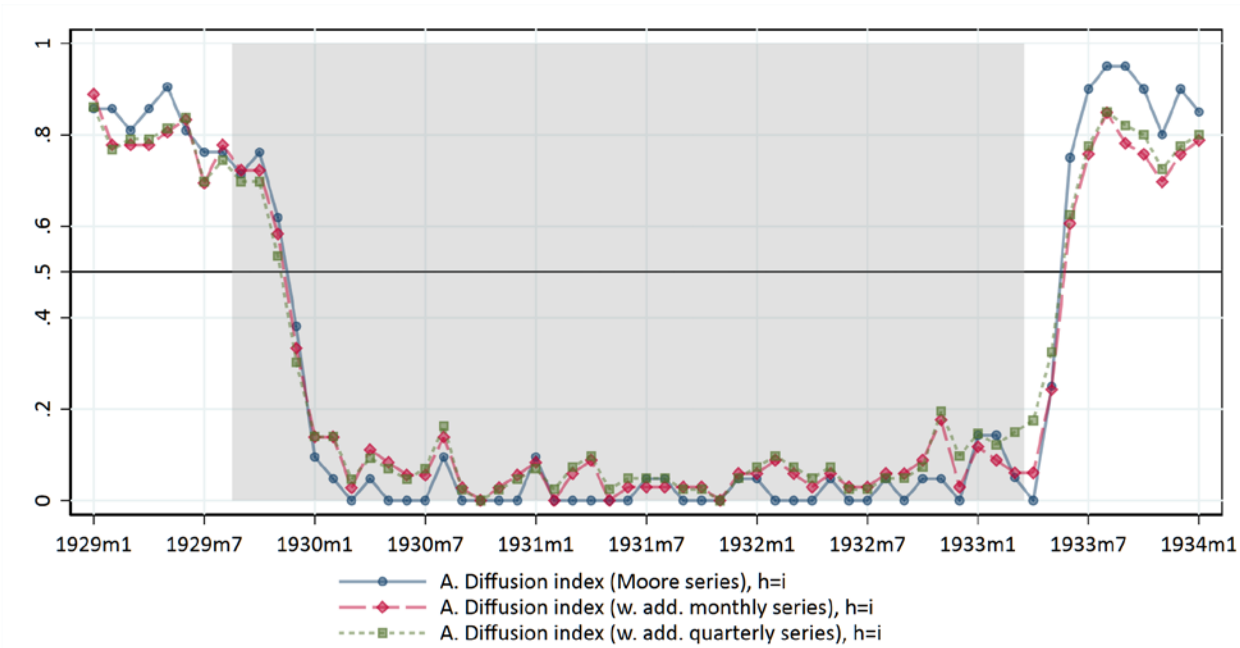

## Figure 5. Comparing the diffusion index with an indefinite horizon and the index for nowcasting

This figure compares the diffusion index with an indefinite horizon and the index for nowcasting, i.e., with a horizon of zero month. Both series are aligned the same way, that is, the horizontal axis shows the date when the forecasts are made (which, given the horizon, is the same as the target date). Shaded areas correspond to NBER recession periods.



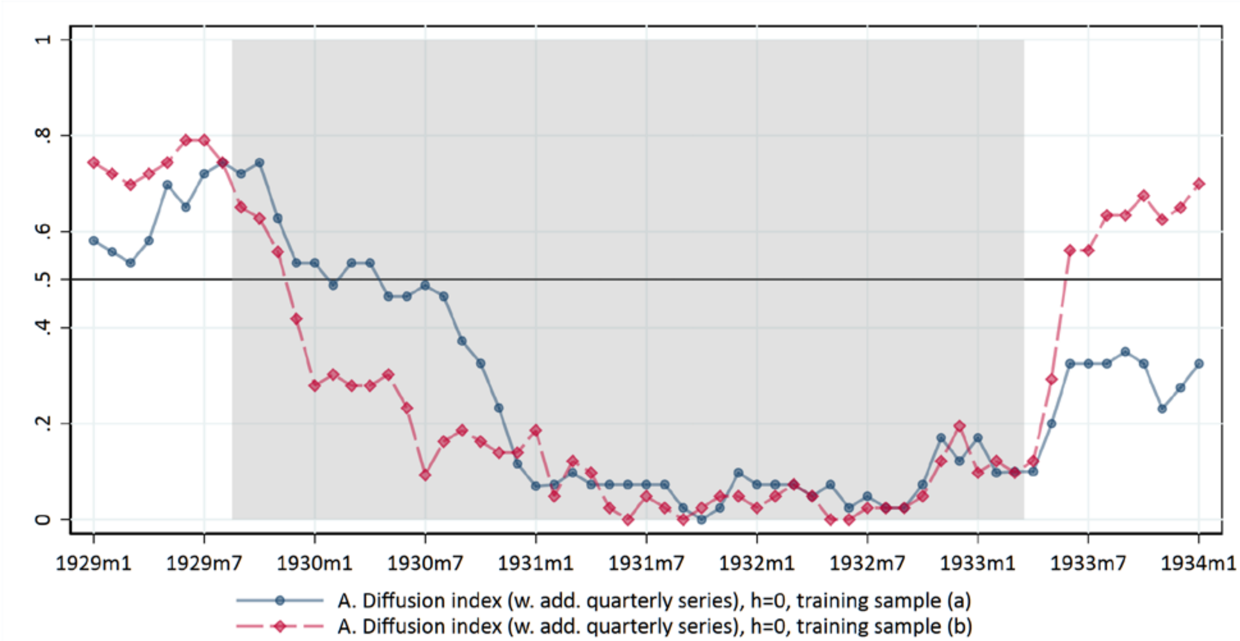
## Figure 6. Comparing the diffusion index constructed using different data sets

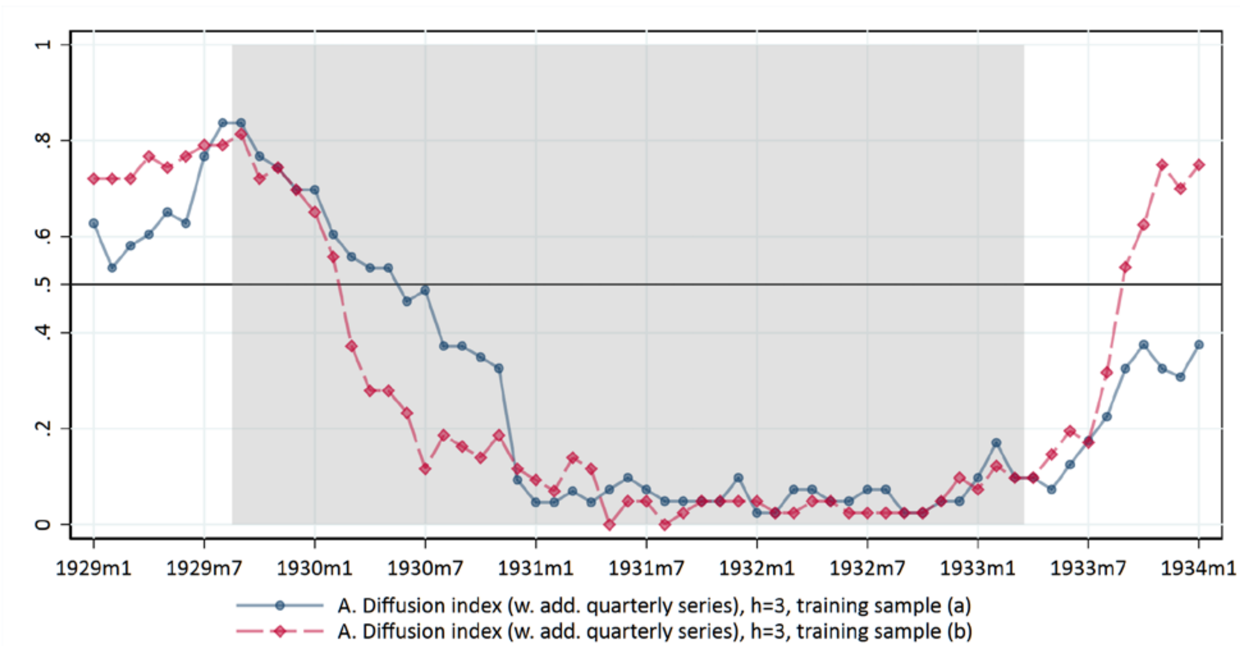
This figure compares the diffusion indexes constructed using the data set containing Moore's 21 series and data sets with the additional monthly and quarterly series we selected. The three data sets are progressively larger, i.e., the forecasts labeled as made with additional quarterly series use all the data we have available. Shaded areas correspond to NBER recession periods.



## Figure 7. Comparing the effect of different training samples on diffusion indexes with specific horizons

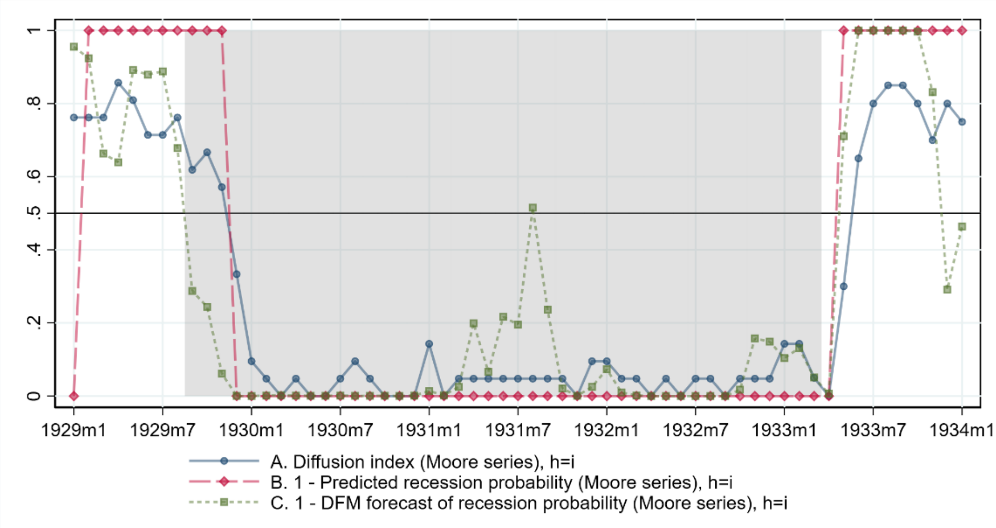
This figure compares the horizon-specific diffusion indexes whose components' lead time is determined using different training samples. Training sample (a) covers the three recessions before the 1929 and training sample (b) covers only the recession immediately preceding that of 1929. The top plot shows the indexes for nowcasting and the bottom plot shows the indexes used to forecast with a horizon of three months. Shaded areas correspond to NBER recession periods.

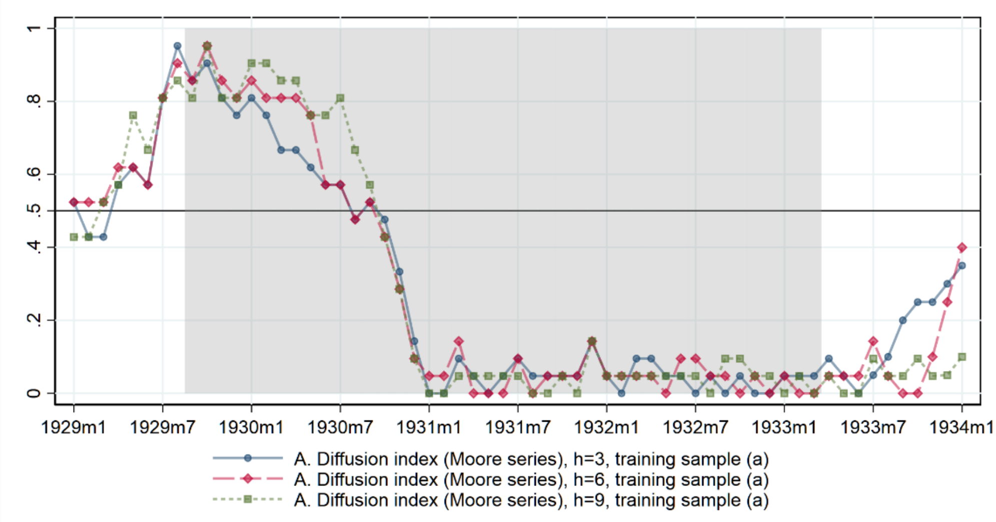


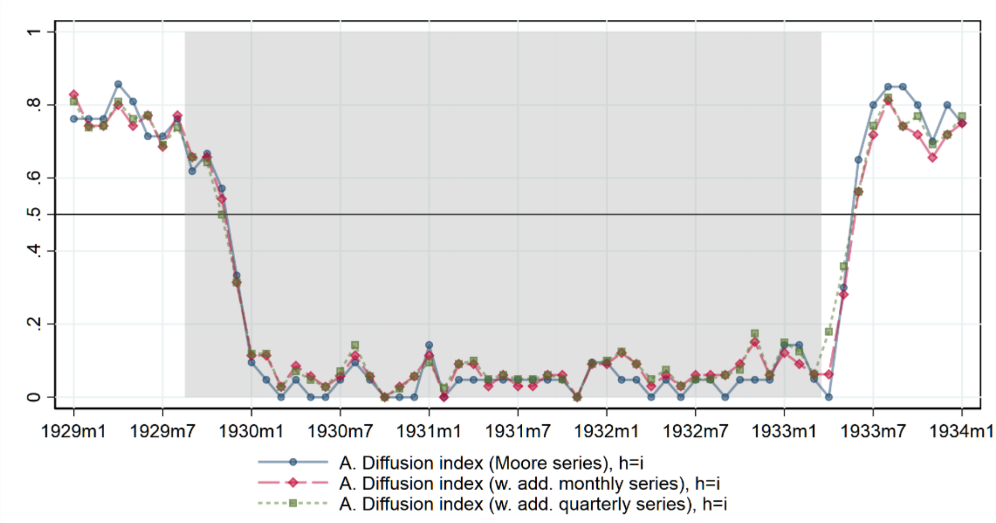



## Figure 8. Indexes constructed using real-time data vintages

In this set of plots, we repeat some of the earlier exercises using forecasts constructed with appropriate real-time data vintages. The top plot repeats the same exercise shown in the top plot of Figure 3. The middle plot repeat the same exercise shown in Figure 4. The bottom plot repeats the same exercise shown in Figure 6. Note that, for ease of interpretation, for indexes B and C, we report one minus the predicted recession probability so that low values in the figure correspond to recession periods. Shaded areas correspond to NBER recession periods.

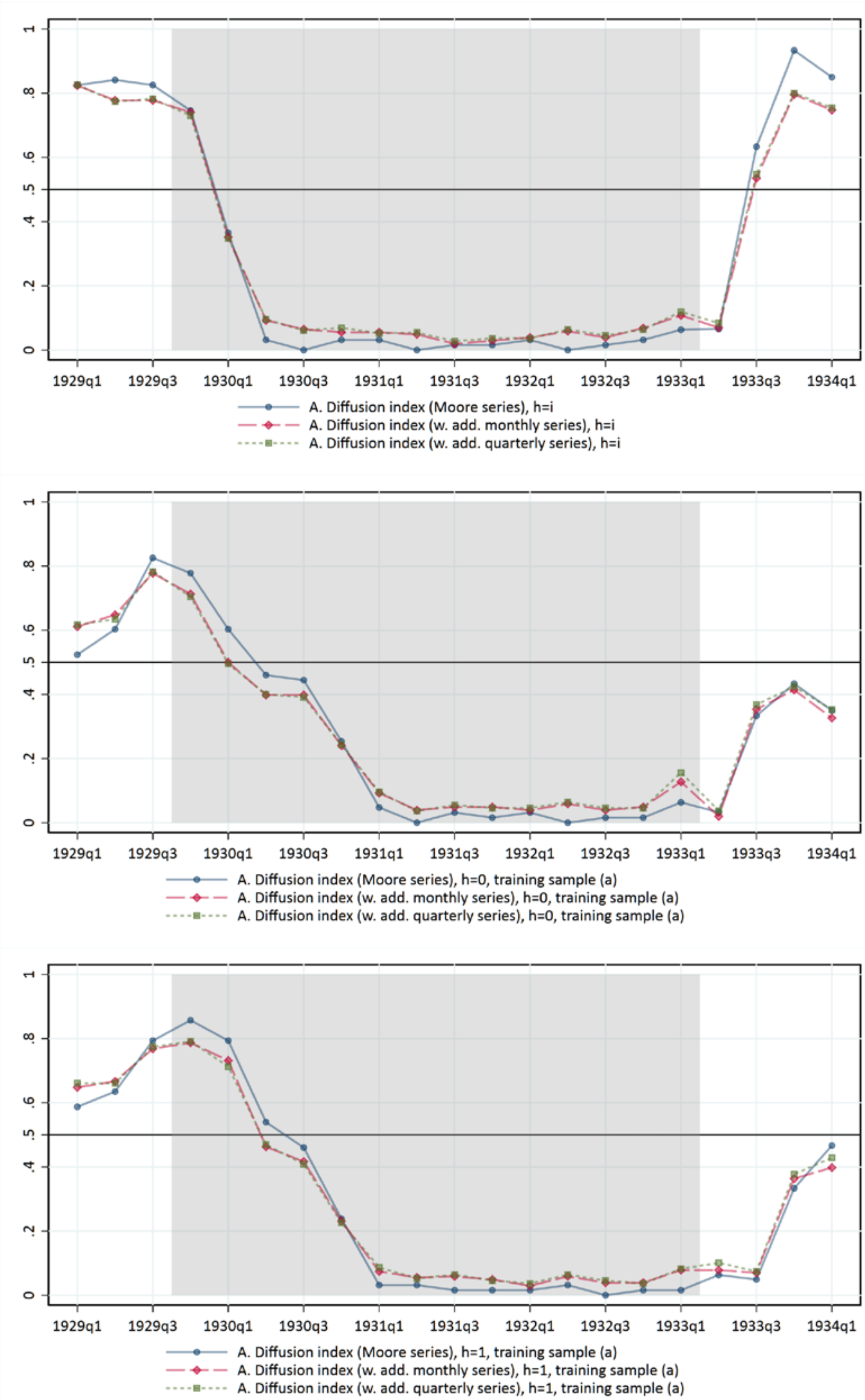


## Figure 9. Comparing quarterly diffusion indexes constructed using different data sets

In this figure, we plot the quarterly diffusion indexes. The top plot compares the three indexes with indefinite horizon that are constructed using the three data sets. The middle plot shows the comparison for the indexes constructed for nowcast (i.e., $h = 0$). The bottom plot shows the comparison for $h = 1$. Note that the horizons are stated in quarters, and given the reduced number of observations, training sample (b) is not applicable here. Shaded areas correspond to NBER recession periods.



## Online Appendix A. Description of the factor model

As reported in the paper, the second set of forecasts we make in addition to the standard diffusion index is based on the dynamic factor model (DFM) used in Zhao (2020), which is built on important earlier contributions including Giannone *et al*. (2008) and Lahiri *et al*. (2016). Specifically, let $t \in \{1, \dots, T\}$ be the time index and let $x_t$ be the $N \times 1$ data vector containing index components described in the paper, i.e., $T_j(c_j), j = 1, 2, \dots, N$. We extract the latent factors $F_t$, an $r \times 1$ vector, from the data and we model $F_t$ as an autoregressive process:

$$x_t = \theta + D \cdot F_t + u_t \tag{A1}$$

$$F_t = A \cdot F_{t-1} + B \cdot v_t \tag{A2}$$

where $\theta$ is an $N \times 1$ vector of constants, $D$ is an $N \times r$ matrix of factor loadings, $u_t$ is an $N \times 1$ vector of index component-specific innovations, $A$ is an $r \times r$ matrix with all roots of $\det(I_r - Az)$ outside the unit circle, $B$ is an $r \times q$ matrix of rank $q$, and $v_t$ is a $q \times 1$ vector of common shocks. As were in the case in the literature, $u_t$ is assumed to be cross-sectionally diagonal and orthogonal to the common shocks $v_t$. All error terms are assumed to be Gaussian. In our implementation, we follow the above-cited literature and set $r = q = 2$. The estimation procedure of the model accounts for any missing observations in any of the index components. In the initial step, principle components are obtained from a fully balanced panel of the index components created by discarding incomplete observations. Then, given these preliminary estimates, we obtain estimates of the model's parameters, which then allows the Kalman smoother to provide estimates of the latent factors for the entire sample period, including any incomplete observations. Both the model's parameters and the factors are consistently estimated.[27] The factor estimates are then used in the maximum likelihood estimation of a standard probit model that connects the state-space model of (A1) and (A2) and our target variable of interest, the binary state of the business cycle $r_t$, where a value of one indicates recession:

[27] On this issue, please refer to Giannone *et al*. (2008), a part of which we quote below:

> Although the model does not allow for cross-sectional and serial correlation of the idiosyncratic component, consistency is achieved under more general assumptions. The key insight to understand this robustness property of the estimator is the same as for simple principal components: due to the law of large numbers, the idiosyncratic component becomes negligible as the cross-sectional dimension increases. As a consequence, as far as it is confined to the idiosyncratic part, the misspecification of the model does not compromise consistency. For the same reason, the estimates of the common factors are likely to be robust to the presence of data revisions provided the revision errors are weakly cross-correlated.

$$r_t^* = \beta \cdot F_t + \mu_t \quad \text{(A3)}$$

where $r_t^*$ is the latent continuous variable with $r_t = \mathrm{I}(r_t^* > 0)$, $\beta$ is a vector of coefficients, and $\mu_t$ is the usual error term.

## Online Appendix B. On predicting the severity of the Great Depression using the factor model

In this appendix, we turn to the challenge of forecasting the severity of the Great Depression. In line with the exercises in existing literature, we consider forecasting the next 18 months at a 3-month interval from Mar 1929 to Mar 1931. In other words, at the end of each quarter during this period, we make a set of forecasts covering the next 18 months. Our target variables for this exercise are industrial production and wholesale prices, percent change from a year ago. Since the diffusion index is devoid of magnitude information from its components, it cannot be used for this purpose. The standard method in the literature uses time series models such as the vector autoregression. We follow this approach in spirit. Specifically, we do not attempt to create a structural model. Instead, we use the factor model that we reported above when forecasting the turning points. This model has been successfully applied to nowcasting and forecasting variables like real GDP and consumption.[28] We use this model to estimate the common factors in the same way as before, and then use the factors in a linear regression model to forecast our target variables. The forecasting equation contains three lags of the dependent variable and three lags of each of the factors. As in our exercises above, we consider three different data sets: The first contains only Moore's series; the second is augmented with additional monthly variables; and the third further augmented with additional quarterly variables. Figure B1 shows these forecasts.

[Figure B1 here]

The key question for us is whether the severity and the continuation of the depression are predictable, and if so, how much in advance. In particular, we consider whether the predictability changes after the stock market crash of October 1929. As discussed in Section 1, conventional wisdom suggests little predictability at longer horizons, while small output declines are foreseeable

[28] See Giannone *et al*.(2008) and Lahiri *et al*. (2016).

at shorter horizons, c.f., Dominguez *et al*. (1988). Our results are largely similar. The models here successfully predict the decline in output even at longer horizons. However, the decline is predicted to be modest in magnitude, and the predictions remained modest even after the stock market crash in Oct 1929. In fact, as shown in Figure B1, output forecasts point to a soon-to-come recovery even after steep and sustained declines in the actual values. This is similar to the qualitative forecasts made by the business press in this period as well, which forecasted recovery "right around the corner" (Mathy and Stekler 2017). The severity of the price drop is better predicted in the sense that none of the forecasts in the year of 1930 point to an imminent recovery. However, at no point in our exercise do we obtain a price forecast as low as the realized value. As for the role of the additional variables we used to augment Moore's data set, Figure B1 shows that the forecasts made with these variables better state the severity of the depression, especially when we focus on the forecasts made in 1929 and early 1930. However, the differences are modest – even with these additional variables, it remains difficult to foretell the true severity and extent of the depression in its early months.

## Online Appendix References


Dominguez, Kathryn M., Ray C. Fair, and Matthew D. Shapiro. "Forecasting the depression: Harvard versus Yale." *The American Economic Review* (1988): 595-612.

Giannone D, Reichlin L, Small D. "Nowcasting: the real-time informational content of macroeconomic data." *Journal of Monetary Economics* 55 (2008): 665–676.

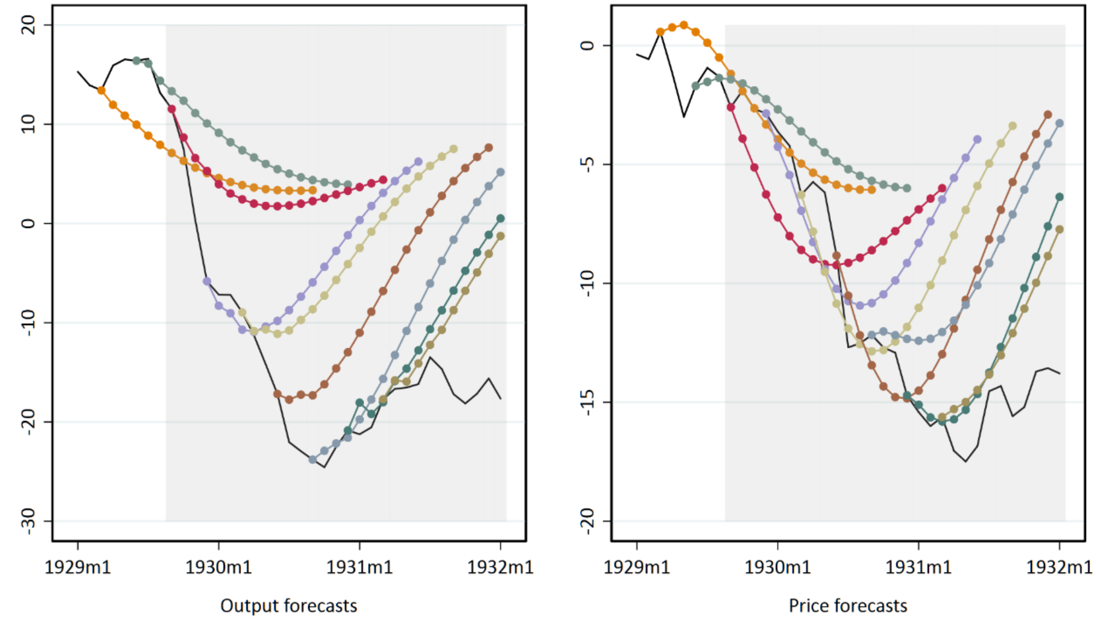
Lahiri, K., Monokroussos, G., and Zhao, Y. "Forecasting Consumption: the Role of Consumer Confidence in Real Time with many Predictors." *Journal of Applied Econometrics* 31 (2016): 1254– 1275.

Mathy, Gabriel, and Herman Stekler. "Expectations and forecasting during the Great Depression: Real-time evidence from the business press." *Journal of Macroeconomics* 53 (2017): 1-15.

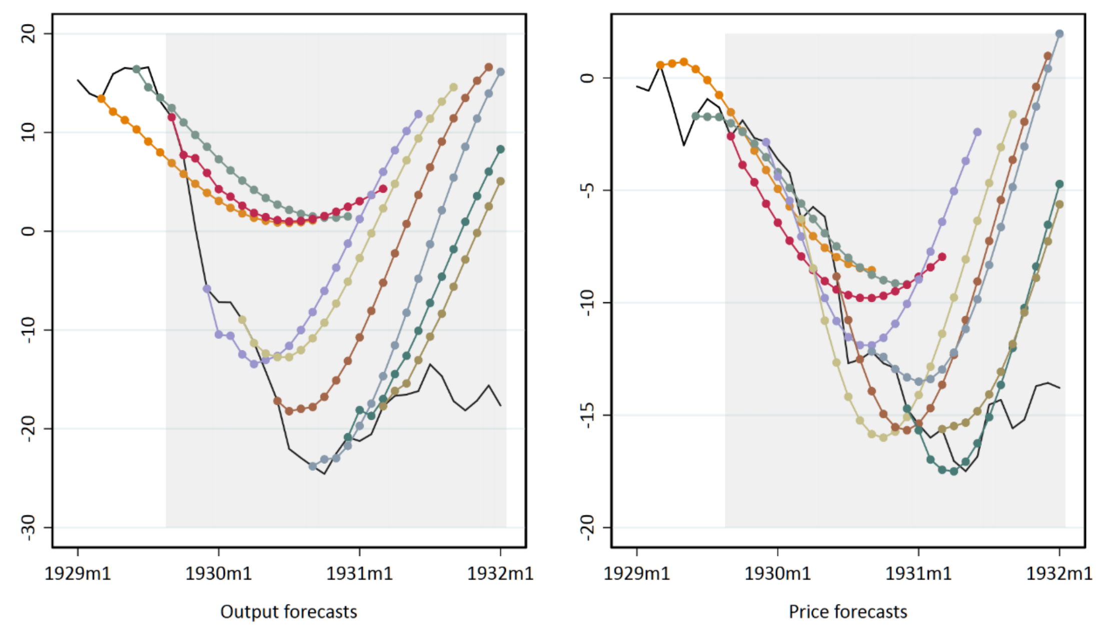
Zhao, Yongchen. "Predicting US Business Cycle Turning Points Using Real-Time Diffusion Indexes Based on a Large Data Set." *Journal of Business Cycle Research* 16, no. 2 (2020): 77-97.

## Figure B1. Comparing quarterly diffusion indexes constructed using different data sets

This figure shows the forecasts of the factor model made at a 3-month intervals for the next 18 months from Mar 1929 to Mar 1931. Our target variables for are industrial production and wholesale prices, percent change from a year ago. In each plot, the left panel shows the output forecasts and the right panel shows the price forecasts. The top plot shows the forecasts made using Moore's 21 series. The bottom plot shows the forecasts made using all available data including the additional monthly and quarterly variables. Shaded areas correspond to NBER recession periods.




Data set contains Moore series.




Data set contains Moore series and additional monthly and quarterly variables.