

**Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.**

**Financial Conditions and Economic Activity: Insights from
Machine Learning**

Michael T. Kiley

2020-095

Please cite this paper as:

Kiley, Michael T. (2020). "Financial Conditions and Economic Activity: Insights from Machine Learning," Finance and Economics Discussion Series 2020-095. Washington: Board of Governors of the Federal Reserve System, <https://doi.org/10.17016/FEDS.2020.095>.

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

Financial Conditions and Economic Activity: Insights from Machine Learning

Michael T. Kiley¹

November 2, 2020

Abstract

Machine learning (ML) techniques are used to construct a financial conditions index (FCI). The components of the ML-FCI are selected based on their ability to predict the unemployment rate one-year ahead. Three lessons for macroeconomics and variable selection/dimension reduction with large datasets emerge. First, variable transformations can drive results, emphasizing the need for transparency in selection of transformations and robustness to a range of reasonable choices. Second, there is strong evidence of nonlinearity in the relationship between financial variables and economic activity—tight financial conditions are associated with sharp deteriorations in economic activity and accommodative conditions are associated with only modest improvements in activity. Finally, the ML-FCI places sizable weight on equity prices and term spreads, in contrast to other measures. These lessons yield an ML-FCI showing tightening in financial conditions before the early 1990s and early 2000s recessions, in contrast to the National Financial Conditions Index (NFCI).

Keywords: Big Data; Recession Prediction; Variable Selection

JEL Codes: E17, E44, C55, E50

¹ The analysis and conclusions set forth are those of the author and do not indicate concurrence by the Federal Reserve Board or other members of its staff. This research has benefitted from the research assistance of Charles Horgan and comments by Michael McCracken and Minchul Shin, as well as workshop participants at the Federal Reserve Board and the 2020 Annual Meeting of the Central Bank Research Association. Email: mkiley@frb.gov.

1. Motivation

A broad set of financial conditions affect household and business demand: Equity prices influence household consumption through wealth effects; interest rates shape mortgage borrowing and home purchases as well as business investment; and exchange rates affect spending on imported and domestic goods and services. Researchers, the private sector, and policy institutions have developed indexes summarizing financial conditions.²

I uses prediction approaches popular in the machine-learning (ML) literature to reconsider construction of a financial conditions index (FCI). Earlier research, including English et al (2005), Hatzius et al (2010), and Aramonte et al (2017), examined whether FCIs constructed from principal-components analysis or similar techniques were useful in prediction equations for economic activity. Building off this work, the approach herein defines the FCI as the aggregate of financial variables selected to predict economic activity by the ML algorithms. The measure of economic activity chosen is the unemployment rate, reflecting the central role this measure plays as a summary statistic for the state of the business cycle in the United States.

The ML algorithms considered include LASSO, elastic net, and random forests. Varian (2014) and Mullainathan and Spiess (2018) provide accessible introductions to these approaches for an economics' audience and suggest references that explore their properties in more detail.³ For the purposes herein, three considerations determined the choice of algorithms. First, they are popular in the prediction literature. Second, LASSO and elastic-net represent algorithms for selection of predictors from a potentially large set of possibilities that have a form similar to econometric techniques commonly used by

² E.g., Guichard, 2009; Hatzius et al, 2010; Brave and Butters, 2011; Matheson, 2012; Koop and Korobilis, 2014; Hatzius and Stehn, 2018 (which presents the Goldman Sachs approach); Bartsch, 2018 (which presents the BlackRock approach); Federal Reserve Bank of Kansas City, 2019; Federal Reserve Bank of Cleveland, 2019; and IMF, 2019.

³ Reichlin et al (2017) and Athey (2018) discuss a broader set of related "big data" issues. The analysis herein focuses on simple LASSO; Diebold and Shin (2019) discuss alternative LASSO approaches that may perform better in certain applications.

economists. Third and finally, the random forest approach represents an approach that differs notably from standard econometric techniques and provides a means to examine the robustness of key results from LASSO and elastic net. Importantly, LASSO and elastic net are linear equations for prediction, whereas a random forest can capture nonlinear relationships.

The resulting ML-FCIs shows sizable differences from other FCIs, notably from the Federal Reserve Bank of Chicago's National Financial Conditions Index (NFCI). A striking difference is the state of financial conditions before the recessions of the early 1990s and the early 2000s. During those periods, the NFCI did not tighten, suggesting that a tightening of financial conditions was not an important precursor to those recessions. In contrast, the ML-FCI tightens during those periods, suggesting that a tightening in financial conditions was a contributor to the subsequent downturns. One reason for this difference is that the variable selection approaches deployed are tuned to be correlated with future unemployment—so the ability to capture business cycle aspects is hard-wired in the approach. An out-of-sample exercise considers the robustness of the results and demonstrates that some of the peak in the ML-FCI before a recession weakens out-of-sample. However, the approach herein reveals a number of useful insights and is not primarily focused on forecast performance.

In particular, the ML-FCI places a significant weight on term spreads and equity prices as a contributor to the financial conditions relevant for economic activity, in contrast to the NFCI. Research has long suggested that a relatively flat or inverted yield curve (i.e., long-term interest rates that are low relative to short-term interest rates) foreshadows a weakening in economic activity (e.g., Stock and Watson, 1989; Estrella and Mishkin, 1996; Rudebusch and Williams, 2009). The ML approach selects the near-term slope of the yield curve as among the variables that predict unemployment, as in Engstrom and Sharpe (2019).⁴ The idea that equity prices are among the key determinants of economic activity

⁴ Other variables emphasized in some research are not significant contributors to the ML-FCIs herein. For example, the NFCI places large weights on measures of corporate bond spreads as

has a long history and central role in the literature on monetary policy transmission (e.g., Boivin, Kiley, and Mishkin, 2010). Paul Samuelson summarized the idea that equity price movements may signal future economic activity in his humorous quip, from 1966, that the stock market has predicted nine of the past five recessions. Despite the role of equity prices as an impetus to activity in theory, several FCIs show little imprint from, for example, the run-up and subsequent decline of equity prices in the late 1990s and early 2000s.

Perhaps more importantly, the ML approach yields these insight for reasons that shed light on important issues for macroeconomics, forecasting, and analyses of big data (specifically, variable selection and dimension reduction). In particular, the ML approach naturally allows consideration of a range of variables and transformations as well as nonlinearity. Data transformations are often chosen on an a priori basis for reasons related to the stationarity of the variables (e.g., Forni et al, 2000; Stock and Watson, 2002, and 2011, 2016; McCracken and Ng, 2016). However, there are many transformations that may render a variable stationary, and these alternative transformations may emphasize different aspects of the series (e.g., introduce peaks in the spectrum of the data at different points in the frequency domain).⁵ As a result, these choices are consequential, and the ML approaches allow an algorithmic approach to consideration of such choices. The ML approach herein considers both high-frequency transformations (e.g., percent changes in equity prices) as well as lower frequency transformations (e.g., twelve-month percent changes or the level of price-to-earnings ratios). The algorithms select the lower-frequency transformations as carrying the information in equity prices that is relevant for predicting unemployment and assigns an important role for equity prices in the ML-FCI. In contrast, the related literature,

determinants of financial conditions—a finding consistent with direct examinations of the predictive power of bond spreads such as Gilchrist and Zakrajsek (2012)—while these measures receive relatively small weights in the ML-FCI. In addition, the ML-FCI places little weight on the level of interest rates, a finding similar to that embedded in construction of the NFCI. This finding is difficult to square with the central role of the level of interest rates in the conduct of monetary policy and could reflect a number of factors, including trend shifts in the equilibrium real interest rate (e.g., Holston et al, 2017; Kiley, 2020a & 2020b).

⁵ The role of transformations is documented in the literature, e.g., in analysis of business-cycle facts and the Hodrick and Prescott (1997) filter (e.g., Cogley and Nason, 1995; and Hamilton, 2018).

such as the NFCI, only considers the transformations that emphasize higher frequencies. Sensitivity of results to variable choice and data transformations is important for other “big data” approaches in economics and finance including dimension-reduction (shrinkage) techniques such as dynamic factor models, as variable transformations are often specified on a priori grounds with limited discussion of the sensitivity of results to such choices (e.g., Stock and Watson, 2002, 2011 and 2016; Reichlin et al, 2017; and Kiley, 2018).

The FCI emerging from the random forest approach also shows striking evidence of nonlinearity. Tight financial conditions are associated with large subsequent increases in the unemployment rate. Accommodative financial conditions are associated with only modest declines in the unemployment rate. These nonlinear relationships are consistent with macroeconomic theories in which financial constraints can weigh heavily on economic activity as they become increasingly binding, as emphasized in Barnichon et al (forthcoming). Despite the emphasis on potential nonlinearities in macroeconomic theory, previous empirical work has not emphasized such nonlinear relationships between financial conditions and economic activity, with limited exceptions including Barnichon et al (forthcoming) and references therein. The findings of nonlinearity demonstrate the value of considering the machine-learning approaches such as a random forest to identify features of the data. Such features cannot be detected by linear approaches including linear approaches that identify the common factor of interest through the correlation with a variable of interest (as is done herein using the ML approaches for the unemployment rate), such as the algorithms in Brauning and Koopmans (2014), Kelley and Pruitt (2015), and Brave et al (2019).

The next section reviews the ML approaches considered. Section 3 presents core results for the LASSO and elastic-net approaches. Section 4 presents from a random forest. Section 5 considers some robustness exercises, and section 6 concludes.

2. Constructing a Financial Conditions Index via Machine Learning

Machine learning techniques take large amounts of data to create tools for decision-making. Their strength lies in their ability to combine large numbers of observations of different types of data – numerical, textual, or other – to produce tools that predict outcomes. A key weakness, especially in economic contexts, is the difficulty that arises in interpreting the resulting prediction algorithm (which may be large, complex, and nonlinear, resulting in a “black box”).⁶

The problem herein is relatively simple. As a result, a simple approach, using a moderately large number of financial indicators, is used. Given a set of financial variables, I search for a summary index of financial conditions—an FCI. I define financial conditions as the combination of the financial variables most closely connected with economic activity—thereby adopting the perspective that defines popular FCIs (e.g., by the investment banking firm Goldman Sachs, as presented in Hatzius and Stehn, 2018) and that underlies analyses linking FCIs to real activity (e.g., Adrian, Boyarchenko, and Giannone, 2019). To operationalize the notion of connection with economic activity, I look for an FCI that is correlated with the unemployment rate one-year ahead; in the monthly dataset (described below), this means looking for an FCI to predict the unemployment rate twelve-months ahead, at $t+12$.

This approach to construction of an FCI looks like a standard econometric problem. The role of ML approaches is to help reduce the dimension of the financial variables entering the FCI, as a large number of financial variables may not be amenable to standard econometric forecasting techniques (e.g., least squares). I will return to the question of whether the ML approaches provide value beyond typical econometric approaches in the robustness analysis, along with a brief discussion of the implications of the results for alternative dimension-reduction techniques (such as principal components or dynamic factor models).

⁶ Varian (2014) and Mullainathan and Spiess (2018).

For the remainder of the analysis, let

- $U(t)$ denote the unemployment rate in month (period) t ;
- $x^i(t)$ denote the i th financial variable (of N total variables) in period t ;
- $X(t)$ denote the $1 \times (N+1)$ vector containing all financial variables and the current (period t) value of the unemployment rate); and
- X denote the $T \times (N+1)$ matrix of containing all observations (T) of financial variables.

The approaches to constructing the FCI follow two different paths. The primary path uses ML techniques most similar to regression – LASSO and elastic net. The robustness of key results from this path are examined through comparison to results from a random forest. I first discuss the data and then turn to the ML algorithms.

2.1 Data

All variables used in the analysis are described in table 1 and are taken from the monthly macroeconomic database of McCracken and Ng (2016), which supports replicability. Data span the period January 1962 to the first half of 2020, with estimation based on the sample ending in December 2019.⁷

The unemployment rate, the variable the ML-FCI is tuned to predict, is measured by the civilian unemployment rate for the non-institutional population age 16 and over. Nominal interest rates are transformed to real interest rates using the trailing 12-month percent change in the all items CPI.

Financial variables are included in the analysis based on a mix of economic and data availability factors. Measures of real interest rates are included given the central role of the real interest rate in macroeconomics and monetary economics (e.g., such rates are traditionally the measure of financial conditions in an IS curve). Equity prices are included

⁷ The sample period does not include the impact of COVID-19, which resulted in a large increase in the unemployment rate. The rise in the unemployment rate induced by COVID-19 is clearly unrelated to financial conditions twelve months prior, and future econometric work will need to address the implications of such an outsized move in the unemployment rate.

as a widely watched indicator with clear macroeconomic interpretation (e.g., as part of Tobin's Q influencing investment). Factors related to the slope of the yield curve are included given the empirical record in predicting recessions. Corporate bond spreads are included in light of their role in the cost of finance for large corporations, the idea that they may capture financial frictions, and previous empirical work. Exchange rates are included given their role in international trade and finance.

The financial predictor variables (i.e., elements of $X(t)$ excluding the unemployment rate) can be organized into five categories.

- **Level of (real) interest rates:** This set includes the federal funds rate, the commercial paper rate, three-month and six-month Treasury bill rates, rates on Treasury securities with maturities of one, five, and ten years, and Moody's Seasoned Corporate Bond yields for Aaa and Baa ratings. In all cases, these nominal interest rates are transformed to "real" interest rates by subtracting the average rate of inflation, as measured by the Consumer Price Index, over the preceding three years (36 months). (Note this differs from the transformation suggested by McCracken and Ng (2016), which involved first differencing the nominal rates.)
- **Term spreads (yield curve slope):** This set includes the three-month, six-month, one-year, five-year, and ten-year Treasury rates, all minus the federal funds rate.
- **Credit spreads:** This set includes the three-month commercial paper rate minus the federal funds rate and Moody's Seasoned Corporate Bond yields for Aaa and Baa rating minus the yield on the ten-year Treasury.
- **Equity Prices:** This set includes the S&P 500 and Industrials Composite Price Indexes and the S&P Composite dividend yield and price-to-earnings ratio. Two transformations for each variable are considered. The low-frequency transformation uses the twelve-month percent change for the price indexes and the levels for the two ratios. The high-frequency transformation uses the simple percent changes for the composite price indexes and the simple change for the dividend yield (as suggested by McCracken and Ng, 2016).

- **Exchange rates:** This set includes the exchange value of the U.S dollar relative to the Swiss Franc, Japanese Yen, U.K. Pound Sterling, and the Canadian dollar. In each case, the set includes two transformations—the twelve-month percent change (to capture lower frequencies) and the simple percent change (as suggested by McCracken and Ng, 2016).

All told, the set of financial variables, including alternative transformations, results in 33 elements of $X(t)$ (in addition to the unemployment rate). This set of predictors is not linearly independent, as, for example, it includes the levels of interest rates and spreads computed from such levels. This is not a technical challenge for the algorithms, although it could lead to problems such as instability in the set of selected predictors—an issue analyzed in the robustness section.

While this set of variables is smaller than that used in some FCIs (e.g., Brave and Butters, 2011), it is similar-in-size or larger than the set used in many FCIs.

2.2. LASSO and Elastic Net

As the goal is to gauge the role of various financial variables as predictors for the unemployment rate one-year ahead, the problem relates to variable selection. The class of penalized regression techniques used in ML applications is well suited to this task. The robustness section will return to the challenges associated with variable selection given limited observations and correlated predictors.

The penalized regression approaches considered choose the weights in the ML-FCI, B , to minimize

$$\sum_{T'} (U(j + 12) - X(j)B)^2 + \lambda \left(\sum_{N+1} (\alpha|B(j)| + (1 - \alpha)B(j)^2) \right)$$

where T' is the sample used to estimate the parameters. Note that this sample size may be less than the full set of observations T , in order to allow assessments of fit outside the estimation window as a tool to avoid overfitting the data.

The parameter λ governs the weight given to the penalty function on parameters. λ equal to 0 returns least-squares estimates. In LASSO, α equals 1—that is, LASSO penalizes the absolute value of coefficients and hence prefers to set coefficients to zero (i.e., not select a variable) if the coefficient is small, all else equal. Elastic net involves choosing α between 0 and 1. Both LASSO and elastic net have been widely used for prediction and variable selection in ML contexts and are computationally efficient to estimate (e.g., Varian, 2014; Mullainathan and Spiess, 2018).⁸

Several observations are important. First, the data used to predict the unemployment rate twelve-months ahead, $X(t)$, includes the current level of the unemployment rate, reflecting the persistence in the level of this series. As the interest herein is in the role of financial conditions and not the intrinsic persistence in the unemployment rate, the ML-FCI is computed as $\hat{X}\hat{B}$, where the “hat” indicates that the column/row in the data/coefficient vector corresponding to the unemployment rate has been deleted.

I implement these approaches in R using the `glmnet` package. Rather than split the data into a training and test datasets, I use k-fold cross validation in the estimation of parameters with five folds—i.e., k-fold cross validation is used to select the subsample for estimation T . I do not use a training set in the baseline results as I am interested in understanding relationships across the full set of data. In order to eliminate randomness in the results related to the choice of folds, I estimate parameters 500 times, creating folds randomly on each iteration and subsequently considering results across these iterations (via averaging). The robustness section below considers a training dataset (which ignores the final 15 years of observations).⁹ Finally, I consider values of α equal to 1 (LASSO) and 0.5 (a specific choice for elastic net), for 1000 values of λ . The results choose the largest value of λ such that the prediction error is within 1 standard error of the minimum value; this is a

⁸ α equal to 0 is ridge regression, which penalizes large value of the coefficients and hence is like a Bayesian approach in which the researcher has a prior that all variables are important, and no variable is likely to be especially important. See Hoerl and Kennard (1970) and Reichlin et al (2017).

⁹ To preview results, the ML approaches are robust along this dimension, while traditional econometric techniques, e.g., least squares, are not.

standard approach in the literature, as a choice of λ away from the value that minimizes the prediction error limits the risk of spuriously overfitting the data.

2.2 An Alternative ML Approach: Random Forests

Both LASSO and elastic net resemble traditional regressions: The equation is linear, and the estimation of parameters involves standard (albeit nonlinear) minimization of an objective function. An alternative ML approach is a random forest. Consideration of an alternative approach yields at least two advantages. First, it may yield similar or different results, thereby speaking to robustness. Second, an approach that may elicit nonlinear relationships has the ability to detect such nonlinearities, which theory suggests may be important.

Figure 1 provides a brief summary of the algorithm. First, the objective and data are defined. Next, observations to train the forest are selected. At the third stage, a subset of potential predictors are randomly chosen. Given these observations and predictors, step 4 finds the variable and associated threshold value that best splits the observations in two groups—i.e., the best branching of the tree is identified, as measured by some loss function. Step 5 involves repeating this process along each resulting branch to select new branches until some predefined number of final branches/nodes is achieved. Finally, the process of randomly selecting a subset of predictors and each subsequent step is repeated to construct another tree. A random forest is the collection of these trees (Breiman, 2001).

<<INSERT FIGURE 1 HERE>>

In the implementation, I use the entire sample to develop the forest, as I am focused on relationships within the full sample (and in parallel to the treatment using LASSO and elastic net). The robustness section considers a forest grown on a training sample that excludes the final 15 years of data. I grow 500 individual forests to construct the random forest. Each forest randomly chooses 10 potential predictors at step 3 (from a possible 34, as discussed above) and I allow the tree to have 8 final nodes. The analysis is conducted in R using the randomForest package (with supporting analysis via the randomForestExplainer

package). Note that an individual tree with 8 final nodes can only predict (up to) 8 values for the unemployment rate. The random forest, of 500 trees, can predict (up to) 4000 values. The presence of many trees in the forest allows the random forest to fit a continuous variable like the unemployment rate via an approach using a decision tree framework.

3. Results from LASSO and Elastic Net

The central results from the LASSO and elastic-net exercises are two-fold. First, LASSO and elastic net produce nearly identical FCIs and in both cases the ML-FCI shows clear cyclical properties, with notable tightening before recessions—in contrast to, for example, the NFCI. Second, equity prices play an important role in driving the ML-FCI, and this result owes importantly to the inclusion of low-frequency transformations. In addition to equity prices, term spreads are an important set of predictors. The following subsections take these points up in turn.

3.1. ML-FCIs

In principle, LASSO and elastic net could provide different estimates of the ML-FCI and/or differ significantly in the set of predictors selected. However, the results suggest that for this objective and dataset, results are very similar. This is illustrated in figure 2, which presents the ML-FCI implied by LASSO and elastic net.

<<INSERT FIGURE 2 HERE>>

In light of these similarities, I focus primarily on LASSO in the remaining discussion.

The pattern of the ML-FCI in figure 2 shows a clear lead relative to the business cycle, as should be expected given that the FCI is constructed to predict the unemployment rate one-year ahead. (Note that k-fold cross validation avoids overfitting, but the entire time period is used and hence these are in-sample results; for understanding relationships, this may be appropriate, and the robustness section considers out-of-sample results.) The cyclical pattern simply indicates that, within sample, the financial variables have predictive

power for the unemployment rate. For example, financial conditions tighten discernably in advance of the 1990s and early 2000s recession and tighten dramatically before and during the Global Financial Crisis.

These patterns contrast with those from the Federal Reserve Bank of Chicago’s NFCI, as shown in figure 3. The NFCI did not tighten appreciably in the late 1980s and early 1990s or in the late 1990s and early 2000s. The ML-FCI differs from the NFCI in two fundamental ways: Different data is used, including data transformations that focus on lower-frequencies; and the ML-FCI is designed to forecast unemployment, whereas the NFCI is from a factor model estimated to account for the comovement among the included financial variables.¹⁰ The cyclical contrast owes to the latter difference—the tuning of the ML_FCI to future unemployment—as is intuitive and will be clear in various robustness exercises in section 5. Finally, note that the NFCI is a dynamic factor model that also uses the entire sample of data.

<<INSERT FIGURE 3 HERE>>

3.2. Identified Important Predictors

As the LASSO and elastic net algorithms select from the set of predictors, it is interesting to consider the variables that are selected and how they contribute to the business-cycle properties just discussed.

Figure 4 presents the ML-FCI (from LASSO) and the contributions to this FCI from each set of predictors—interest rates, term spreads, exchange rates, equity prices, and credit spreads. Two classes of aggregates stand out as especially important. First, equity-price measures play a large role in fluctuations in the ML-FCI, especially since the 1990s. Recall that the set of variables related to equity prices include lower-frequency measures, such as the twelve-month changes in the equity price indexes or level of the price-to-earnings ratio, as well as higher-frequency measures. It turns out that the low-frequency measures—

¹⁰ Note that algorithms other than machine learning can also be tuned to a forecast an objective, e.g., Brauning and Koopmans (2014), Kelly and Pruitt (2015) and Brave et al (2019).

measures not common in factor model analyses such as that underlying the NFCI—are the important drivers. I return to this below in the random forest results, where related summary statistics are intuitive.

<<INSERT FIGURE 4 HERE>>

The second set of variables important for the FCI are term spreads. Spreads are more clearly important in the 1970s and 1980s and somewhat less important subsequently. The sizable non-zero coefficients are typically selected for the near-term yield-curve slope variables, consistent with research emphasizing the slope of the yield curve in recession prediction such as Engstrom and Sharpe (2019).

Note that other variables contribute relatively little. The set of real interest rates contributes essentially nothing, as does the set of exchange rates. The set of corporate bond spreads also contribute relatively little—a result seemingly in contrast to that of Gilchrist and Zakrajsek (2012) and related research.¹¹

4. Results from a Random Forest

The core results from LASSO (and elastic net) are both novel and intuitive: Equity-price measures have long been viewed as key macroeconomic factors (e.g., Boivin et al, 2011), but played little role in (many) FCIs. It is plausible that the high-frequency transformations popular in big-data applications hide the role of equity prices, as month-to-month changes are volatile while indicators such as twelve-month changes, dividend yields or price-to-earnings ratios capture important business-cycle components. The ML algorithms select equity-price indicators with the intuitive properties.

However, a review of the degree to which the properties of an ML-FCI are robust along several dimensions is important. First, LASSO (and elastic net) are very similar to standard

¹¹ Gilchrist and Zakrajsek (2012) find evidence for bond spreads but emphasize their measure of the excess bond premium, which removes default risk from bond spreads. This measure is not readily available in databases of macroeconomic indicators (e.g., McCracken and Ng, 2016) and hence is not included in the analysis for replicability reasons.

econometric approaches and comparison with an ML-FCI based on alternative algorithms, especially algorithms that allow for nonlinear interactions, is revealing.

A random forest differs substantially from LASSO or elastic net. In its search for predictors, an individual tree creates branches that predict whether the outcome variable (i.e., the unemployment rate one-year ahead) should be placed in one of two buckets/branches, and then branches again. While this approach may appear most natural for outcomes with discrete values/buckets, these “regression” trees can be effective prediction algorithms for continuous outcomes because many trees are grown within the forest, and hence there are many final buckets that are predicted across the forest.

Finally, this approach allows nonlinear interactions and a way to assess the importance of an indicator by examining how close an indicator is to the “trunk” of the tree—that is, an indicator is important if it is very early in the branch structure of the tree. Figure 5 provides an illustrative example. In this example, the predictor closest to the trunk of the tree is the twelve-month percent change in the S&P 500 composite index. The second most important predictors selected are the near-term slope of the yield curve (the difference between the 6-month Treasury rate and the federal funds rate) or the Baa corporate bond spread—that is, these financial conditions contain information on whether unemployment is likely to rise or fall over the coming year on top of equity prices. This branching continues until the tree is fully-grown. As outlined in section 2, this process is repeated to grow 500 trees with 8 final nodes, via random selection of potential predictors, and the random forest prediction is the average value predicted from these trees.

<<INSERT FIGURE 5 HERE>>

Recall that in the LASSO case, the ML-FCI only accounts for the role of the financial variables in predicting the unemployment rate; the role of the unemployment rate itself, owing to serial correlation, is ignored. For the random forest, an analogous approach is used. However, the approach needs to be different, as a random forest which included the unemployment rate as a predictor would include nonlinear interactions between current

unemployment and the financial variables. A simple alternative is to estimate an equation that predicts the unemployment rate twelve months ahead using only the current unemployment rate, and to fit the random forest to predict the residuals from this first stage regression using only financial variables. I use this approach, which is based on the full sample; the robustness exercise that considers out-of-sample results in section 5 uses residuals from rolling (out-of-sample) forecasts.

Figure 6 presents the resulting ML-FCIs from LASSO and the random forest. Both FCIs share similar business cycle properties, as should be expected given that they are designed to predict unemployment. Relatedly, the ML-FCI from the random forest looks much more like the ML-FCI from LASSO than the NFCI.

However, there is one striking difference: the ML-FCI from the random forest is very nonlinear. Tight financial conditions are associated with a high value for the unemployment rate (as that is what a high value of the FCI implies). In contrast, the FCI only becomes modestly accommodative—indeed, it reaches a floor implying a modest low value for the unemployment rate. This pattern cannot emerge from a linear approach (such as LASSO, elastic net, or a dynamic factor model). It is also consistent with macroeconomic theory predicting such nonlinearities and a limited set of previous empirical work (e.g., Barnichon et al, forthcoming).

<<INSERT FIGURE 6 HERE>>

While the similarity in business-cycle properties is expected given the common objective (predict unemployment) behind both ML-FCIs, the degree to which the algorithms emphasize similar predictors is of central interest. Figure 7 presents the branch level of the top 5 predictors from the random forest—a branch level of one indicates the predictor is at the trunk and hence very important for classifying outcomes, whereas a high branch level indicates the predictor is less important. Within financial variables, measures of the term spread (e.g., the six-month Treasury yield relative to the federal funds rate and others) are four of the top five predictors. The top predictor is the twelve-month percent change in

the S&P 500 (and two other equity price indicators fall in the top ten predictors). These results echo those from LASSO along the critical dimensions highlighted earlier—term spreads and equity prices are key measures of financial conditions, and the low-frequency transformations of equity prices are the important measures among that set of indicators.

<<INSERT FIGURE 7 HERE>>

5. Robustness

The dataset used is small (by ML standards, albeit not by macroeconomic standards) and results on variable selection can be fragile with correlated predictors (e.g., de Mol, 2008; and Mullainathan and Spiess, 2018). As a result, it is useful to examine the degree to which the results are robust by ignoring the final 15 years of data. It is also useful to compare the ML-FCI results to a simple least-squares prediction and the fragility of that approach relative to the ML approaches. Finally, the machine-learning approach focuses on variable selection. An alternative approach to “big data” in macroeconomics is dimension reduction—indeed, the NFCI is an example of that approach. The final robustness exercise demonstrates how this approach can be dependent on choices regarding variable inclusion and transformations—thereby demonstrating the value of the approaches herein that allow selection across alternative transformations.

5.1. Stability: Training and Test Data

Variable selection among correlated predictors is inherently challenging and subject to instability: Because certain sets of predictors are highly correlated, it can be largely irrelevant which individual predictors are chosen and/or the set of chosen predictors can shift with small changes in specification or sample period (even though the resulting predictions may be very similar) (as discussed in de Mol, 2008, and Mullainathan and Spiess, 2018). These fragilities underlie the preference in some econometric applications for factor-model approaches (e.g., Forni et al, 2000; Stock and Watson, 2002, 2011, and 2016; de Mol et al, 2008). The random forest results provide some evidence of robustness in variable selection to change in specification.

A further robustness check involves examination of results ignoring some of the data in estimation. To this end, the analysis was executed ignoring the final 15 years of data in the estimation step (and thereby excluding the most severe tightening in financial conditions in many decades, as the estimation sample ends in 2004, four years before the most intense phase of the Global Financial Crisis). The projections were then created for the full sample. Figure 8 presents the resulting ML-FCI from LASSO and the random forest, in each case comparing the full-sample version to the restricted-sample version. There are some important differences. Most notably, the out-of-sample performance is clearly worse, as the tightening in financial conditions around the Global Financial Crisis is less severe and occurs later when the crisis is not in the estimation sample. At the same time, the character of the ML-FCI is broadly similar and remains distinctly different from that of the NFCI. In this sense, the ML algorithms deliver robust results across the full and restricted samples. Moreover, the nonlinear character of the results from the random forest is preserved.

<<INSERT FIGURE 8 HERE>>

Evidence of robustness in variable selection to sample period is shown in figure 9, which presents the contributions to the LASSO ML-FCI from each set of variables. As in the full sample, term spreads are important in the estimation window. Equity-price measures are also important in the estimation window, especially in the loosening of financial conditions in the late 1990s and subsequent tightening. That said, the results out of sample demonstrate that the role of equity prices in the tightening around the Global Financial Crisis estimated using the full sample only comes through partially when that episode is not included in the estimation window.

<<INSERT FIGURE 9 HERE>>

Figure 10 presents the top five predictors from the random forest for the full sample and restricted samples. Again, equity price measures are among the most important, with the 12-month percent changes in both the S&P industrials index and the S&P 500 index within

the top five predictors. In addition, yield-curve slope variables are consistently in the top five. These results point further to robustness in the ML approach.

<<INSERT FIGURE 10 HERE>>

A final consideration with regard to subsample results is whether the use of machine learning techniques provides value beyond more standard approaches such as least squares. The nonlinearity of the random forest FCI provides one insight that least squares cannot. Focusing on other dimensions, figure 11 compares an FCI constructed via least squares and the ML-FCI (from LASSO). The upper panel presents results for the full sample, and the lower panel for the restricted sample. Over the full sample, the least squares approach matches that of the ML approach. But these results are fragile, with the restricted sample results performing somewhat poorly out of sample (where poor performance, in this case, is defined as swings that do not match common views on the degree of tightening in financial conditions during the Global Financial Crisis and developments thereafter). These exercises highlight the value of using ML approaches to find robust relationships within the data to assess financial conditions without spuriously overfitting the researcher's objective.

<<INSERT FIGURE 11 HERE>>

5.2. Variable transformations and dimension reduction

A final set of considerations focuses on whether the focus on machine learning techniques provide variable beyond more commonly-used approaches to big data in macroeconomics that employ dimension reduction, such as dynamic factor models or principal component analysis (as discussed in, most prominently, Stock and Watson, 2002, 2011, and 2016).

Along one dimension, the machine learning techniques clearly point in a new direction: the random forest approach points to strong nonlinear relationships, consistent with macroeconomic theory and suggesting avenues for further exploration. Dynamic factor models are linear.

However, the insights gleaned from the analysis herein arguable go further. The machine learning techniques considered a set of alternative transformations for certain variables and demonstrated that these choices are consequential. For example, the default transformation for equity prices (and many other variables) is typically a simple change or simple percent change (log-difference), as captured in the suggestions of McCracken and Ng (2016). But the machine learning algorithm preferred a lower frequency transformation.

This insight suggests that a priori choices—sometimes made by default—may drive some results. For example, the NFCI is a dynamic factor model and it contains many measures of credit spreads and no measures of the level of interest rates. Taking the first principal component of the 33 financial variables used herein—a simple version of a dynamic factor model—yields figure 12. It is clear that this first principal component captures the level of real interest rates. This result is somewhat mechanical: the levels of real interest rates are highly correlated, and the data includes both levels and several spreads relative to the federal funds rate—implying that the federal funds rate is very important in the dataset used. But the mechanical relation emphasizes the need to transformations and variable selection to be purposeful and/or for techniques to be robust to such choices. For example, standard approaches that do not include the level of real interest rates—as in the default choices suggested by McCracken and Ng (2016)—would not even look at the information in the level of real interest rates.

<<INSERT FIGURE 12 HERE>>

These results highlight how choices regarding variables to include and transformations can drive results. Approaches that allow greater flexibility to consider alternative data transformations and let the data speak, as used herein, may be useful in future applications.

6. Conclusions

Assessments of financial conditions are important in policy discussions, forecasting exercises, and efforts to understand the economic mechanisms driving business and household decisions.

I used machine-learning techniques to construct a financial conditions index, selecting the components of the FCI based on their ability to predict the unemployment rate one-year ahead. The ML-FCI shows a tightening in financial conditions before the early 1990s and early 2000s recessions, in contrast to the Federal Reserve Bank of Chicago's National Financial Conditions Index. This finding owes, in part, to the tuning of the ML-FCI to predict unemployment. The ML-FCI places sizable weight on equity prices, in contrast to the NFCI. This occurs because the ML approaches select lower-frequency changes in equity prices as important predictors, whereas the dynamic factor model underlying the NFCI mechanically considers only high-frequency movements in equity prices.

These results highlight the importance of objectives and variable transformation in construction of FCIs, issues important more broadly in prediction with large datasets. Other research has similarly emphasized how *a priori* choices regarding data transformations can affect assessments of the links between financial conditions and economic outcomes: for example, Kiley (2018) illustrates the sensitivity of crisis prediction models to such considerations. The analysis herein highlights how machine learning and big data may allow researchers to consider a wider range of variables and transformations, thereby improving both the measurement and theory of economic and financial conditions. A related area that may benefit from this approach involves financial-stability assessments, which have often relied on approaches more similar to factor or principal-components analysis (e.g., Aikman et al, 2017; and Lee et al, 2018).

Finally, there is strong evidence of nonlinearity in the relationship between financial variables and economic activity from a random forest. Tight financial conditions are associated with sharp deteriorations in economic activity and accommodative conditions

are associated with only modest improvements in activity. These results highlight how nonlinear approaches in the machine learning literature are useful for macroeconomists and deserve further exploration (as in, for example, Pike et al, 2019). The results also emphasize the need for empirical macroeconomics to consider the nonlinearities implied by macroeconomic theories more thoroughly.

References

- Adrian, Tobias, Nina Boyarchenko, and Domenico Giannone. "Vulnerable growth." *American Economic Review* 109, no. 4 (2019): 1263-89.
- Aikman, David, Michael Kiley, Seung Jung Lee, Michael G. Palumbo, and Missaka Warusawitharana. "Mapping heat in the US financial system." *Journal of Banking & Finance* 81 (2017): 36-64.
- Athey, Susan. "The impact of machine learning on economics." In *The economics of artificial intelligence: An agenda*. University of Chicago Press, 2018.
- Barnichon, R., C. Matthes and A. Ziegenbein (forthcoming) "Are the Effects of Financial Market Disruptions Big or Small?" *Review of Economics and Statistics*.
- Bartsch, Elga. "A better gauge of financial conditions." BlackRock Investment Institute/Markets. Published December 12, 2018
- Boivin, J., Kiley, M. T., & Mishkin, F. S. (2010). How has the monetary transmission mechanism evolved over time?. In *Handbook of monetary economics* (Vol. 3, pp. 369-422). Elsevier.
- Bräuning, F., and S. J. Koopman, 2014, "Forecasting macroeconomic variables using collapsed dynamic factor analysis," *International Journal of Forecasting*, Vol. 30, No. 3, July–September, pp. 572–584.
- Brave, Scott A., and R. Butters. "Monitoring financial stability: A financial conditions index approach." *Economic Perspectives* 35, no. 1 (2011): 22.
- Brave, Scott A. & R. Andrew Butters & David Kelley, 2019. "A New "Big Data" Index of U.S. Economic Activity," *Economic Perspectives*, Federal Reserve Bank of Chicago, issue 1, pages 1-30.
- Breiman, Leo. "Random forests." *Machine learning* 45, no. 1 (2001): 5-32.
- Cogley, Timothy, and James M. Nason. "Effects of the Hodrick-Prescott filter on trend and difference stationary time series Implications for business cycle research." *Journal of Economic Dynamics and control* 19, no. 1-2 (1995): 253-278.
- Lawrence J. Christiano & Terry J. Fitzgerald, 2003. "The Band Pass Filter," *International Economic Review*, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, vol. 44(2), pages 435-465, May. <https://ideas.repec.org/a/ier/iecrev/v44y2003i2p435-465.html>

- Cogley T. and Nason J.M., 1995. Effects of the Hodrick-Prescott filter on trend and difference stationary time series implications for business cycle research. *Journal of Economic Dynamics and Control* 19: 253–278.
- De Mol, Christine, Domenico Giannone, and Lucrezia Reichlin. "Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components?." *Journal of Econometrics* 146, no. 2 (2008): 318-328.
- Diebold, Francis X., and Minchul Shin. "Machine learning for regularized survey forecast combination: Partially-egalitarian LASSO and its derivatives." *International Journal of Forecasting* 35, no. 4 (2019): 1679-1691.
- English, William, Kostas Tsatsaronis, and Edda Zoli. "Assessing the predictive power of measures of financial conditions for macroeconomic variables." *BIS Papers* 22 (2005): 228-252.
- Engstrom, Eric C., and Steven A. Sharpe. "The near-term forward yield spread as a leading indicator: A less distorted mirror." *Financial Analysts Journal* (2019).
- Estrella, Arturo, and Frederic S. Mishkin. "The yield curve as a predictor of US recessions." *Current issues in economics and finance* 2, no. 7 (1996).
- Federal Reserve Bank of Cleveland, Cleveland Financial Stress Index (DISCONTINUED) [CFSI], retrieved from FRED, Federal Reserve Bank of St. Louis; <https://fred.stlouisfed.org/series/CFSI>, October 18, 2019.
- Federal Reserve Bank of Kansas City, Kansas City Financial Stress Index [KCFSI], retrieved from FRED, Federal Reserve Bank of St. Louis; <https://fred.stlouisfed.org/series/KCFSI>, October 18, 2019.
- Federal Reserve Bank of St. Louis, St. Louis Fed Financial Stress Index [STLFSI], retrieved from FRED, Federal Reserve Bank of St. Louis; <https://fred.stlouisfed.org/series/STLFSI>, October 18, 2019.
- Forni, Mario, Marc Hallin, Marco Lippi, and Lucrezia Reichlin. "The generalized dynamic-factor model: Identification and estimation." *Review of Economics and statistics* 82, no. 4 (2000): 540-554.
- Gilchrist, Simon, and Egon Zakrajšek. "Credit spreads and business cycle fluctuations." *American Economic Review* 102, no. 4 (2012): 1692-1720.
- Guichard, Stéphanie, David Haugh, and David Turner. "Quantifying the effect of financial conditions in the euro area, Japan, United Kingdom and United States." (2009).

- Hamilton, James D. "Why you should never use the Hodrick-Prescott filter." *Review of Economics and Statistics* 100, no. 5 (2018): 831-843.
- Hatzius, Jan, Peter Hooper, Frederic S. Mishkin, Kermit L. Schoenholtz, and Mark W. Watson. *Financial conditions indexes: A fresh look after the financial crisis*. No. w16150. National Bureau of Economic Research, 2010.
- Hatzius, Jan, and Sven Jari Stehn. "The Case for a Financial Conditions Index," Published by Goldman Sachs Research on 16 July 2018
- Hodrick R.J., and E.C. Prescott, 1997. Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking* 29: 1–16.
- Hoerl, Arthur E., and Robert W. Kennard. "Ridge regression: Biased estimation for nonorthogonal problems." *Technometrics* 12, no. 1 (1970): 55-67.
- Holston, Kathryn, Thomas Laubach, and John C. Williams. "Measuring the natural rate of interest: International trends and determinants." *Journal of International Economics* 108 (2017): S59-S75.
- International Monetary Fund (IMF), 2019. *Global Financial Stability Reports*.
<https://www.imf.org/en/Publications/GFSR/>
- Kelly, Bryan, and Seth Pruitt. "The three-pass regression filter: A new approach to forecasting using many predictors." *Journal of Econometrics* 186, no. 2 (2015): 294-316.
- Kiley, Michael T. *What Macroeconomic Conditions Lead Financial Crises?*. No. 2018-038. Board of Governors of the Federal Reserve System (US), 2018.
- Kiley, Michael T. (2020a) *The Global Equilibrium Real Interest Rate: Concepts, Estimates, and Challenges*. *Annual Review of Financial Economics* 2020 12:1
- Kiley, M. T. (2020b). *What Can the Data Tell Us About the Equilibrium Real Interest Rate*. *International Journal of Central Banking*. Previously available as FEDS Working Paper, (2015-77)
- Koop, Gary, and Dimitris Korobilis. "A new index of financial conditions." *European Economic Review* 71 (2014): 101-116.
- Lee, Seung Jung, Kelly E. Posenau, and Viktors Stebunovs. "The Anatomy of financial vulnerabilities and banking crises." *Journal of Banking & Finance* (2018).
- Matheson, Troy D. "Financial conditions indexes for the United States and Euro area." *Economics Letters* 115, no. 3 (2012): 441-446.

McCracken, Michael W., and Serena Ng. "FRED-MD: A monthly database for macroeconomic research." *Journal of Business & Economic Statistics* 34, no. 4 (2016): 574-589.

Mullainathan, Sendhil, and Jann Spiess. "Machine learning: an applied econometric approach." *Journal of Economic Perspectives* 31, no. 2 (2017): 87-106.

Pike, Tyler, Horacio Saprizza, and Tom Zimmermann (2019). "Bottom-up leading macroeconomic indicators: An application to non-financial corporate defaults using machine learning," Finance and Economics Discussion Series 2019-070. Washington: Board of Governors of the Federal Reserve System, <https://doi.org/10.17016/FEDS.2019.070>.

Reichlin, L., C. De Mol, E. Gautier, D. Giannone, S. Mullainathan, H. van Dijk, and J. Wooldridge. "Big data in economics: evolution or revolution?." (2017): 612-632.

Rudebusch, Glenn D., and John C. Williams. "Forecasting recessions: the puzzle of the enduring power of the yield curve." *Journal of Business & Economic Statistics* 27, no. 4 (2009): 492-503.

Stock, James H., and Mark W. Watson. "New indexes of coincident and leading economic indicators." *NBER macroeconomics annual* 4 (1989): 351-394.

Stock, James H., and Mark W. Watson. "Forecasting using principal components from a large number of predictors." *Journal of the American statistical association* 97, no. 460 (2002): 1167-1179.

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

Stock, James H., and Mark W. Watson. "Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics." In *Handbook of macroeconomics*, vol. 2, pp. 415-525. Elsevier, 2016.

Varian, Hal R. "Big data: New tricks for econometrics." *Journal of Economic Perspectives* 28, no. 2 (2014): 3-28.

Table 1: Variables Used in the Analysis

Description	FRED mnemonic	Baseline transformation	High-frequency transformation	Transformation in FRED-MD
Interest rates				
Effective Federal Funds Rate	FEDFUNDS	$x_t - \Delta_{12} \ln(p_t)$.	$\Delta(x_t)$
3-Month AA Financial Commercial Paper Rate	CP3Mx	$x_t - \Delta_{12} \ln(p_t)$.	$\Delta(x_t)$
3-Month Treasury Bill:	TB3MS	$x_t - \Delta_{12} \ln(p_t)$.	$\Delta(x_t)$
6-Month Treasury Bill:	TB6MS	$x_t - \Delta_{12} \ln(p_t)$.	$\Delta(x_t)$
1-Year Treasury Rate	GS1	$x_t - \Delta_{12} \ln(p_t)$.	$\Delta(x_t)$
5-Year Treasury Rate	GS5	$x_t - \Delta_{12} \ln(p_t)$.	$\Delta(x_t)$
10-Year Treasury Rate	GS10	$x_t - \Delta_{12} \ln(p_t)$.	$\Delta(x_t)$
Moody's Seasoned Aaa Corporate Bond Yield	AAA	$x_t - \Delta_{12} \ln(p_t)$.	$\Delta(x_t)$
Moody's Seasoned Baa Corporate Bond Yield	BAA	$x_t - \Delta_{12} \ln(p_t)$.	$\Delta(x_t)$
Term spreads				
3-Month Treasury C Minus FEDFUNDS	TB3SMFFM	x_t	.	x_t
6-Month Treasury C Minus FEDFUNDS	TB6SMFFM	x_t	.	x_t
1-Year Treasury C Minus FEDFUNDS	T1YFFM	x_t	.	x_t
5-Year Treasury C Minus FEDFUNDS	T5YFFM	x_t	.	x_t
10-Year Treasury C Minus FEDFUNDS	T10YFFM	x_t	.	x_t
Credit spreads				
3-Month Commercial Paper Minus FEDFUNDS	COMPAPFFx	x_t	.	x_t
Moody's Aaa Corporate Bond Minus GS10	AAAFFM-T10YFFM	x_t	.	x_t
Moody's Baa Corporate Bond Minus GS10	BAAFFM-T10YFFM	x_t	.	x_t
Exchange rates				
Switzerland / U.S. Foreign Exchange Rate	EXSZUSx	$\Delta_{12} \ln(x_t)$	$\Delta \ln(x_t)$	$\Delta \ln(x_t)$
Japan / U.S. Foreign Exchange Rate	EXJPUSx	$\Delta_{12} \ln(x_t)$	$\Delta \ln(x_t)$	$\Delta \ln(x_t)$
U.S. / U.K. Foreign Exchange Rate	EXUSUKx	$\Delta_{12} \ln(x_t)$	$\Delta \ln(x_t)$	$\Delta \ln(x_t)$
Canada / U.S. Foreign Exchange Rate	EXCAUSx	$\Delta_{12} \ln(x_t)$	$\Delta \ln(x_t)$	$\Delta \ln(x_t)$
Equity prices				
S&P's Common Stock Price Index: Composite	S&P 500	$\Delta_{12} \ln(x_t)$	$\Delta \ln(x_t)$	$\Delta \ln(x_t)$
S&P's Common Stock Price Index: Industrials	S&P: indust	$\Delta_{12} \ln(x_t)$	$\Delta \ln(x_t)$	$\Delta \ln(x_t)$
S&P's Composite Common Stock: Dividend Yield	S&P div yield	$\Delta_{12} \ln(x_t)$	$\Delta(x_t)$	$\Delta(x_t)$
S&P's Composite Common Stock: Price-Earnings Ratio	S&P PE ratio	$\Delta_{12} \ln(x_t)$	$\Delta \ln(x_t)$	$\Delta \ln(x_t)$
Other variables used in analysis				
Civilian Unemployment Rate	UNRATE	x_t	.	$\Delta(x_t)$
CPI : All Items	CPIAUCSL	$\Delta_{12} \ln(x_t)$.	$\Delta \Delta(x_t)$

The FRED mnemonic column gives mnemonics in the Federal Reserve Bank of St. Louis's FRED database. In the transformation columns "Δ" refers to the change from the previous month, and "Δ₁₂" refers to the change from 12 months earlier. For additional details on the data and the construction of the FRED-MD database, see the working paper version of McCracken and Ng (2017), [https://research.stlouisfed.org/wp/2015-012.pdf](https://research.stlouisfed.org/wp/2015/2015-012.pdf).

Figure 1

Growing a Random Forest

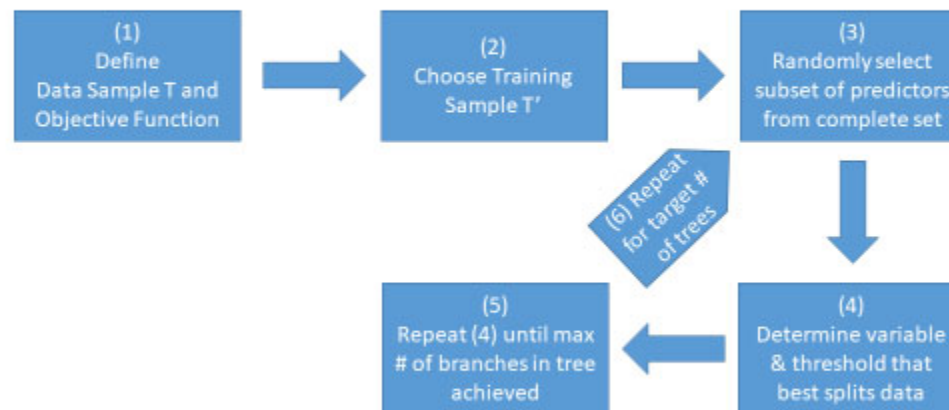
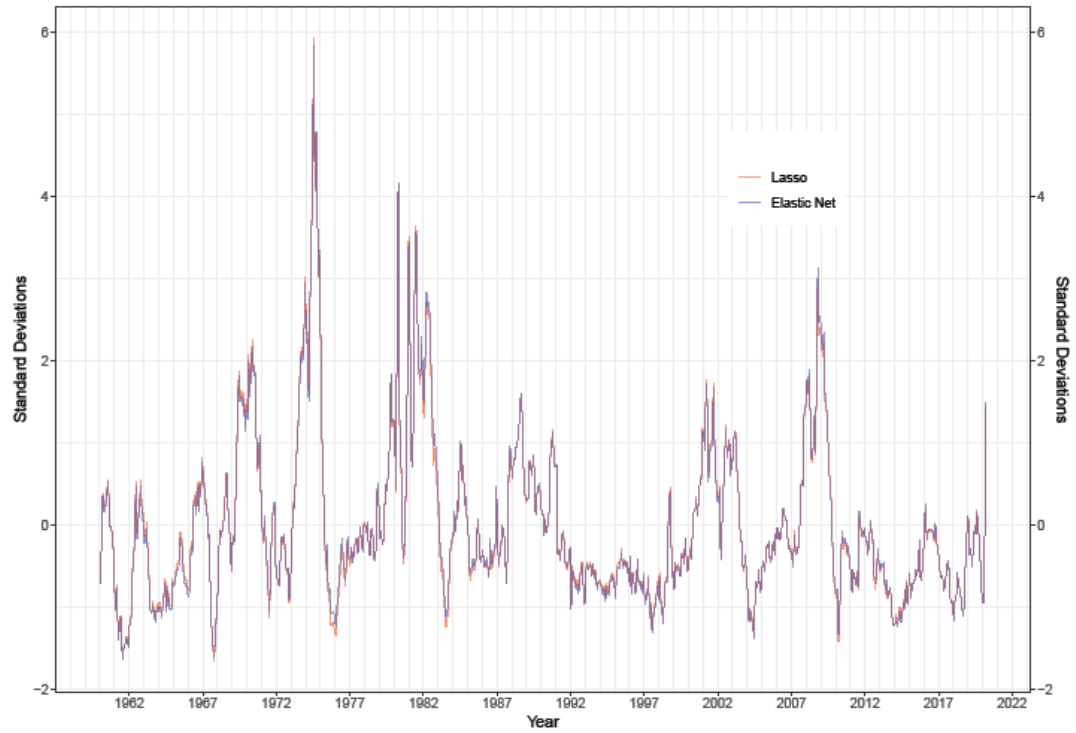
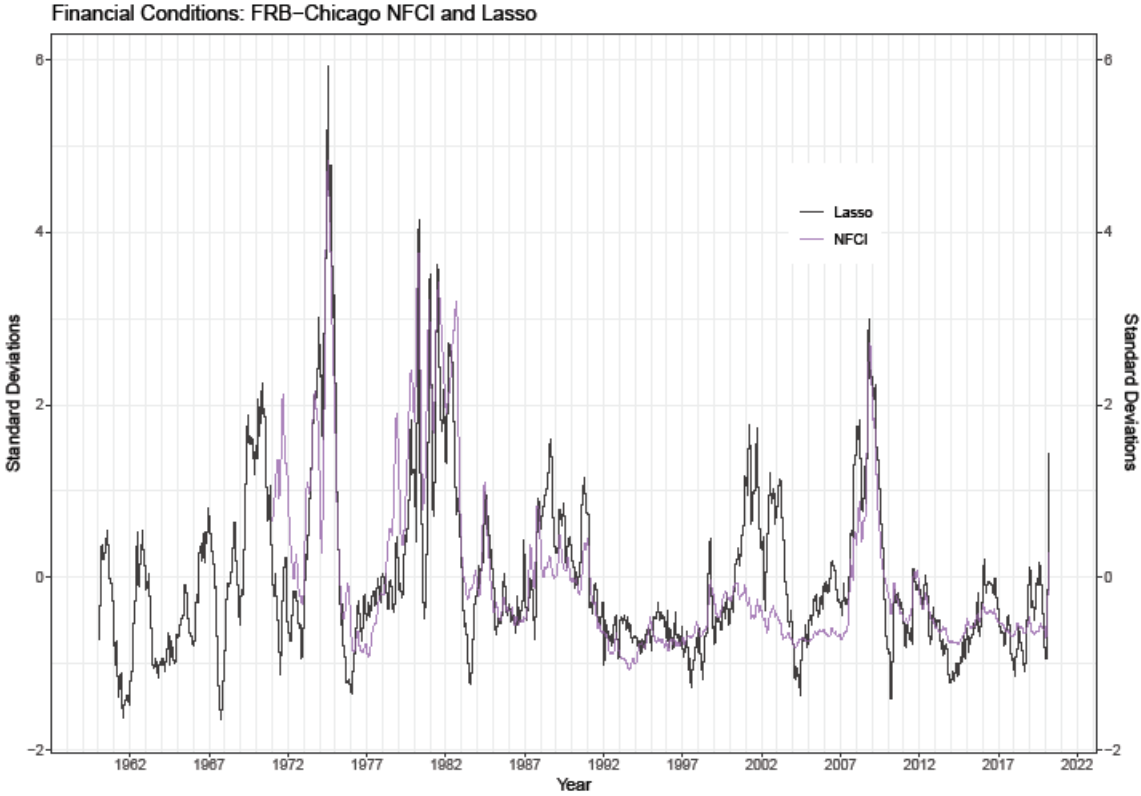


Figure 2: Financial Conditions Indexes from LASSO and Elastic Net



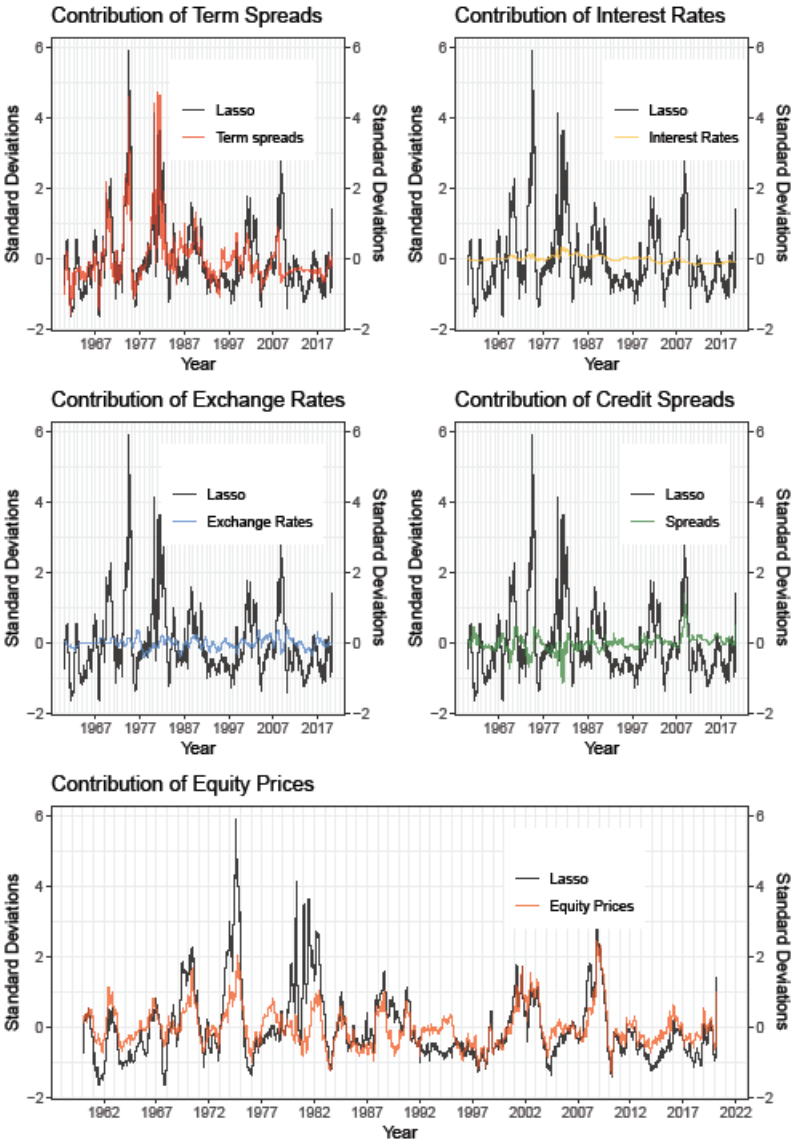
Source: Author's calculations

Figure 3: Financial Conditions Indexes from LASSO and FRB-Chicago NFCI



Source: Author’s calculations and Federal Reserve Bank of St. Louis FRED.

Figure 4: Financial Conditions Indexes from LASSO and Contributions



Source: Author's calculations

Figure 5

An example tree

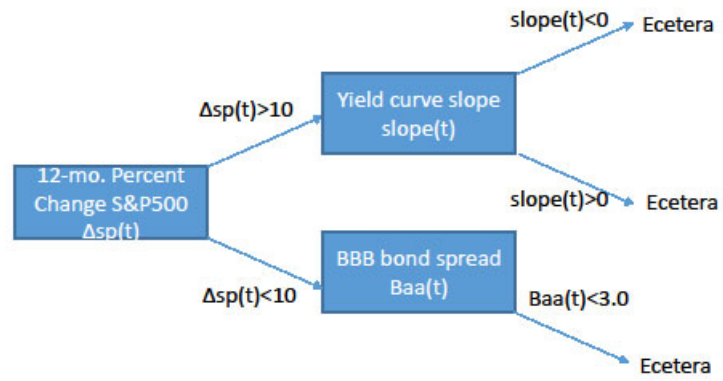
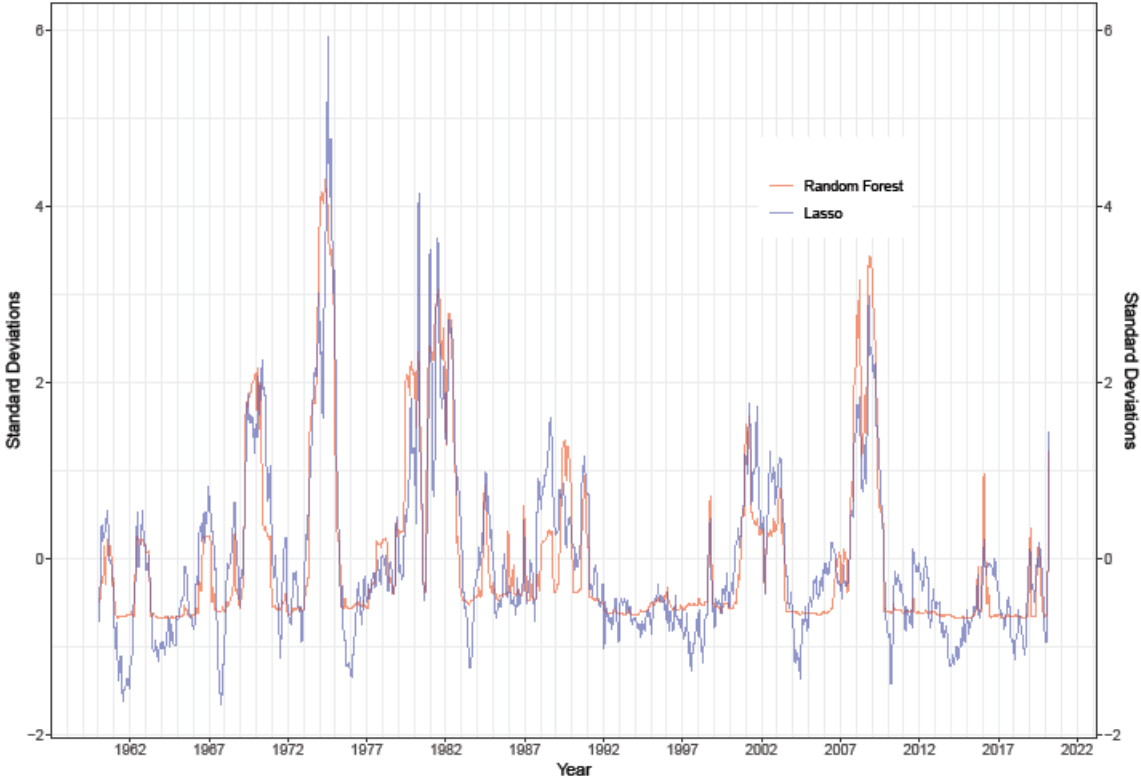
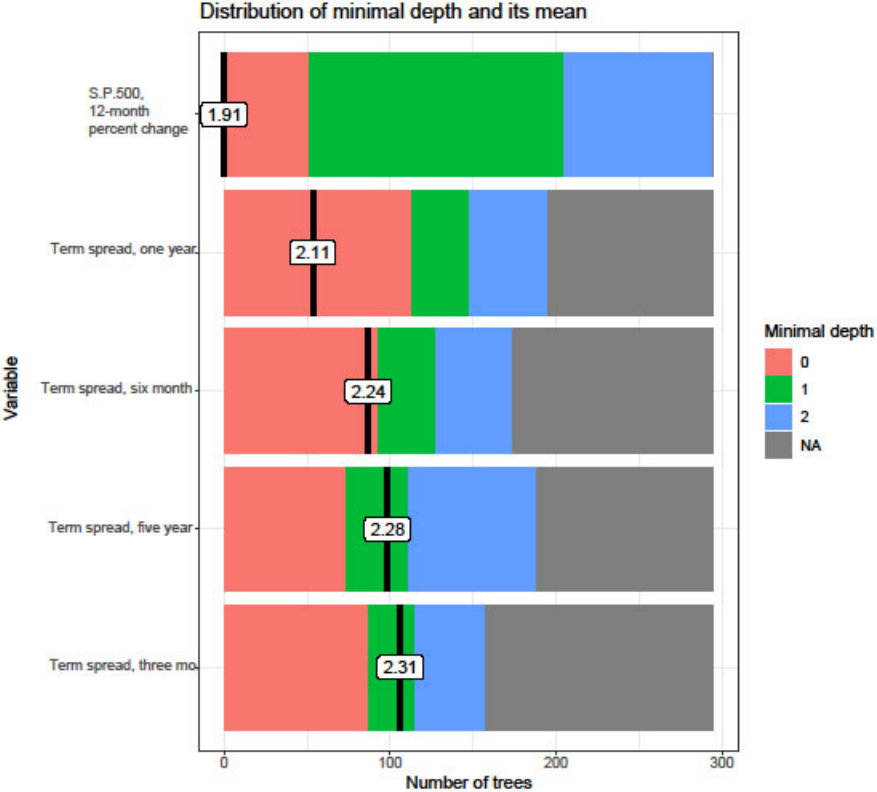


Figure 6: Random Forest FCI vs LASSO FCI



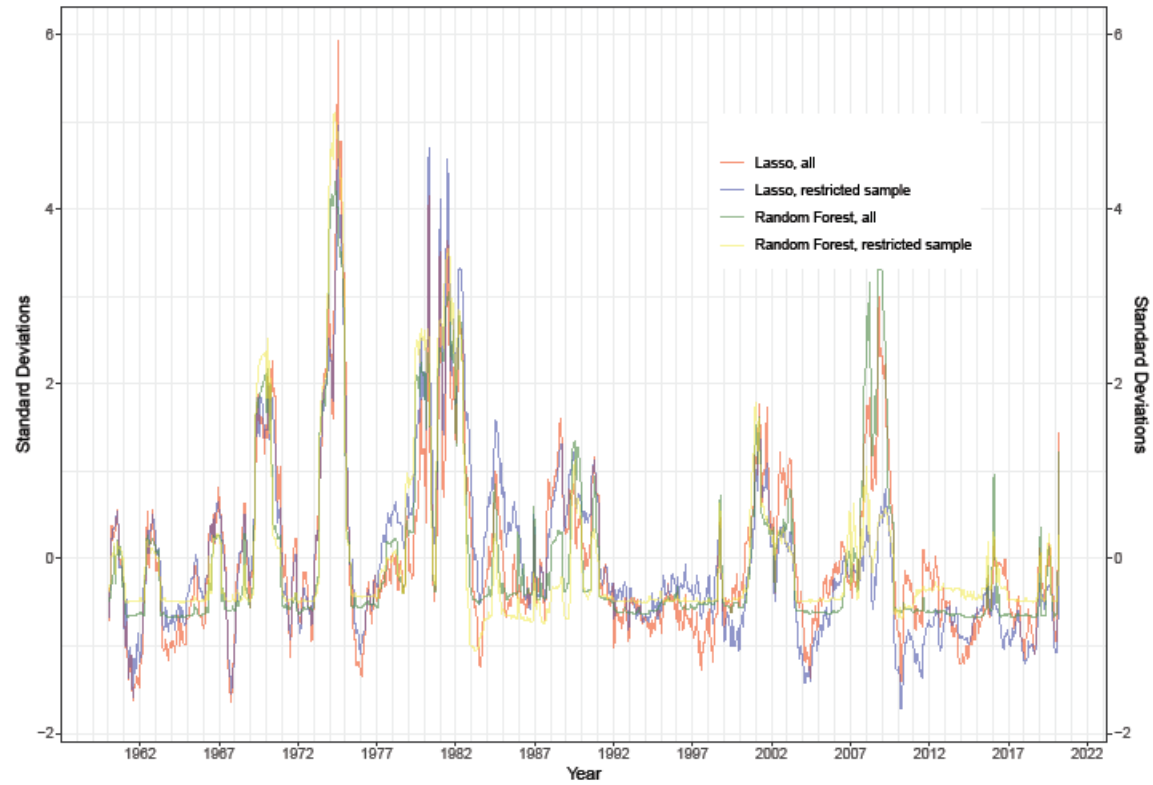
Source: Author's calculations

Figure 7: Random Forest Results on Variable Selection



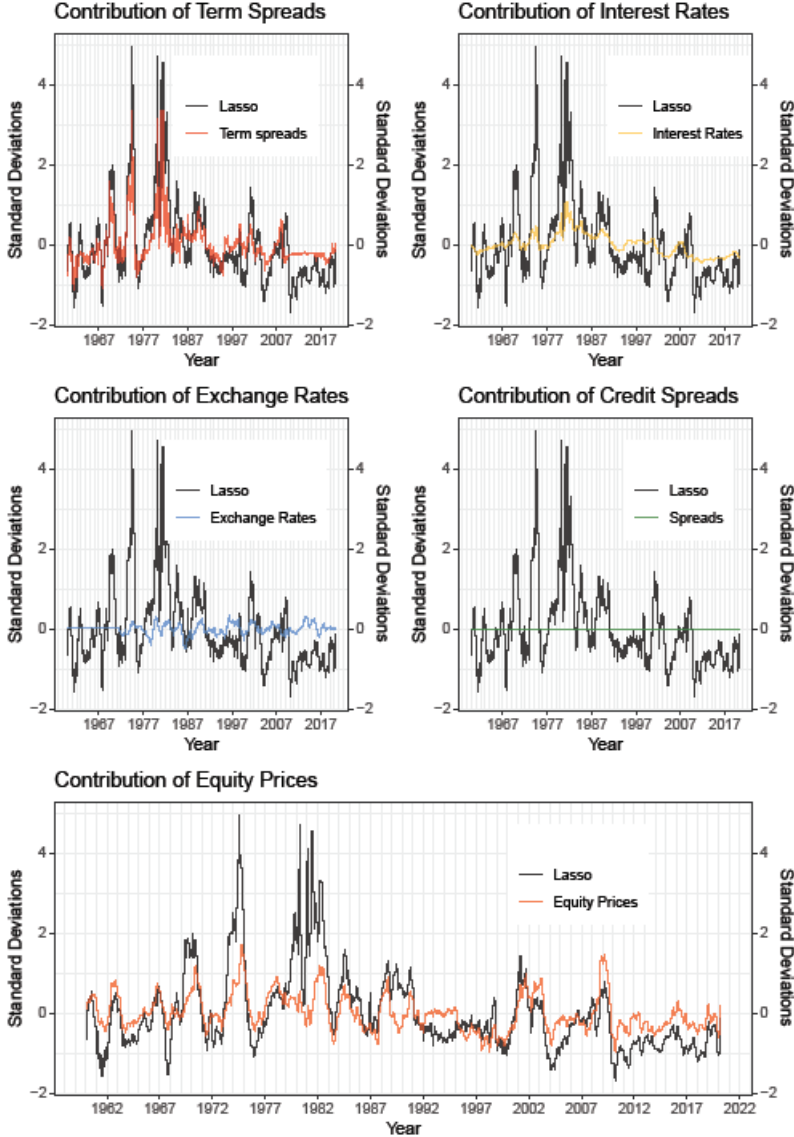
Source: Author’s calculations

Figure 8: ML-FCI from LASSO and Random Forest: Full Sample vs Restricted Sample (ending December 2004)



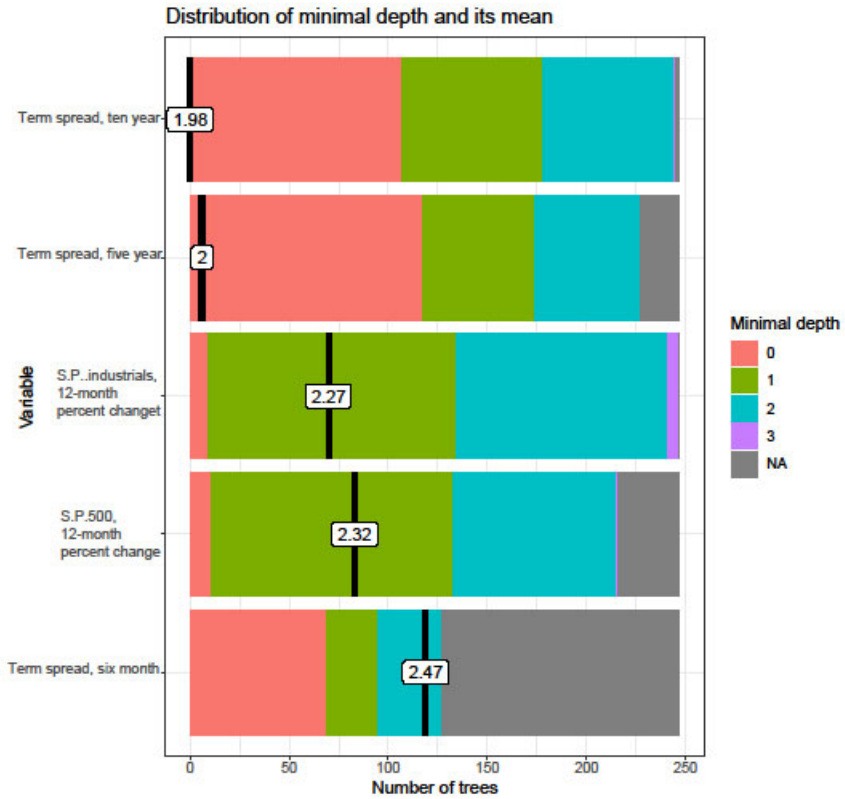
Source: Author's calculations

Figure 9: ML-FCI from LASSO, Contributions in restricted sample (ending December 2004)



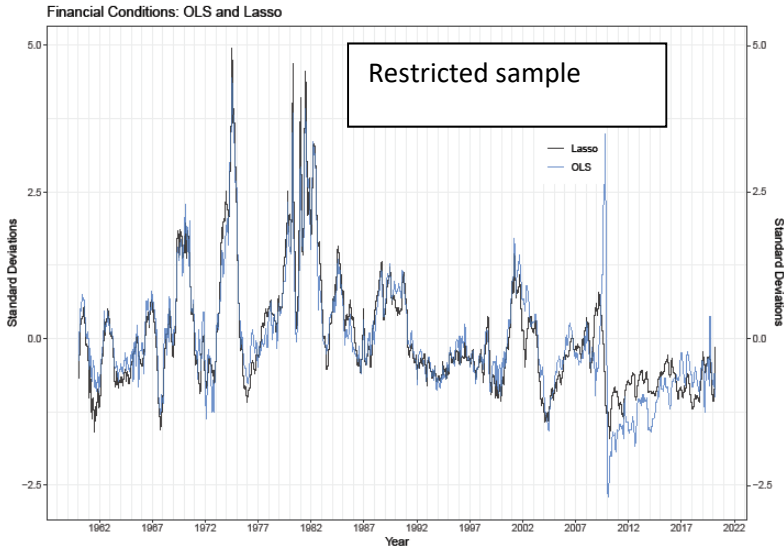
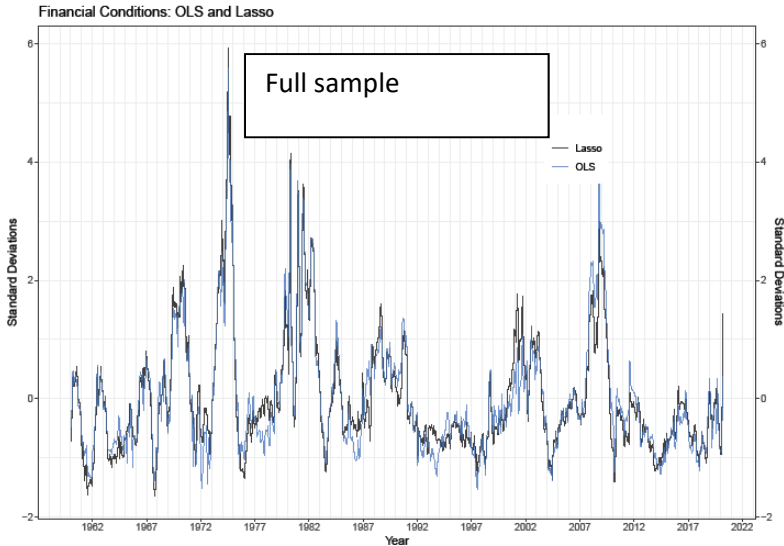
Source: Author’s calculations

Figure 10: Random Forest Results on Variable Selection: Restricted Sample (ending December 2004)



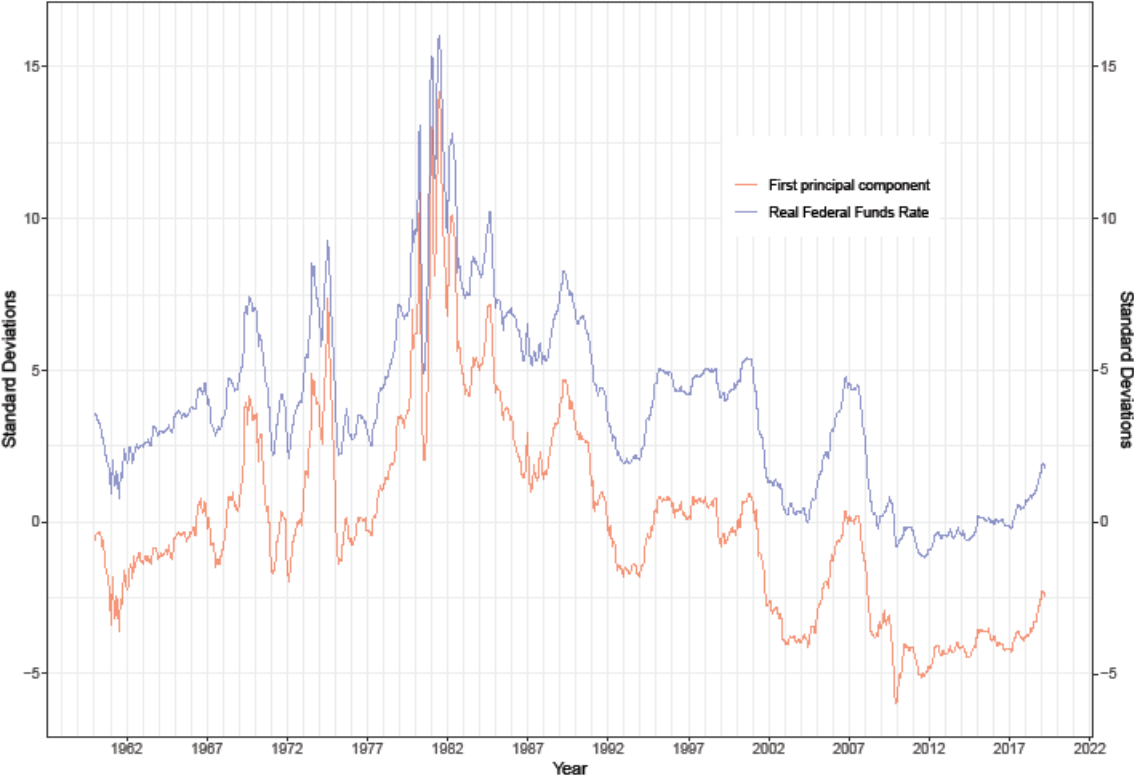
Source: Author's calculations

Figure 11: ML-FCI vs Least Squares FCI: Full Sample and Restricted Sample



Source: Author's calculations

Figure 12: First Principal Component of Financial Variables and Real Federal Funds Rate



Source: Author's calculations