Shapes and Transitions of the Interest Rate Term Structure*

Biwei Chen†

Graduate Center, The City University of New York, 365 5th Ave, NY 10016, USA

October, 2017

Abstract

This paper classifies five shapes of the U.S. Treasury yield curve and examines their links to the macroeconomy. The analysis shows that these shapes have strong linkages to the different states of inflation and real production growth, and variations of the shapes have reliably predicted subsequent recessions. By casting the yield curve shape transition into Markov chains, the paper estimates the transition probability from one yield curve shape to another, and finds strong momentum and asymmetry in the shape migration. The Markov model can serve as a basis for predicting future yield curve movements and business cycle transitions.

JEL classification: E32, E43, E47, G12, C52, C53, C82.

Key words: Business cycles; Macroeconomic states; Markov chain model estimation, evaluation, and forecast; K-fold cross-validation; U.S. Treasury yield curve.

*I am grateful to Prof. Liuren Wu for his intellectual inspiration and guidance shaping the paper. I thank Prof. Wim Vijverberg, Prof. George Vachadze, Prof. Chu-Ping Vijverberg, Prof. Merih Uctum, Prof. Chun Wang, and Prof. Thom Thurston for valuable comments and discussions. All errors and omissions are my own responsibility.

†Email: bchen@gradcenter.cuny.edu

Contents

List of Tables

List of Figures

1 Introduction

Ever since the pioneering study of Kessel (1965) on the term structure patterns over the U.S. business cycles, scholars have enthusiastically investigated the information contained in the term structure for macroeconomic predictions. A vast literature has shown that the term spread, or the slope of the yield curve, is useful for predicting future economic activity, such as real GDP growth, inflation, and business cycle fluctuation.[1] To explore more information beyond the term spread, recent studies have begun to extract unobservable factors (level, slope, and curvature) from the term structure data and found improved forecasting performance for real output growth and business cycle timing.[2]

While the predictive power of the yield curve has become a "stylized fact" among macroeconomists and investors, it only represents a one-way information flow from the yield curve to the macroeconomy. Several studies also consider responses of the yield curve to the macroeconomy and the connections between the two.[3] In this literature, the yield curve level and slope factors are found to relate to inflation and real output variables, respectively, what affects the curvature factor remains uncertain (Diebold et al. 2006, 2013). More fundamentally, the macro-finance factor models do not provide enough intuition on why the yield curve can exhibit certain shapes, what these shapes mean about the macroeconomy, and how they evolve over time. This paper addresses these questions surrounding yield curve shapes.

The contribution of this paper to the term structure literature is two-fold. The first is the exploration of the links between various yield curve shapes and the macroeconomy. By directly focusing on the shape of the yield curve, we adopt a new angle to examine the old topic of the term structure. Second, this paper models and estimates the shape transition process of the term structure. Traditionally, the transition dynamics of the yield curve

[1] Harvey (1988, 1989, 1991), Mishkin (1990a,b), Estrella and Hardouvelis (1991), Frankel and Lown (1994), Plosser and Rouwenhorst (1994), Estrella and Mishkin (1996, 1997, 1998), Kamara (1997), Roma and Torous (1997), Hamilton and Kim (2002), Stock and Watson (2003), Bordo and Haubrich (2008), Rudebusch and Williams (2009); See Wheelock and Wohar (2009) for an comprehensive literature survey.

[2] Monech (2012), Abdymomunov (2013), Argyropoulos and Tzavalis (2016), Chauvet and Senyuz (2016)

[3] Ang and Piazzesi (2003), Diebold et al. (2006), Dewachter et al. (2006), Ang et al. (2007), Rudebusch and Wu (2008), Lu and Wu (2009), Bibkov and Chernov (2010), Lange (2013), Chauvet and Senyuz (2016)

have been ignored in term structure studies. Our estimation results can shed light on the underlying transition patterns.

Some key findings are as follows. First, the classified shape of the U.S. Treasury yield curve has the following frequency distribution: The upward yield curve is the most typical and it accounts for 72.22% of the total sample observations; the hump shape and downward yield curve are observed much less frequently, each occurring about 10% of the time in the sample; and the bowl shape and flat yield curves are observed least frequently—only 5.29% and 2.12% of the time, respectively.[4]

Second, in terms of their ability to predict U.S. recessions in the past 60 years, the four less frequent shapes of the yield curve are very informative. This result contrasts with the over-emphasis on the downward yield curve in the literature. Moreover, after the 1982 recession, their joint predictive power becomes more impressive: They occur only before each recession, not during and after.

Third, each yield curve shape may correspond to certain economic states: If an upward sloping yield curve is set as a benchmark state of low inflation rate and mediocre real output growth rate, the steep upward yield curve would correspond to the lowest level of inflation rate and real output growth rate among all states. The flat yield curve may match a state of a moderate inflation rate and the highest output growth rate. The downward yield curve may reflect a state of the highest inflation rate and a mediocre output growth rate. The hump and bowl shape yield curves may indicate some states in between: Relatively high inflation and low output growth.

The estimation for the shape transition process shows that it displays significant momentum: The yield curve is more likely to retain its shape than changing shapes. In addition, the transition is asymmetric: The yield curve is more likely to transition to adjacent (e.g., from the upward to hump shape curves) than non-adjacent shapes (e.g., from the upward to flat, bowl, or downward curves). Four types of transitions never happen: From upward to

[4]This result is obtained from the monthly Treasury yield data, April 1953 to March 2016. Similar frequency distribution is obtained from the daily and weekly data.

downward, from bowl to hump, from downward to upward, and from hump to bowl.

2 Data and methodology

The empirical examination of this study combines three monthly datasets from different sources. They cover a period from April 1953 to March 2016. The first set is the U.S. Treasury yield data that can be found from the Federal Reserve Board H.15 interest rate statistics. It includes nominal yields on eleven Treasury securities of different maturities. The second set is the chronology of the U.S. business cycle, which is maintained by the Business Cycle Dating Committee of the National Bureau of Economic Research (NBER). It identifies the dates of peaks and troughs that define economic recession or expansion. The third set, which can be downloaded from the Federal Reserve Economic Data (FRED-St. Louis Fed), merges macroeconomic state variables. It includes four measures for inflation rate and four measures for real production activity. All these datasets are available from the corresponding institution's website.

The innovation of this study starts from a classification algorithm that identifies five patterns of yield curves. With the first dataset, we reduce the dimension of Treasury yields to three averaged yields (short, median, and long) and define the yield curve shapes from their relations. The advantage of this data-driven, algorithm-based classification methodology is tremendous: It enables quantitative comparisons among different shapes of the yield curve and generates a categorical sequence for statistical modeling.

To explore the links between yield curve shapes and macroeconomic states, we utilize the second and third dataset. The NBER business cycle chronology enables us to identify recession-related periods in which we document the timing and frequencies of each shape of the yield curve. The merged FRED macroeconomic dataset allows us to associate each shape of the yield curve with the production and inflation states. We model the term structure transition dynamics in Markov chains and estimate the shape transition probabilities via maximum likelihood methods. We consider Markov chains of alternative orders and test their predictive performance via cross-validation.

3 Classification of yield curve shapes

This section introduces the basic concepts and classification method for the U.S. Treasury yield curves. Details are in Appendix A.[5]

The central concept in term structure analysis is the yield to maturity, henceforth YTM. As the most widely referenced interest return on a bond, the YTM measure implies an average rate of return if the investors hold the bond to maturity. This implied interest rate is "backed out" from the market price of the corresponding security.

In theory, the YTM satisfies the bond pricing equation:

$$P = \sum_{t=1}^{T} \frac{CF_t}{(1+Y_t)^t} \quad or \quad P = \sum_{t=1}^{T} CF_t * e^{-Y_t t} \tag{3.1}$$

where P is the market price of the security, Y_t is the YTM corresponding to the maturity t cash flow, CF_t is the cash flow payment at time t. The two different versions merely represent discrete or continuous compounding methods.

All else equal (risk factors such as creditworthiness, liquidity, callable features, tax treatment, coupon payment schemes), relating yield Y on a security to its maturity T constitutes the term structure analysis. In this paper, we analyze yields on the U.S. Treasury securities: At the short end, Treasury bills are money market assets with maturities of one year or less, sold at a discount from par and do not bear periodic interest payments. Treasury notes are median-term coupon securities with maturities from two to ten years. Treasury bonds have maturities of more than ten years. Treasury notes and bonds are capital market assets carrying periodic coupon payments. Yields of all these securities are therefore regarded as short-term, median-term, and long-term rates, respectively. By market convention, a yield

[5] Available upon request.

curve is derived by plotting the cross-sectional yields against their corresponding maturities.

The observed Treasury yield curve varies its shapes from time to time, reflecting particular conditions in financial markets and general states in the economy. If significant links between various shapes of the yield curve and unobservable economic states exist, it is essential to distinguish the types of yield curve first, then to model the changing states of the economy.

The classification algorithm is summarized in Table 1. First, the algorithm reduces data dimension from 11 U.S. Treasury yields to three averaged yields—Y_s, Y_m, Y_l, which are the simple average of Treasury bill, note, and bond yields, respectively. Second, it sets the flat yield curve as a benchmark with a threshold value of ten basis points for the differences among averaged yields. Then, it defines other shapes of the yield curve based on the slope and curvature relations. While the classification of the flat, upward, and downward yield curves allows weak mathematical inequalities among yields (Y_s, Y_m, Y_l), the yield relations of hump and bowl are strictly unequal. A hump shape yield curve is high in the middle but low on both ends; a bowl shape yield curve is low in the middle but high on both ends. The algorithm is effective in generating a sequence of mutually exclusive and exhaustive yield curve shapes.

[Table 1 about here.]

The classified yield curve statistics are shown in Table 2. Sample observations of the classified shapes are shown in Figure 1. In a descending order of their frequencies, we observe the upward yield curve (U 72.22%), the hump yield curve (H 10.32%), the downward yield curve (D 10.05%), the bowl yield curve (B 5.29%), and the flat yield curve (F 2.12%). From the table, we also observe that yield curves not only distinguish themselves by their shapes but also by their average yield levels. A detailed comparison among these shapes is in the Appendix A4.

In the following section, we will explore the links between classified yield curve types and general economic states. Readers who are more interested in modeling shape transition may skip to Section 5 without any loss of continuity.

[Table 2 about here.]

4 A tale of shapes and states

Why do we observe various shapes of the yield curve in the U.S. Treasury securities market? The expectation hypotheses argue that the shape of the yield curve is determined by the expectation of future interest rates and a risk premium.[6] Measurement and statistical issues aside, the empirical evidence is quite mixed and even confusing. There is as much evidence rejecting the hypotheses as supporting them (e.g., Bekaert and Hodrick, 2001). More fundamentally, the expectation theories and tests provide few meaningful answers to the question. In order to understand the driving forces behind the term structure, the behavior of the yield curve is our focus of attention. The classified shapes of the yield curve need to be analyzed within a standard economic context.

In this section, we examine how yields vary over different stages of the U.S. business cycles, investigate the changing patterns of yield curve shapes, and map the shapes onto their corresponding inflation and production states. These steps serve to establish potential links between yield curve shapes and the macroeconomy.

4.1 Links to business cycles

Figure 2 shows the U.S. Treasury yield co-movements for the period from April 1953 to March 2016. Among the yields, three level measures (average short, median, and long yields) are very persistent with upward trends before the fall of 1981 and downward trends thereafter. The yield spread measure fluctuates around zero and remains positive most of the time. The long swing in interest rates is often referred to as the property of slow mean reversion. Fama (2006) argues that this phenomenon is due to two permanent expected inflation shocks. The

[6] Three versions of the hypotheses inliterature are the Liquidity Premium Theory by Lutz (1940) and Hicks (1946), the Market Segmentation Theory by Culbertson (1957), and the Preferred Habitat Theory by Modigliani and Sutch (1966).

first positive one arises from the Federal Reserve's minimal prior experience in managing a fiduciary currency. The succeeding negative shock owes to Federal Reserve's strong monetary stance and drastic policies that bring down long-term inflation expectation.

[Figure 2 about here.]

To detect yield patterns over business cycle fluctuations, we enhance the approach of Kessel (1965, p. 65) by dividing the business cycles into four stages and examine how seven yield measures vary across these stages.[7] Since the monthly yield data is available from April 1953 and its subsequent recession starts in July 1953, we start our analysis from July 1953 and end it at the most recent 2007–09 recession. The NBER chronology helps us segment the entire business cycles into pre-recession, recession, post-recession, and other stages.[8] To better capture the synchronous effects of yields over different stages, we focus on an 18-month window preceding recessions, and a 12-month window following recessions. The choice of 18 and 12 is similar to existing findings on the predictive power of yield spreads—negative yield spreads tend to precede recessions and yield spreads widen immediately after recessions.[9]

The key statistics of the average monthly yield measures are reported in Table 3. These statistics serve to inform us of the variation in yield measures associated with different stages of the business cycle. The yield patterns are striking: On average, yield levels hit their plateau during the 18-month periods before recession, remain at a high level during the recessions, and enter a trough in the post-recession 12-month periods. Yield spreads all turn negative piror to a recession, turn back to positive in a recession, and widen in the post-recession 12-month periods. Correspondingly, we observe the downward yield curve most often during the pre-recession periods and the upward yield curve most often in other periods.

[7] In the study, Kessel segmented business cycles into peaks and troughs, and calculated changes in average yields in each trough-to-peak and peak-to-trough segment. Examining government obligations in the U.S. 1945 to 1961 business cycle, Kessel found that the peaks and troughs in yields were roughly synchronous with those in business activity.

[8] For defining the business cycles, refer to the http://www.nber.org/cycles/main.html.

[9] Estrella and Trubin (2006) tabulated term spread signals 12 months before each recession since 1968; Estrella and Mishkin (1996; 1997) find that term spread outperforms other financial and other macroeconomic indicators in predicting recessions two to six quarters ahead; Estrella and Hardouvelis (1991) find robust significance of term spread in real output prediction up to two years ahead.

It is also noted that, on average, yield spreads are largest during the post-recession recovery periods, equivalent to a steep upward yield curve. Compared with the recession-related three stages, yield levels are more volatile during the non-recession periods, as measured by standard deviation. In particular, volatility of all yields increases before recessions, becomes most volatile during recession periods, and then decreases after recessions. The yield curve tends to increase its concavity (hump) as the business cycle moves from pre-recession to post-recession.

[Table 3 about here.]

To further investigate the links between each type of yield curve and the business cycles, we document the frequency of each yield curve type in different stages over the ten U.S. business cycles from 1953 to 2010. Extracting the signals this way will help predict recessions from the yield curve shapes.

In Table 4, the occurrence of each yield curve is numbered for three stages in each cycle. During the 18 months preceding each recession, the frequency of the upward (U) curve shrinks by more than 50% compared with its full sample frequency. On the contrary, all four "minor" types of yield curve become significantly more prominent—their relative frequency more than doubled compared with their overall occurrence frequency. More interestingly, the bowl (B) yield curve predominates before the 2007 financial crisis.

[Table 4 about here.]

During the recessions, though the signal weakens, the H and B curves still occur more often than their full sample counterparts. The relative frequency of the D curve is close to its overall sample frequency. However, the F curve disappears in recession periods. Extraordinarily, the B curve occurs 12 times in the 1973 to 1975 recession.

In the post-recession 12-month periods, the U curve carries some good news about economic recovery. The F and B curves clearly are absent. The H curve signal further weakens, but the ten D curve counts is quite unusual after the 1980 recession. In fact, if we adjust for the overlap in between the 1980 and 1981 recessions, the ten D curve counts can be considered as strong signals before the 1981 recession. Similarly, if we extend the prediction horizon to more than 18 months, the two and four counts of the H curve might also be regarded as pre-recession signals. Therefore, after adjusting for the recession overlapping and timing, all four "minor" types of yield curve "disappear" right after recessions.

[Table 5 about here.]

Given the frequency count for each shape over different stages of the business cycle, it is also informative to calculate the probability of the economy being in a business cycle stage in the future given an observed yield curve shape today. For the sample period from April 1953 to March 2016, Table 5 panel A computes these k-month ahead conditional probabilities. Conditional on the U curve, the upcoming four months are more likely to be in a non-recession stage; however, conditional on the four "minor" shapes, the next four months are most likely to be in a pre-recession stage. Extending the in-sample forecast horizon to five to six quarters, conditional on the four current "minor" shapes, the business cycle stage is most likely to be in recession.

Panel B further computes the conditional probability of a recession starting in a future month. The D curve is most likely to precede a recession by three to four months. The B and H are most likely to precede a recession by five quarters. The F curve is likely to precede a recession by six to fourteen months. The last row in the table computes the cumulative probabilities of a recession starting within the next six quarters, conditional on each shape. Given the four "minor" shapes, the cumulative probabilities that a recession starts within future six quarters are quite high—45% to 70%, whereas the corresponding recession probability conditional on the U curve is only 7.1%. Overall, the four "minor" shapes do carry significant information in forecasting upcoming recessions in the sample. An evident change in the shape pattern is discovered after the 1981 to 1982 recession: Compared with their prior performance, "minor" types never send "wrong" signals in any recession and

12-month post-recession periods. This fact makes their predictive power more promising in the future.

Related to yield curve shapes, a common explanation for the ability of the term spread to forecast recessions is the Fed's monetary policy. The logic is as follows: Most of the time, monetary policies alter interest rates at the short end but the long rates are relatively fixed. In and after recessions, when the Fed cuts interest rates, we observe steep upward yield curve more often. During the booms (pre-recession periods), interest rate levels are often high because interest rates are pro-cyclical and the Fed tends to raise rates to halt inflation. Hence, yield curves tend to flatten or slope downward before recessions. However, the monetary policy explanation is usually stated with little theory (Wheelock and Wohar, 2009). Moreover, the explanation does not account for the bowl and hump yield curves, not to mention the time-varying predictive power of the shapes.

4.2 Links to economic states

As a general barometer of macroeconomic condition, a business cycle study aggregates real output variables into a single recession index. It provides a useful tool to understanding macroeconomic phenomena. However, we could miss the elephant in the room if our attention only centers around the business cycle and real output.

The other pertinent economic state variable that has been well identified as a driving force of the level of yield curve is inflation. Thus, we two economy-wide state variables that are closely linked to yield curve dynamics. Essentially, we need to understand whether the five types of yield curve are tied to the two states and in what manner. Our dividing unit of analysis is no longer business cycle stages but rather the shapes of the yield curve. In our investigation on the yield patterns over the business cycle, yield spreads tend to widen in each post-recession period. For a more effective comparison, we will divide upward yield curves into steep and usual ones (a cutoff value of 200 basis points average long-short spread is chosen). In this examination, if inflation and output state variables display significant differences among various yield curve shapes, these efforts may be effective to establish some links.

A simple approach is adopted to track major output and inflation state variables. The most relevant macroeconomic indicators are selected, which cover the period of April 1953 to March 2016. The four inflation variables are consumer price index rate for all urban consumers and all items (CPIA), consumer price index rate less food and energy (CPIC), producer price index rate for finished goods (PPIA), and producer price index rate for capital equipment (PPIC). Four output variables are real GDP growth rate (RGDP), unemployment rate (UNEM), industrial production index rate (INPR), and capacity utilization rate (CUR). Due to data availability, CUR starts from January 1968 and CPIC starts from January 1958. The data are downloaded from FRED - St. Louis, and are seasonally adjusted. The index rate is calculated as percentage growth from 12 months ago. The quarterly real GDP growth rate is assumed to be uniform within each quarter. To better infer the macroeconomic states, the four inflation measures are compared. The four output-related variables also capture slightly different dimensions of real production activity. Table 6 reports the mean and the standard deviation of the state variables. More robust statistics are provided in Appendix B.

[Table 6 about here.]

Conditional on each yield curve type, inflation state varies. If we set the normal upward yield curve U1 as a benchmark of low inflatio, on average, the steep yield curve U2 is related to the lowest rate of inflation (around 2.5%). The D and B curve are associated with high inflation rates. The H and F curves are associated with a median rate of inflation. More significantly, the inflation rate associated with the D curve (above 6% on average) has the highest level (about 2 to 3% higher than the inflation rate 3.5-4% associated with the H and F curves). These results are robust across mean, median, and trimmed mean statistics, though conditional volatility of inflation rate varies across yield curve shapes and seems to be proportional to the yield level.

The conditional real output related statistics present certain interesting patterns as well. Setting the U1 curve as a benchmark, the steep upward yield curve U2, on average, corresponds to the most miserable level of real production (very low RGDP with the highest UNEM, lowest INPR and CUR); if we rely on RGDP and UNEM statistics alone, the F curve corresponds to the most active state of the real activity, though INPR and CUR are a little lower than the D curve counterparts. When it comes to the B curve, the UNEM and CUR indicators suggest a relatively active production state though in a lower the real GDP growth environment. The lowest RGDP (1.97%) associated with the B curve seems inconsistent with other indicators. Further investigation reveals a 14-month period deep recession in the 1974–75 associated with the B curve. The indicators associated with the H curve suggest a relatively low level of real production growth.

Combining the findings in conditional inflation and output statistics, a general mapping from yield curve shapes to economic states can be sketched out. Marking the state of the usual yield curve as low inflation rate and median output growth rate, we find that a steep upward yield curve corresponds to the lowest level of inflation rate and output growth rate among all states. The flat yield curve matches a state of a moderate inflation rate and the highest output growth rate. The downward yield curve reflects a state with the highest inflation rate and mediocre output growth rate. The hump and bowl shape yield curves reflect some state in between: relatively high inflation but low output growth.

A word of caution is necessary when interpreting these comparisons among states. First and foremost, as noted in Lu and Wu (2009), non-synchronously released monthly series could lead to misleading or even erroneous conclusions. Although inflation and unemployment state measurements are less subject to this issue, the real output variables must be examined with greater care. (The sample correlations among inflation rate measures are all above 0.9 whereas the correlation between real GDP growth rate and unemployment rate is only -0.2.) Second, due to lags and leads in the macroeconomic interactions, certain macroeconomic variables, such as real GDP and unemployment, can even reflect converse states of the economy in certain periods. For instance, they move in tandem during 2002 to 2008. Last but not least, the terms "high," "low," and "moderate" are all relative to the chosen benchmark, not to a particular state in any specific period.

While the "minor" types of yield curve can signal recessions strongly and reliably, the link between shapes and states is less definite. The attempt to link shapes to states in this section only serves as a starting point for future research. In the next section, statistical methods are applied to model the transition dynamics of the yield curve shapes, estimate the transition probabilities, and test competing models.

5 Markov chain models

As a subset of larger macroeconomic phenomena, yield curve dynamics may correspond to various states in the system. While different shapes of observed yield curves are largely driven by macroeconomic factors, the yield curve in turn feeds back information on the state of the economy (Lu and Wu 2009). Therefore, a state transition model of the dynamic yield curves may capture the economy's changing pattern and also serve to gauge the general state of the economy. This simplification calls for a stochastic process characterized by probability distributions of the state variables and their transitions.

This section begins with key elements of the Markov chain models, followed by a brief introduction to some popular techniques in estimating the transition probabilities of the chain. More technical notes for Markov chain properties and estimation strategies are in Appendix C.

5.1 First-order Markov chain

Let S be defined as defined as $S = \{S_1, ..., S_i, ..., S_N\}$ where S_i is a state of the nature and N is total number of the states. A stochastic process $\{X_t, t \geq 0\}$ on state space S is said to be a first-order Markov chain (MC) if, for all i and j in S, the conditional probability satisfies

$$P(X_{t+1} = S_j | X_t = S_i, X_{t-1}, ..., X_0) = P(X_{t+1} = S_j | X_t = S_i) = P_{ij}^{(t)}. \tag{5.1}$$

A Markov chain $\{X_t, t \geq 0\}$ is said to be time homogenous if, for all $t = 0, 1, 2, ..., T$

$$P(X_{t+1} = S_j | X_t = S_i) = P(X_1 = S_j | X_0 = S_i). \tag{5.2}$$

The five shapes of yield curve constitute the state space $S = \{U, H, F, B, D\}$ in a MC. The conditional probability P_{ij} measures the transition probability from one state to another. All these transition probabilities can be collected into a square matrix $\boldsymbol{P}$. For any row in this transition matrix, the elements in each row must be non-negative and add up to one: $P_{ij} \geq 0$ and $\sum_{j \in S} P_{ij} = 1$ for all S_i and S_j in S.

Another key element of a MC is its initial distribution, or the starting probabilities of each state. Notationally, an initial distribution is the set $\{\pi_i^{(0)} : S_i \in S\}$, where $\pi_i^{(0)} = P(X_0 = S_i)$. Similar to transition probabilities, initial state probability must be non-negative $\pi_i^{(0)} \geq 0$ and add up to one $\sum_{i \in S} \pi_i^{(0)} = 1$. In the context of the yield curve transition process, the initial distribution my be written as a vector $\boldsymbol{\pi^{(0)}} = (\pi_U^{(0)}, \pi_H^{(0)}, \pi_F^{(0)}, \pi_B^{(0)}, \pi_D^{(0)})$.

Thus, a Markov chain is a sequence $\{X_t\}$ of random variables, with an initial distribution $P(X_0 = S_i) = \pi_i^{(0)}$ and transition probabilities $P(X_{t+1} = S_j | X_t = S_i) = P_{ij}$. With this minimum amount of information, we can compute the state distribution $\{\pi_j^{(t)} : S_j \in S\}$ at time t by a theorem: $\pi_j^{(t)} = P(X_t = S_j) = \sum_i^N P(X_0 = S_i) P(X_t = S_j | X_0 = S_i) = \sum_i^N \pi_i^{(0)} P_{ij}^{(t)} = \sum_i^N \pi_i^{(0)} P_{ij}^t$. This theorem is a direct application of the law of total probabilities but contains more information on how to compute t-step transition probabilities of a MC. A proof of the theorem employs the basic Markov property and time homogeneity assumption.

5.2 Estimation methods

The most commonly adopted technique for estimating the MC transition matrix is the maximum likelihood estimation (MLE). This subsection derives the maximum likelihood (ML) estimators and, to save space, lightly touches bootstrapped MLE, Laplace smoothing, and Bayesian estimation. Their detailed derivation, properties, and estimation procedure are covered in Appendix C.3.

A MC process $\{X_t, t = 0, 1, 2, ..., T\}$ evolves in a finite state space $\{S_i \in S\}$ with transition probabilities $P_{ij} = P(X_t = S_j|X_{t-1} = S_i) = P_{ij}(t)$ for all t, and $0 \leq P_{ij} \leq 1$ and $\sum_j P_{ij} = 1$ for all $S_i \in S$. The joint probability of an ordered sequence for such a Markov chain may be written as

$$P(X_0, X_1, ..., X_T) = P(X_0) \prod_t P(X_t|X_{t-1}) = P(X_0) \prod_t P_{ij}(t) = P(X_0) \prod_{i,j} P_{ij}^{n_{ij}} \tag{5.3}$$

where n_{ij} is the total number of events over $t = 1, ..., T$ for which $X_{t-1} = S_i$ and $X_t = S_j$. Equation (5.3) hence derived is the likelihood function of an observed trajectory $(X_0, X_1, ..., X_T)$. To maximize (5.3) or just the product term $\prod_{i,j} P_{ij}^{n_{ij}}$ with respect to P_{ij} subject to probability constraints $0 \leq P_{ij} \leq 1$ and $\sum_j P_{ij} = 1$, the Lagrange multipliers method may be applied. The objective function is set up by taking logs on both sides of (5.3):

$$\mathcal{L}(P_{ij}) = logP(X_0) + \sum_{i,j} n_{ij} * logP_{ij} - \sum_i \lambda_i(\sum_j P_{ij} - 1) \tag{5.4}$$

By the first order condition with respect to P_{ij}, we obtain $\frac{n_{ij}}{P_{ij}} = \lambda_i$. Together with the probability constraint $\sum_j P_{ij} = 1$, the maximum likelihood estimator of the transition probabilities P_{ij} is derived:

$$\hat{\boldsymbol{P}} = (\hat{P}_{ij}) = (\frac{n_{ij}}{\lambda_i}) = (\frac{n_{ij}}{\sum_j n_{ij}}) \geq 0 \tag{5.5}$$

Interestingly, the MLE solution turns out to be very simple and intuitive. The transition probability from state i to state j can be computed from the proportion of its transition count to the total transition count from i to all other states. The standard error of the MLE can be derived from the information matrix:

$$\hat{\sigma}(P_{ij}) = \frac{\hat{P}_{ij}^{MLE}}{\sqrt{n_{ij}}} \tag{5.6}$$

The ML estimator is consistent and converges almost surely to the true transition probability. $\hat{\boldsymbol{P}}$ can also be shown to be asymptotically normal with a rate of convergence $\sqrt{T}$ (Athreya and Fuh, 1992).

In practice, depending on data availability and research purposes, additional techniques are used to enhance the performance of regular MLE. The bootstrap method can help determine the sampling distribution of the MLE by resampling from the estimated MC but its performance is subject to a limited data problem. The Laplace smoothing adds a strictly positive parameter to the MLE so that all state transition probabilities become positive. This method overcomes a limited sample estimation bias when the true transition is most likely to happen given a large sample. The Bayesian method combines prior knowledge of the transition density and sample information to generate an estimator that is usually more precise in terms of a smaller estimation error.

5.3 Model extension

The first-order Markov chain assumes that the future state only depends on the present state, independent of all past memory. However, given any Markov chain, one is not sure about its underlying dependence structure. In essence, higher-order Markov chains assume the transition probabilities are multi-step dependent. Although model complexity increases, it might capture more features of the data and potentially be a better fit. The selection of Markov chain order can be assigned to statistical methods by comparing models' forecast performances.

Introduced in Ching, et al. (2013), a generalized version of Raftery (1985) higher-order

Markov chain model is $\boldsymbol{\pi}^{(t+M+1)} = \sum_{m=1}^{M} \lambda_m \boldsymbol{\pi}^{(t+M+1-m)} \boldsymbol{P_m}$, where time $t+M+1$ state distribution $\boldsymbol{\pi}^{(t+M+1)}$ is a weighted average of the past M state distribution vectors $\boldsymbol{\pi}^{(t+M)}$, $\boldsymbol{\pi}^{(t+M-1)}$, ..., $\boldsymbol{\pi}^{(t+1)}$. Here, M is the highest order of the Markov chain. The weight parameter is λ_m and the m-step transition matrix is $\boldsymbol{P_m}$. The higher-order dependence of $\boldsymbol{\pi}^{(t+N+1)}$ on $\boldsymbol{\pi}^{(t+M+1-m)}$ is relayed by λ_m and P_m in the model. The model also assumes that the weight λ_m is non-negative and $\sum_{m=1}^{M} \lambda_m = 1$. An advantage of the model is that it nests all lower-order Markov chains up to M.

Depending on the order M, the total number of parameters is $(M + M * N^2)$. There are M weight parameters associated with each stochastic matrix and N^2 transition probabilities within each stochastic matrix. While we can construct the m-step transition matrix by computing the relative transition frequency within each state, the estimation of weight parameters λ_m must resort to constrained optimization under the MC equilibrium. Consider the following l_1-norm minimization problem:

$$\underset{\lambda}{\mathrm{Min}}\{\sum_{i=1}^{N} |(\sum_{m=1}^{M} \lambda_m \hat{\boldsymbol{\pi}} \hat{\boldsymbol{P}}_{\boldsymbol{m}} - \hat{\boldsymbol{\pi}})|_1\}$$

subject to $\sum_{m=1}^{M} \lambda_m = 1$ and $\lambda_m \geq 0$, $\forall n$. $\hat{\boldsymbol{\pi}}$ is estimated from the proportion of the occurrence of each state in the Markov chain sequence. $\hat{\boldsymbol{P}}_{\boldsymbol{m}}$ is estimated from the contingency counts. The problem can also be formulated as a linear programming problem. Solution for parameters λ_m is then obtained. Readers may refer to Ching, et al. (2013) for details.

6 Estimation and forecast

In this section, we focus on the most pivotal estimation result for the first-order yield curve transition dynamics and compare it with that of higher-order models. We evaluate and test their forecast performance through k-fold cross-validation. The selected model is then chosen to illustrate how to perform forecast. The details of model estimation, validation, and results for other data frequencies are collected in Appendix D, E and F.

6.1 Estimation results

The regular MLE yields reliable estimates for transition probabilities among various estimation methods. Table 7 presents the results; Figure 3 displays the transition diagram. The transition probabilities in each row add up to 1, and the standard deviation never exceeds the corresponding probability estimate.

[Table 7 about here.]

First and foremost, the within-shape transitions display significant momentum—the diagonal elements dominate each row in the transition probabilities matrix. This implies that a change in the shape of the yield curve from month to month is not frequently observable. The estimated self-transition probabilities with weekly and daily data are even higher. Also, these self-transition probabilities are not positively related to the states' relative sample frequencies. Whereas the D curve has a lower frequency (10.05%) to be observed in the sample than the H curve (10.32%), in one-step transition it has a higher probability (82.89%) of recurring than the H curve (69.23%).

Second, the between-shape transitions are highly asymmetric. In monthly transitions, the U curve is more likely to transition to a H curve than to others, the H curve is more likely to transition to the U curve than to others, the F curve is much more likely to transition to the U or B curve than to the H or D curve, the B curve is much more likely to transition to the U curve than to others, and the D curve is much more likely to transition to the H curve than to others.

Third, several shape transitions are statistically insignificant. Two zero transition probabilities indicate that an upward yield curve never leads a downward yield curve, neither does a hump yield curve lag a bowl shaped yield curve. Among the statistically insignificant transitions (P_{HB}, P_{FH}, P_{FD}, P_{BF}, P_{DU}, and P_{DF}), two are economically significant ($D \to U$ and $H \to B$). In April to May of 1980, we witness one of the sharpest declines in yield levels in the transition of $D \to U$ (around 400 basis points at the short maturity). The transition of $H \rightarrow B$ in November to December 1957 also shifts down, but to a less extent, in yield levels of all maturities (ranging around 20 to 60 basis points). In daily estimation, the transition probabilities from $D \rightarrow U$ and $H \rightarrow B$ are zero, however, which indicates that the two monthly transitions are measurement errors due to over-smoothing in monthly data.

With the stochastic matrix being estimated, we infer that the yield curve Markov chain will converge to a stationary state distribution in the long run: $P_U = 0.7219, P_H = 0.1033, P_F = 0.0212, P_B = 0.0530, P_D = 0.1007$, which is close enough to its sampling frequencies. The convergence time takes 5 to 6 years.

[Table 8 about here.]

We now turn to analyze the results for higher-order Markov chains with $M = 3$. Since higher-order Markov chains nest lower-order ones, the full sample estimation for third-order MC can be shown succinctly in a single table below. First, the estimation assigns equal weights to each transition matrix ($\lambda_1 = \lambda_2 = 0.5$ for second-order MC and $\lambda_1 = \lambda_2 = \lambda_3 = 0.3333$ for third-order MC). Second, the estimated m-step transition probabilities matrices also display sufficient transition momentum but weaken as the order increases (the diagonal elements of P_3 are smaller than their P_2 and P_1 counterparts). Last but not least, when higher-order transition dependence is considered, there are fewer zero entries associated with higher-order transition matrices, which means $U \rightarrow D$ occurs in two steps and $B \rightarrow H$ occurs in three steps. Correspondingly, the self-transition probabilities decrease as the chain order increases, because the model assigns weights to the higher-order transitions. For higher-order models, the shape transitions also show a high degree of momentum and asymmetry.

6.2 Model validation

Is it worth modeling higher-order dependences of the shape transitions? We delegate the task of model evaluation/selection to k-fold cross-validation (henceforth C.V.).[10] Customizing k-

[10] To save space, details are provided in the appendix. Refer to chapter 2.2 and 5.1 of James et al. (2013) for an introduction to this approach. In a nutshell, by comparing average prediction errors of different models,

fold C.V. to the setting of Markov chain models, we make two assumptions to simplify the evaluation process. Since the state distribution at time $t+1$ can be computed by the product of time t state distribution and the stochastic matrix: $\boldsymbol{\pi}^{(t+1)} = \boldsymbol{\pi}^{(t)} \times \boldsymbol{P} = \boldsymbol{\pi}^{(0)} \times \boldsymbol{P}^{t}$, the nature of prediction depends on the choice of initial state distribution $\boldsymbol{\pi}^{(t)}$ and stochastic matrix $\boldsymbol{P}$. The first assumption treats $\boldsymbol{\pi}^{(t)}$ as deterministic for all t, which means our forecaster judges the predicted state by setting the highest probability entry in the state distribution vector to one and others zero. The second assumption treats the estimated stochastic matrix as fixed in recursive predictions, which is less expensive than real time forecast.

Essentially, the Markov chain model evaluation amounts to accessing the prediction error given a model. By the first assumption, we can compute the average prediction error rate for a hold-out test dataset by

$$Ave.(I(S_t \neq \hat{S}_t)) = \frac{1}{T} \sum_{t=1}^{T} I(S_t \neq \hat{S}_t) \tag{6.1}$$

Here $\hat{S}_t$ is the predicted state for the tth observation in the chain, T is the number of data points in the test set, and $I(S_t \neq \hat{S}_t)$ is an indicator variable that equals 1 if $S_t \neq \hat{S}_t$ and 0 if $S_t = \hat{S}_t$. Hence equation (6.1) computes the fraction of incorrect predictions in the validation set. With k-fold C.V., the mean of all k dataset average prediction error rates is used to access the model prediction accuracy. In practice, the choice of k should balance the bias-variance in test errors. Typically, one chooses $k = 5$ or 10, as these values perform well in empirical estimation (James et al., 2013, p184).

[Table 9 about here.]

For Markov chain models of different orders, we compute their respective mean prediction error rates with the result summarized in Table 9. Regardless of the choice of k, the lowest

k-fold cross-validation serves as a general statistical approach to detecting problems of model overfitting and excess data-mining.

prediction error rate is always associated with the first-order Markov chain. This result is somewhat counter-intuitive, as we usually expect that more flexible models (higher-order MC) can capture more transition information and hence fit the data better. However, the validity of this intuition depends on the true data generating process. The power of k-fold cross-validation lies in selecting a model with the lowest test set error rate balancing the bias-variance tradeoff. If the first-order MC is closer to the true yield curve transition dynamics, then k-fold cross-validation would suggest the best alternative model. More attractive, the first-order Markov chain beats higher-order ones for its simplicity. Hence, we must reject higher-order Markov chains and select the most parsimonious model.

Tightly related to the forecast application, k-fold cross-validation reveals a riveting fact that carries economic significance. In 2-fold and 5-fold validations, using 1984 to 2016 data as the test set, the average prediction error rates are on average four times smaller than their training set counterparts for Markov chain models of three orders (see Table E2 and E3 in the Appendix). The estimated transition probability matrix using the 1953 to 1984 sub-sample shows much stronger transition momentum for the B and D curves than that using the 1984 to 2016 sub-sample. The low test error rate also implies that Markov chain models are more successful in predicting future term structure states since the mid-1980s. A possible economic explanation can be attributed to the "Great Moderation"—a period of moderate economic growth along with less fluctuation and reduced inflation risk.

6.3 Forecast exercise

Our purpose is to forecast future yield curve shapes and corresponding macroeconomic states. In Table 10, the forecast horizon is from 1-month to 5-year and the chain starts with deterministic initial state distribution. For instance, starting with a U curve, the probability of observing a U curve next month is 95.96% and zero for a D curve; in a 2-month forecast, the probability of observing a U curve declines to 92.72%, but the probability of observing a D curve increases a bit from 0 to 0.41%. As the forecast horizon extends, the probability of observing a U curve decreases but still dominates other states. It is also noted that, starting from a U curve, the decrease and increase in predicted future state probabilities are monotonic (this is not true if starting from other state distributions).

[Table 10 about here.]

There are some similar evolution patterns of the predicted state distribution even if the initial distributions differ. The most eminent feature is the transition "inertia": a higher chance of remaining in the same state. Starting with a deterministic state distribution vector, the U curve is predicted to dominate with the highest probability, the D curve lasts for four months, the H and B curves last for three months, and the F curve only lasts for one month. The monthly rate of change of transition probability seems to inversely relate to the state empirical frequency: the U curve changes less likely and less dramatically to other states, whereas the opposite is true of the F curve.

Given sufficient passage of time, the predicted state distribution will converge to the stationary distribution of the chain. The convergence property holds for different initial distributions.[11] The last block in the table shows the equilibrium distribution of the transition.

The forecast exercise in Table 10 is static in nature because it assumes both fixed transition matrix and initial state distribution. Preferably, the forecast exercise would be performed in real time with all updated information. While re-estimating the transition probabilities matrix recursively is not always necessary, the inclusion of current and expected future market conditions into the initial state distribution is essential for a superior forecast (i.e., update the state vector in real time forecast).

7 Conclusion

The observed Treasury yield curves vary in shape from time to time, reflecting particular conditions in financial markets and general states in the economy. This paper studies the

[11] This property results from the irreducibility and aperiodicity of the Markov chain. Appendix C.1 provides further details.

links between yield curve shapes and the macroeconomy, as well as models shape transitions.

While the yield curve is upward sloping most often, its relative frequency declines substantially in the pre-recession periods. In contrast, the four less frequent yield curves all carry useful information useful for predicting the U.S. recessions, especially after the mid-1980s. Moreover, the variations in the shapes of the yield curve may reflect the changing production and inflation states in the economy. Given the mapping between shapes and states, the term structure transition process can be modeled in Markov chains. Estimation shows that the shape transitions display significant momentum and asymmetry.

The findings on the time-varying predictive power of various shapes of the yield curve are interesting but puzzling; the exploration of the links between yield curve shapes and macroeconomic states is still at its prenatal stage, and the shape transition estimation of the yield curve reveals new term structure dynamics. We hope to construct more precise shape-state mapping and gain deeper insight into the driving forces of the term structure dynamics in future research.

References

Abdymomunov, A. (2013), "Predicting output using the entire yield curve," *Journal of Macroeconomics*, 37, 333-344.

Argyropoulos, E., Tzavalis, E. (2016), "Forecasting economic activity from yield curve factors," *North American Journal of Economics and Finance*, 36, 293-311.

Ang, A., and Piazzesi, M. (2003), "A no-arbitrage vector autoregression of term structure of dynamics with macroeconomic and latent variables," *Journal of Monetary Economics*, 50, 745-787.

Ang, A., Piazzesi, M., and Wei, M. (2006), "What does the yield curve tell us about GDP growth?" *Journal of Econometrics*, 131, 359-403.

Ang, A., Bekaert, G., and Wei, M. (2007)," Do macro variables, asset markets or surveys forecast inflation better?" *Journal of Monetary Economics*, 54, 1163-1212.

Athreya, K. B., and Fuh, C. D. (1992), *Bootstrapping Markov chains.* New York: John Willey & Sons.

Benati, L., and Goodhart, C. (2008), "Investigating time-variation in the marginal predictive power of the yield spread," *Journal of Economic Dynamics and Control*, 32(4), 1236-1272.

Bekaert, G., and Hodrick, R. J. (2001), "Expectation hypotheses tests," *Journal of Finance*, 56(4), 1357-1394.

Bibkov, R., and Chernov, M. (2010), "No-arbitrage macroeconomic determinants of the yield curve," *Journal of Econometrics*, 159, 166-182.

Bordo, M. D., and Haubrich, J. G. (2008), "The yield curve as a predictor of growth: long-run evidence, 1875-1997," *Review of Economics and Statistics*, 90(1), 182-85.

Chauvet, M., and Senyuz, Z. (2016). "A dynamic factor model of the the yield curve components as a predictor of the economy," *International Journal of Forecasting*, 32, 324-343.

Campbell, J. Y., and Shiller, R. J. (1991), "Yield spreads and interest rate movements: a bird's eye view," *Review of Economic Studies*, 58, 495-514.

Ching, W. K., Huang, X., Ng, M. K., and Siu, T. K. (2013), *Markov chains, International series in operations research & management science 189*, Springer Science.

Culbertson, J.M . (1957), "The Term Structure of interest Rates," *Quarterly Journal of Economics*, 71, 485—517.

Dewachter, H., Lyrio, M., and Maes, K. (2006), "A joint model for the term structure of interest rates and the macroeconomy," *Journal of Applied Econometrics*, 21, 439-462.

Diebold, F., Rudebusch, and G., Aruoba, S., (2006), "The macroeconomy and the yield curve: a dynamic latent factor approach," *Journal of Econometrics*, 131, 309-338.

Diebold, F., and Rudebusch, G. (2013), *Yield Curve Modeling and Forecasting.* Princeton University Press.

Estrella, A., and Hardouvelis, G. A. (1991), "The term structure as a predictor of real economic activity," *Journal of Finance*, 46(2), 555-576.

Estrella, A., and Mishkin, F. S. (1996), "The yield curve as a predictor of U.S. recessions," Federal Reserve Bank of New York *Current Issues in Economics and Finance*, 2(7).

— — (1997), "The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank," *European Economic Review*, 41, 1375-1401.

— — (1998), "Predicting US recessions: Financial variables as leading indicators," *The Review of Economics and Statistics*, 80, 45-61.

Estrella, A., Trubin, M. R. (2006), "The yield curve as a leading indicator: Some practical issues," Federal Reserve Bank of New York *Current Issues in Economics and Finance*, 12(5).

Fama, E. F. (2006), "The behavior of interest rates," *The Review of Financial Studies*, 19, 359-379.

Frankel, J., and Lown, C. (1994), "An indicator of future inflation extracted from the steepness of the interest rate yield curve along its entire length," *Quarterly Journal of Economics*, 109, 517-530.

Hamilton, J., and Kim, D. H. (2002), "A reexamination of the predictability of economic activity using the yield spread," *Journal of Money, Credit, and Banking*, 34, 340-360.

Harvey, C. (1988), "The real term structure and consumption growth," *Journal of Financial Economics*, 22, 305-333.

— — (1989), "Forecasts of economic growth from the bond and stock market," *Financial Analysts Journal*, September/October: 38-45.

— — (1991), The term structure and world economic growth. *The Journal of Fixed Income*, 1, 7-19.

Hicks, J. (1939), *Value and capital*, Oxford Univeristy Press

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013), *An introduction to statistical learning with applications in R.* Springer.

Kamara, A. (1997), "The relation between default-free interest rates and expected economic growth is stronger than you think," *Journal of Finance*, 52, 1681-1694.

Kessel, R. (ed.) (1965), *The cyclical behavior of the term structure of interest rates.* New York: National Bureau of Economic Research.

Lange, R. (2013), "The Canadian macroeconomy and the yield curve: A dynamic latent factor approach," *International Review of Economics and Finance*, 27, 261?274.

Lu, B., and Wu, L. (2009), "Macroeconomic releases and the interest rate term structure," *Journal of Monetary Economics*, 56, 872-884.

Lutz, F. A. (1940), "The Structure of Interest Rates," *The Quarterly Journal of Economics*, 55(1), 36-63.

Mishkin, F. (1990a), "What does the term structure tell us about future inflation?" *Journal of Monetary Economics*, 25, 77-95.

— — (1990b), “The information in the longer-maturity term structure about future inflation,” *Quarterly Journal of Economics*, 55, 815-828.

Modigliani, F. and Sutch, R. (1966), “Innovations in interest rate policy,” *American Economic Review*, 56(1), 178-197.

Moench, E. (2012), ”Term structure surprises: The predictive content of curvature, level, and slope,” *Journal of Applied Econometrics*, 27, 574-602.

Plosser, C., and Rouwenhorst, K. (1994), “International term structures and real economic growth,” *Journal of Monetary Economics*, 33, 133-155.

Raftery, A. E. (1985), “A model for higher-order Markov chains,” *Journal of the Royal Statistical Society. Series B*, 47(3), 528-539.

Roma, A., and Torous, W. (1997), “The cyclical behavior of interest rates,” *Journal of Finance*, 52, 1519-1542.

Rudebusch, G., and Wu, Tao. (2008), “A macro-finance model of the term structure, monetary policy and the economy,” *The Economic Journal*, 118, 906-926.

Rudebusch, G., and Williams, J. (2009), “Forecasting Recessions: The Puzzle of the Enduring Power of the Yield Curve,” *Journal of Business and Economic Statistics*, 27(4), 492-503.

Stock, J. H., and Watson, M. W. (2003), “Forecasting output and inflation: the role of asset prices,” *Journal of Economic Literature*, 41(3), 788-829.

Wheelock, D. C., and Wohar, M. E. (2009), “Can term spread predict output growth and recessions? A survey of the literature,” *Federal Reserve Bank of St. Louis Review*, 91(5), 419-440.

Table 1: Classification algorithm for five shapes of the yield curve

Yield curve shapes	Term structure relations with 0.1 percent threshold	Sample
Upward (U)	$(Y_m - Y_s > 0.1 \,\&\, Y_m \leqq Y_l)$ or $(Y_s \leqq Y_m \,\&\, Y_l - Y_m > 0.1)$	2016-03
Hump (H)	$(Y_m - Y_s > 0.1 \,\&\, Y_m > Y_l)$ or $(Y_s < Y_m \,\&\, Y_m - Y_l > 0.1)$	1982-07
Flat (F)	$\lvert Y_m - Y_s \rvert \leqq 0.1$ and $\lvert Y_l - Y_m \rvert \leqq 0.1$	2006-02
Bowl (B)	$(Y_s - Y_m > 0.1 \,\&\, Y_m < Y_l)$ or $(Y_s > Y_m \,\&\, Y_l - Y_m > 0.1)$	2006-08
Downward (D)	$(Y_s - Y_m > 0.1 \,\&\, Y_m \geqq Y_l)$ or $(Y_s \geqq Y_m \,\&\, Y_m - Y_l > 0.1)$	1982-02

Notes: Y_s, Y_m, Y_l are the average of Treasury bill, note, and bond yields, respectively.

Table 2: Classification result and key statistics for monthly yield curves (1953.4–2016.3)

Shapes	Ocurrence	$\overline{Y}_s$	$\overline{Y}_m$	$\overline{Y}_l$	$\overline{Y}_m - \overline{Y}_s$	$\overline{Y}_l - \overline{Y}_m$	$\overline{Y}_l - \overline{Y}_s$	$\overline{Y}_m - \frac{\overline{Y}_s + \overline{Y}_l}{2}$
Upward (U)	546 (72.22%)	3.97	5.07	5.93	1.10	0.86	1.96	0.12
		(2.76)	(2.78)	(2.52)	(0.61)	(0.68)	(1.15)	(0.29)
Hump (H)	78 (10.32%)	6.34	6.72	6.49	0.38	-0.23	0.15	0.31
		(3.02)	(3.16)	(3.11)	(0.39)	(0.15)	(0.43)	(0.21)
Flat (F)	16 (2.12%)	5.96	5.99	5.98	0.03	-0.01	0.02	0.02
		(2.13)	(2.11)	(2.11)	(0.06)	(0.07)	(0.09)	(0.04)
Bowl (B)	40 (5.29%)	6.67	6.29	6.53	-0.38	0.24	-0.14	-0.31
		(1.62)	(1.50)	(1.49)	(0.28)	(0.16)	(0.34)	(0.15)
Downward (D)	76 (10.05%)	8.75	8.22	7.77	-0.53	-0.45	-0.99	-0.04
		(3.79)	(3.40)	(3.25)	(0.52)	(0.31)	(0.70)	(0.25)
Full sample	756 (100%)	4.88	5.64	6.20	0.76	0.56	1.33	0.10
		(3.26)	(3.00)	(2.67)	(0.81)	(0.77)	(1.47)	(0.30)

Notes: $\overline{Y}_s$, $\overline{Y}_m$, $\overline{Y}_l$ are the sample means of the averaged Treasury bill, note, and bond yields in percentage, respectively. Standard deviations in the parentheses. Source: The original monthly yields data are downloaded from the Federal Reserve Board H.15 Treasury nominal yield statistics, covering Apr.1953 to Mar.2016.

Table 3: Yield statistics over business cycle stages (1953.7–2010.6)

I: all pre-recession 18-month periods (155 months)
II: all recession periods (121 months)
III: all post-recession 12-month periods (101 months)
IV: all remaining periods (307 months)

Statistics	I	II	III	IV
Y_s	6.71 (2.95)	5.94 (4.32)	3.92 (2.85)	4.94 (2.07)
Y_m	6.63 (2.61)	6.51 (3.95)	5.31 (2.80)	5.90 (2.22)
Y_l	6.52 (2.48)	6.91 (3.49)	6.12 (2.61)	6.50 (2.22)
$Y_m - Y_s$	-0.08 (0.56)	0.57 (0.82)	1.38 (0.54)	0.95 (0.71)
$Y_l - Y_m$	-0.11 (0.32)	0.40 (0.76)	0.81 (0.70)	0.61 (0.61)
$Y_l - Y_s$	-0.19 (0.76)	0.97 (1.42)	2.19 (1.17)	1.56 (1.23)
$Y_m - \frac{Y_s+Y_l}{2}$	0.02 (0.26)	0.09 (0.35)	0.29 (0.22)	0.17 (0.24)

Note: Y_s, Y_m, Y_l are the averages of Treasury bill, note, and bond yields, respectively. Mean and standard deviation (in parentheses) of different yield measures. Source: The original monthly yields data are downloaded from the Federal Reserve Board H.15 Treasury nominal yield statistics.

Table 4: Yield curve type frequency counts through the U.S. business cycles (1953.7–2010.6)

NBER recessions	Pre-recession (18)					In recession					Post-recession (12)				
(duration in months)	U	H	F	B	D	U	H	F	B	D	U	H	F	B	D
Jul.1953 : May.1954 (11)	-	-	-	-	-	11	0	0	0	0	12	0	0	0	0
Aug.1957 : Apr.1958 (9)	3	13	2	0	0	4	2	0	1	2	10	2	0	0	0
Apr.1960 : Feb.1961 (11)	5	7	0	0	6	8	3	0	0	0	12	0	0	0	0
Dec.1969 : Nov.1970 (12)	0	3	0	0	15	1	8	0	0	3	8	4	0	0	0
Nov.1973 : Mar.1975 (17)	5	5	0	6	2	5	0	0	12	0	12	0	0	0	0
Jan.1980 : Jul.1980 (7)	1	0	1	2	14	3	0	0	0	4	1	1	0	0	10*
Jul.1981 : Nov.1982 (17)	4	1	0	0	13	5	8	0	0	4	12	0	0	0	0
Jul.1990 : Mar.1991 (9)	4	6	6	2	0	9	0	0	0	0	12	0	0	0	0
Mar.2001 : Nov.2001 (9)	6	6	0	4	2	9	0	0	0	0	12	0	0	0	0
Dec.2007 : Jun.2009 (19)	6	0	0	12	0	19	0	0	0	0	12	0	0	0	0
Column sum frequency:	34	41	9	26	52	74	21	0	13	13	103	7	0	0	10
Local relative frequency:	.21	.25	.06	.16	.32	.61	.17	0	.11	.11	.86	.06	0	0	.08
Global relative frequency:	.06	.53	.56	.65	.68	.14	.27	0	.33	.17	.19	.09	0	0	.13

Note: Recessions are dated by NBER; yield data from H.15 statistics and author's classification. The asterisk indicates an overlap with the coming pre-recession. The Local relative frequency is defined as the column sum over total counts of all types within respective business cycle stage; global relative frequency is defined as the column sum over total counts of the respective type in the entire 1953 to 2016 monthly sample.

Table 5: Conditional probability of business cycle stage on yield curve shapes (1953.4–2016.3)

Panel A: Given a yield curve shape at t, the probability a business cycle stage is underway in the $t+k$ period, for $k = 1, 2, 3, 4, 15, 16, 17, 18$ month.

	$Prob(Stage_{t+1}\|Shape_t)$				$Prob(Stage_{t+2}\|Shape_t)$				$Prob(Stage_{t+3}\|Shape_t)$				$Prob(Stage_{t+4}\|Shape_t)$			
	Pre-rec.	Rec.	Post-rec.	Oth.	Pre-rec.	Rec.	Post-rec.	Oth.	Pre-rec.	Rec.	Post-rec.	Oth.	Pre-rec.	Rec.	Post-rec.	Oth.
U	.070	.125	.178	.629*	.077	.114	.176	.632*	.087	.103	.175	.635	.098	.090	.175	.637*
H	.513*	.295	.051	.141	.500*	.295	.064	.141	.487*	.295	.077	.141	.474*	.308	.077	.141
F	.562*	.000	.000	.438	.562*	.000	.000	.438	.625*	.000	.000	.375	.688*	.000	.000	.312
D	.605*	.211	.000	.184	.566*	.263	.000	.171	.500*	.329	.000	.171	.447*	.382	.000	.171
B	.625*	.350	.000	.250	.575*	.400	.000	.025	.550*	.425	.000	.025	.500*	.450	.000	.050
	$Prob(Stage_{t+15}\|Shape_t)$				$Prob(Stage_{t+16}\|Shape_t)$				$Prob(Stage_{t+17}\|Shape_t)$				$Prob(Stage_{t+18}\|Shape_t)$			
	Pre-rec.	Rec.	Post-rec.	Oth.	Pre-rec.	Rec.	Post-rec.	Oth.	Pre-rec.	Rec.	Post-rec.	Oth.	Pre-rec.	Rec.	Post-rec.	Oth.
U	.203	.038	.094	.665*	.209	.042	.083	.666*	.214	.047	.074	.665*	.214	.057	.066	.663*
H	.282	.359*	.295	.064	.256	.397*	.269	.077	.256	.385*	.256	.103	.256	.372*	.244	.128
F	.188	.438*	.062	.312	.250	.375*	.125	.250	.188	.375*	.188	.250	.188	.312*	.250	.250
D	.224	.447*	.158	.171	.224	.395*	.211	.171	.224	.368*	.250	.158	.224	.342*	.289	.145
B	.125	.525*	.350	.000	.075	.525*	.400	.000	.050	.525*	.425	.000	.050	.500*	.425	.025

Panel B: Given a yield curve shape at t, the probability of a recession starting in the $(t+k)^{th}$ month, $Prob(Rec._{t+k}^{*}|Shape_t)$, for $k = 1:16$.

	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10	k=11	k=12	k=13	k=14	k=15	k=16
U	.005	.004	.002	.004	.004	.005	.002	.002	.004	.004	.004	.004	.005	.005	.000	.004
H	.026	.013	.026	.026	.013	.000	.013	.026	.038	.038	.038	.038	.026	.013	.051*	.038
F	.000	.000	.000	.000	.000	.063	.063	.063	.063	.063	.063	.000	.063	.063	.000	.000
D	.039	.053	.066*	.066*	.053	.053	.039	.039	.026	.026	.026	.026	.026	.026	.026	.013
B	.025	.050	.025	.000	.050	.025	.075*	.050	.025	.025	.025	.050	.025	.050	.075*	.075*

Given a yield curve shape at t, the cumulative sum probability of a recession starting within the future q quarters, for $q = 2, 4, 6$.

	q=2	q=4	q=6		q=2	q=4	q=6		q=2	q=4	q=6		q=2	q=4	q=6		q=2	q=4	q=6
U	.024	.042	.071	H	.103	.295	.449	F	.063	.375	.563	D	.329	.513	.671	B	.175	.425	.700

Note: Yield curve shape notation follows U—upward, H—hump, F—flat, B—bowl, D—downward; Recessions are dated by NBER and business cycle stages consist of pre-recession 18-month period (pre-rec.), recessions (rec.), post-recession 12-month period (post-rec.), and other periods. Panel A calculates the k-month ahead conditional probability of a business cycle stage on a particular yield curve shape if the stage is underway in $t+k$, $Prob(Stage_{t+k}|Shape_t)$, for k=1:16. Panel B calculates the conditional probability $Prob(Recession_{t+k}^{*}|Shape_t)$ of a future recession on a particular yield curve shape if the recession starts in the $(t+k)^{th}$ for $k = 1:16$ month ahead), and the cumulative probability of a recession starting within the future q quarters, for $q = 2, 4, 6$. Asterisk* indicates the maximum probability within each row.

Table 6: Statistics for macroeconomic state variables conditional on yield curve shapes

U1: Usual upward shape (301). U2: Steep upward shape (245). H: Hump shape (78). F: Flat yield curve (16). B: Bowl shape (40). D: Downward shape (76)

Statistics	U1	U2	H	F	B	D
CPIA	3.01 (2.55)	2.58 (1.49)	3.82 (2.05)	3.64 (1.65)	5.74 (3.27)	6.63 (3.73)
CPIC	3.42 (2.41)	2.81 (1.38)	4.42 (2.25)	3.27 (1.76)	4.39 (2.37)	6.36 (3.37)
PPIA	2.46 (3.35)	1.79 (2.70)	3.09 (1.88)	3.91 (1.88)	7.00 (4.96)	6.06 (4.03)
PPIC	2.76 (3.70)	1.77 (1.37)	4.03 (2.36)	3.65 (2.38)	4.92 (4.80)	5.47 (3.29)
RGDP	3.56 (2.51)	2.51 (2.17)	2.91 (2.73)	4.39 (2.35)	1.97 (2.44)	3.26 (2.23)
UNEM	5.74 (1.39)	7.00 (1.42)	5.18 (1.62)	4.79 (0.63)	4.88 (0.54)	5.06 (1.50)
INPR	3.26 (5.90)	1.48 (4.93)	2.71 (5.51)	3.77 (3.95)	2.85 (3.13)	4.24 (3.49)
CUR	0.81 (0.03)	0.78 (0.04)	0.82 (0.04)	0.83 (0.02)	0.84 (0.03)	0.84 (0.03)

Note: U2 yield curve has a long-short spread larger than 200 basis points. Source: FRED - St. Louis. CPIA–CPI all items inflation rate, CPIC–core CPI inflation rate, PPIA–PPI all item inflation rate, PPIC–core PPI inflation rate, RGDP–real GDP growth rate, UNEM–unemployment rate, INPR–industrial production growth rate, CUR–capacity utilization rate. Except for unemployment rate and capacity utilization rate, all other economic indicators are expressed in annual percentage rate calculated from seasonally adjusted data.

Table 7: Maximum likelihood estimates of transition probabilities matrix

States	U	H	F	B	D
U	0.9596 (0.0420)	0.0257 (0.0068)	0.0073 (0.0037)	0.0073 (0.0037)	0.0000 (0.0000)
H	0.1410 (0.0425)	0.6923 (0.0942)	0.0256 (0.0181)	0.0128 (0.0128)	0.1282 (0.0405)
F	0.1875 (0.1083)	0.0625 (0.0625)	0.5000 (0.1768)	0.1875 (0.1083)	0.0625 (0.0625)
B	0.1750 (0.0661)	0.0000 (0.0000)	0.0250 (0.0250)	0.7500 (0.1369)	0.0500 (0.0354)
D	0.0132 (0.0132)	0.1184 (0.0395)	0.0132 (0.0132)	0.0263 (0.0186)	0.8289 (0.1044)

Note: Yield curve shape notation follows U—upward, H—hump, F—flat, B—bowl, D—downward. The monthly yield curve Markov chain in estimation is a categorical five-state sequence classified by the effective algorithm. Standard errors are in parentheses. Source: The original monthly yields data are downloaded from the Federal Reserve Board H.15 Treasury nominal yield statistics, covering Apr.1953 to Mar.2016.

Table 8: Parameter estimates of higher-order Markov chains for M=3

	P1 transition matrix					P2 transition matrix					P3 transition matrix				
	U	H	F	B	D	U	H	F	B	D	U	H	F	B	D
U	.960	.026	.007	.007	.000	.936	.039	.011	.011	.004	.921	.046	.013	.015	.006
H	.141	.692	.026	.013	.128	.218	.564	.051	.013	.154	.244	.474	.051	.038	.192
F	.188	.063	.500	.188	.063	.313	.125	.250	.188	.125	.375	.186	.186	.186	.063
B	.175	.000	.025	.750	.050	.225	.000	.025	.650	.100	.275	.025	.025	.576	.100
D	.013	.118	.013	.026	.829	.053	.145	.013	.053	.737	.092	.158	.013	.066	.671

First order MC $\lambda_1 = 1$; Second-order MC $\lambda_1 = \lambda_2 = 0.5$; Third-order MC $\lambda_1 = \lambda_2 = \lambda_3 = 0.3333$.

Note: Higher-order Markov chain model $\boldsymbol{\pi}^{(t+M+1)} = \sum_{m=1}^{M} \lambda_m \boldsymbol{\pi}^{(t+M+1-m)} \boldsymbol{P_m}$. Yield curve shape notation follows U—upward, H—hump, F—flat, B—bowl, D—downward. The monthly yield curve Markov chain in estimation is a categorical five-state sequence classified by the effective algorithm. Source: The original monthly yields data are downloaded from the Federal Reserve Board H.15 Treasury nominal yield statistics, covering Apr.1953 to Mar.2016.

Table 9: K-fold C.V. prediction error rates

	1st-order MC	2nd-order MC	3rd-order MC
k=1	0.1020	0.1154	0.1261
k=2	0.1257	0.1984	0.2169
k=5	0.1020	0.1271	0.1444

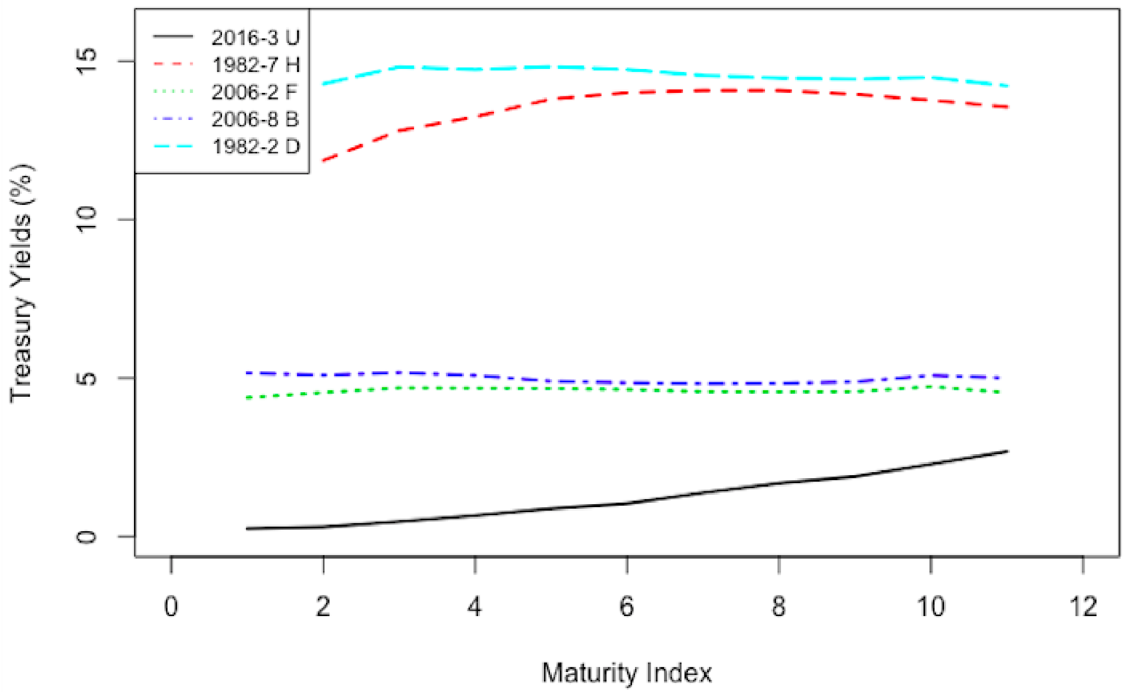
Note: The monthly yield curve Markov chain in estimation is a categorical five-state sequence classified by the effective algorithm. Data are from Federal Reserve Board H.15 interest rate statistics, covering Apr.1953 to Mar.2016.

Table 10: Full-sample monthly forecast with deterministic initial distributions

Forecast	U	H	F	B	D	U	H	F	B	D
Horizon	$\boldsymbol{\pi}^{(0)} = (1,0,0,0,0)$					$\boldsymbol{\pi}^{(0)} = (0,1,0,0,0)$				
1-month	0.9596	0.0257	0.0073	0.0073	0.0000	0.1410	0.6923	0.0256	0.0128	0.1282
2-month	0.9272	0.0429	0.0116	0.0143	0.0041	0.2417	0.4997	0.0336	0.0277	0.1973
3-month	0.9005	0.0547	0.0141	0.0203	0.0103	0.3162	0.3776	0.0347	0.0405	0.2311
6-month	0.8431	0.0741	0.0175	0.0333	0.0320	0.4557	0.2132	0.0299	0.0612	0.2399
1-year	0.7810	0.0893	0.0196	0.0450	0.0650	0.5907	0.1389	0.0246	0.0649	0.1809
2-year	0.7367	0.0997	0.0208	0.0512	0.0916	0.6883	0.1115	0.0220	0.0568	0.1213
5-year	0.7221	0.1033	0.0212	0.0530	0.1005	0.7213	0.1034	0.0212	0.0530	0.1010
	$\boldsymbol{\pi}^{(0)} = (0,0,1,0,0)$					$\boldsymbol{\pi}^{(0)} = (0,0,0,1,0)$				
1-month	0.1875	0.0625	0.5000	0.1875	0.0625	0.1750	0.0000	0.0250	0.7500	0.0500
2-month	0.3161	0.0867	0.2585	0.2382	0.1004	0.3045	0.0120	0.0332	0.5698	0.0805
3-month	0.4071	0.0962	0.1411	0.2332	0.1225	0.4009	0.0277	0.0344	0.4381	0.0988
6-month	0.5592	0.1045	0.0402	0.1545	0.1416	0.5686	0.0688	0.0291	0.2151	0.1184
1-year	0.6635	0.1094	0.0233	0.0756	0.1281	0.6778	0.1005	0.0231	0.0833	0.1155
2-year	0.7095	0.1060	0.0215	0.0549	0.1080	0.7144	0.1047	0.0214	0.0546	0.1049
5-year	0.7217	0.1034	0.0212	0.0530	0.1008	0.7217	0.1033	0.0212	0.0530	0.1007
	$\boldsymbol{\pi}^{(0)} = (0,0,0,0,1)$					$\boldsymbol{\pi}^{(0)} = \boldsymbol{\pi}^{*} = \boldsymbol{\pi}(\infty)$				
1-month	0.0132	0.1184	0.0132	0.0263	0.8289	0.7219	0.1033	0.0212	0.0530	0.1007
2-month	0.0473	0.1813	0.0213	0.0456	0.7045	0.7219	0.1033	0.0212	0.0530	0.1007
3-month	0.0922	0.2115	0.0260	0.0594	0.6105	0.7219	0.1033	0.0212	0.0530	0.1007
6-month	0.2407	0.2182	0.0303	0.0788	0.4320	0.7219	0.1033	0.0212	0.0530	0.1007
1-year	0.4681	0.1672	0.0273	0.0771	0.2603	0.7219	0.1033	0.0212	0.0530	0.1007
2-year	0.6565	0.1193	0.0228	0.0605	0.1408	0.7219	0.1033	0.0212	0.0530	0.1007
5-year	0.7208	0.1036	0.0212	0.0531	0.1013	0.7219	0.1033	0.0212	0.0530	0.1007

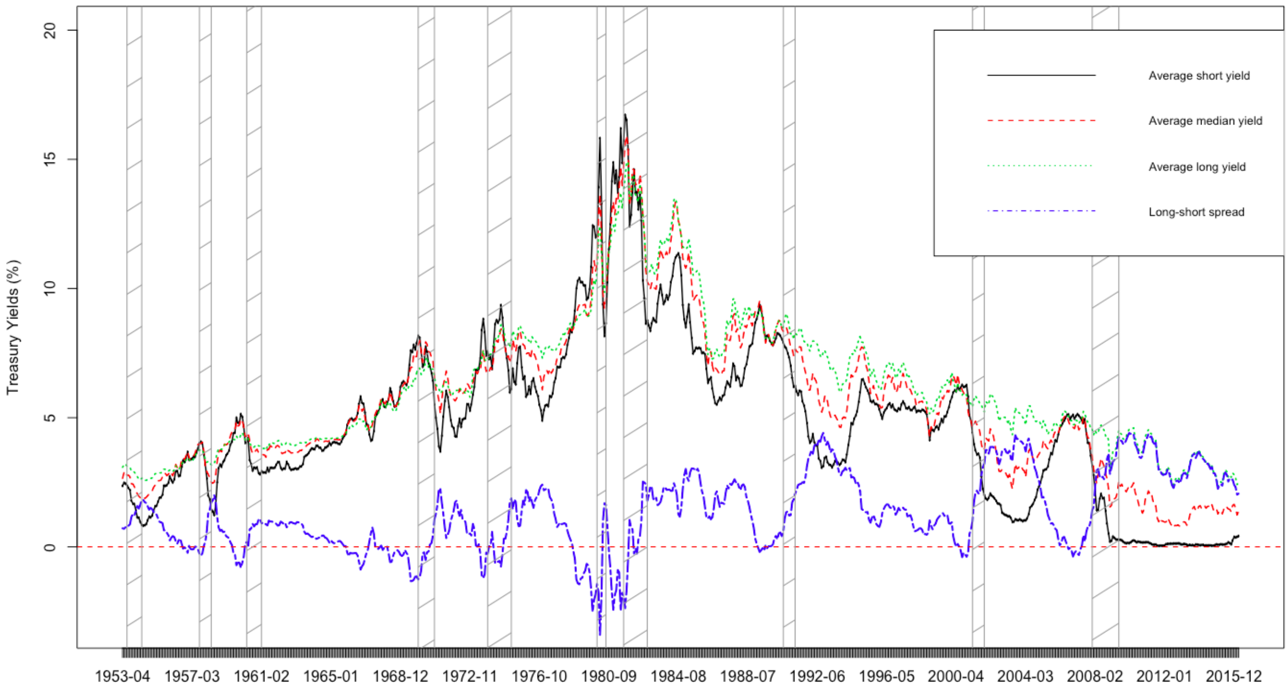
Note: The forecasting equation is $\boldsymbol{\pi}^{(t+1)} = \boldsymbol{\pi}^{(t)} \times \boldsymbol{P} = \boldsymbol{\pi}^{(0)} \times \boldsymbol{P}^{t}$, where $\boldsymbol{\pi}^{(0)}$ is the initial distribution and $\boldsymbol{P}$ is the transition probabilities matrix. The forecast is performed in first-order Markov chain model with regular maximum likelihood estimation. We assume $\boldsymbol{P}$ is fixed and $\boldsymbol{\pi}^{(t)}$ is known in forecasting $\boldsymbol{\pi}^{(t+1)}$. $\boldsymbol{\pi}^{(0)}$, $\boldsymbol{\pi}^{*}$, $\boldsymbol{\pi}^{(\infty)}$ stand for the initial, stationary, and limiting distributions, respectively.

Figure 1: Observation samples of the U.S. Treasury yield curves.



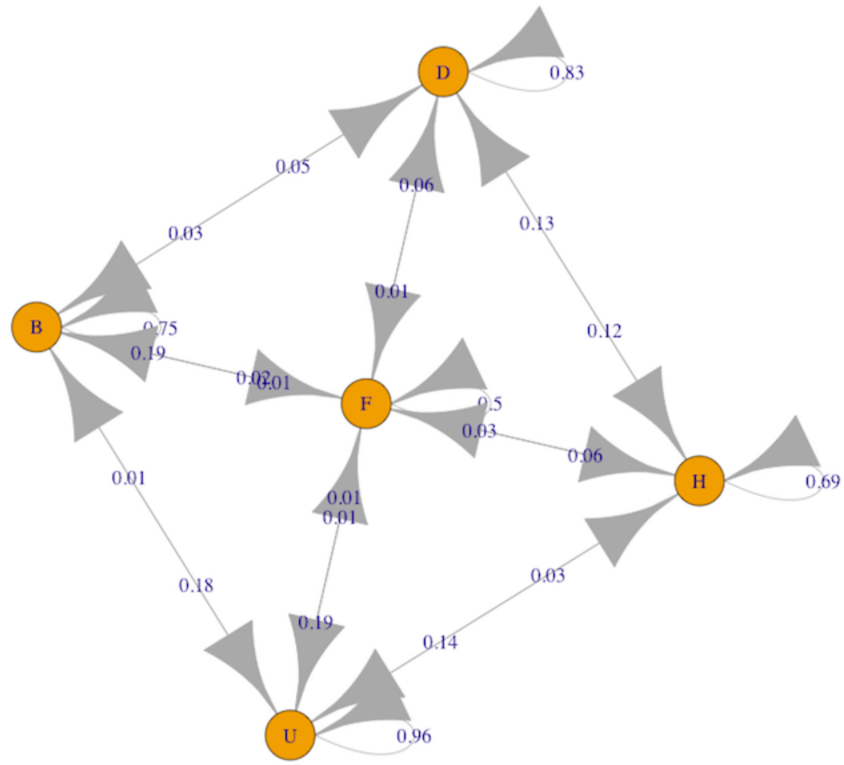
Note: Yield curve shape notation follows U—upward, H—hump, F—flat, B—bowl, D—downward. Data are from Federal Reserve Board H.15 interest rate statistics, covering Apr.1953 to Mar.2016. Shapes are classified by the algorithm described in Section 3.

Figure 2: U.S. Treasury yield co-movements over 1953 to 2016



Note: Author's calculation and graphing using yield data from the Federal Reserve Board H.15 interest rate statistics. Shaded areas are NBER dated recessions.

Figure 3: Transition diagram of the U.S. Treasury yield curve shapes



Note: Yield curve shape notation follows U—upward, H—hump, F—flat, B—bowl, D—downward. Five circles represent the classified shapes of the yield curve and arrows direct the shape transitions. The fractional numbers are estimated transition probabilities from one shape to another using monthly yield data. Data are from Federal Reserve Board H.15 interest rate statistics, covering Apr.1953 to Mar.2016.