The Topology of Time Series: Improving Recession Forecasting from Yield Spreads

Pawel Dlotko[1] and Simon Rudkin*[2]

[1]Mathematics Department, Swansea University, Bay Campus, Swansea, SA8 1EN, United Kingdom
[2]School of Management, Swansea University, Bay Campus, Swansea, SA8 1EN, United Kingdom

June 3, 2019

Abstract

Recession forecasting ranges from simplistic inference from the inversion of the yield curve to sophisticated models drawing data from across the macroeconomic and financial spectra. Each has advantages, in simplicity and informativeness respectively, but each suffers for these. Demonstrating how the properties of yield spread time series themselves can foretell of impending recessions we introduce data topology to economics. Through an exploration of the topology of time series we highlight an untapped source of information with the potential to significantly improve understanding of the economy without risking the overfitting of introducing other variables.

1 Introduction

Questions are regularly asked about the timing of the next recession; business is wary and investors are keeping a very close eye on events for a signal that the latest growth period is over. Recession prediction is an established challenge for the forecasting literature and yet the yield spread inversion remains a go to rule of thumb for determining when the recession will come. In an era of big data and fast computing there is much that can be done to augment our understanding. A risk of replacing a lack of information with a danger of over-fitting presents itself, however. Within the time frame with sufficient available data there have been just 7 recessions, and two of those are the double-dip of the early 1980s. Relative to the more than 50 years of available spread data this is a very low number of positive outcome. Accuracy of forecasts naturally suffers. As we reflect upon the poor performance of recession prediction to date (An et al., 2018) there is an early movement to look back inside the simple yield spread for answers(Benzoni et al., 2018; Kozlowski and Sim, 2019). This paper is another step in that direction.

Data topology locates patterns within a time series to unearth further information about its evolving dynamics. Topological data analysis (TDA), as the study of the shape of data point clouds, can give deep insights on time series in much the way it has been applied across the physical sciences (Perea, 2019). Recent advances in the capturing of periodicity, and in the understanding of the point clouds of longitudinal data, suggest this appreciation for the shape of data can bring value within recession forecasting; Gidea and Katz (2018); Gidea et al. (2018) work on stock market crashes for

*Corresponding author. Tel: (0044)1792 606159, Email:s.t.rudkin@swansea.ac.uk

example. Literature discussing periodicity on the exit from recessions (Clark, 2004; Wosnitza and Sornette, 2015), and the chaotic dynamics on the run into downturns all offer basis for there to be interest in the behaviour of the yield spread series relative to recessions. This paper contributes an evaluation of that potential against the most popular of recession forecast models. Specific focus is placed on the landscape of the point cloud, captured by its norm, and the periodicity of the yield spread extrapolated from the time series using proprietary code.

Recognising the desirability of using minimal numbers of series for maximal benefit this paper offers an alternative enhancement to the yield spread forecast models. In so doing three contributions to the literature from this paper are made. Firstly we demonstrate that it is not just the spread that foretells of impending recession; its behaviour along its evolution is also a factor. Secondly, we demonstrate within economics for the first time, an approach that has already made inroads into time series understanding in chemistry, biology and the physical sciences. The advancement of TDA methodology employed in this paper takes lessons from extant work in the interdisciplinary sphere to produce a deeper evaluation of the properties of time series. Finally, in addition to gaining improved recession prediction by augmenting the TDA, we contribute to the wider discussion on the periodicity, and behaviour, of the yield spread.

The remainder of the paper proceeds as follows. Section 2 discusses the present state of the literature on recession forecasting before 3 introduces the topology of time series and the understanding gained therefrom. We introduce our interest rate data in Section 4 before expositing the topology of the US series in Section 5. Section 6 presents the results of our forecasting, with Section 7 reviewing the implications of our work. Conclusions and directions for future work are offered in Section 8.

2 Forecasting Recessions

2.1 Predicting with the Yield Spread

Mishkin (1990) and Mishkin et al. (1990) identify the links between the yield curve, inflation and real economic activity as direct derivatives of macroeconomic theory and the Fisher equation. Empirical support to this is offered by works such as Estrella and Mishkin (1997). On broader economic activity early work by Harvey (1988), Laurent et al. (1989) and Chen (1991) all extended the focus in the realm of economic and financial activity. Yield spreads were to be regarded as a powerful predictor of a large range of activities. Within the theory that had motivated this initial strength of application also lies an acknowledgement that the relationships are likely to be temporally variant and so more recent work has questioned the strength of the earlier findings (Ang et al., 2006, for example). A call emerges to expand the explanatory set beyond the yield curve to counter the changing relationship with output (Ang et al., 2006). For our purposes these additional links represent channels motivating an ability to predict recessions.

Estrella and Mishkin (1996, 1998) and others work on recession prediction notes the long association of yield spread inversions with the periods determined by the National Bureau for Economic Research (NBER) as being recession dates. This enduring relationship means that the treasury yield spreads continue to be the mainstay of works seeking to predict when economies will turn downward. Further when they do turn onto negative growth trajectories when will they recover? For this again the yield spread finds value. In addition to the inflation channel that drove the early work on yield spreads, further motivation for the value of the yield spread lies in its relationship to investor expectations about the future economic position of the market.Negative yield spreads intuitively work from a lack of faith in the delivery of short term returns versus those promised on longer investments. Movement in the yield curve may thus be associated with uncertainty about the future, uncertainty that then manifests in recession. Estrella and Mishkin (1996) remark that the yield spread already represents a forward looking expected change in interest rates.

There is also a growing body of work that concentrates on the point at which recession is entered. Stock and Watson (2014) comprehensive review of dating methodologies, dissecting identification into either average-then-date or date-then-average approaches. NBER processes of first considering series then looking for the date are examples of the former, whilst the increasing wealth of multivariate works discussed subsequently exemplify the latter. For most studies the NBER dates remain the target; they offer ready comparison with the existing literature and deliver an understood output to quantify improvement against.

Recession prediction modelling traditionally employs binary outcome models, particularly the probit model, to work out the probability that there will be a recession a fixed number of periods ahead. Such frameworks apply naturally where the outcome is a binary one; the economy is in recession or it is not. In this paper we also utilise a similar approach for comparability. A ready ability to construct predicted recession probabilities and an ease of interpretation of the contribution of in dependent variables thereto, further commend simpler approaches. However, Chauvet and Potter (2005) identifies an inherent flaw in the binary approach; information about recession status will affect the likelihood of continuation and is accrued after the time at which the predictions are made. A proposal for probability of first entering recession, and one for continuing are proposed (Chauvet and Potter, 2005). With just 7 recessions this further reduces the number of positive outcomes and presents a modelling challenge beyond this paper.

Binary statuses also lend well to classification modelling; machine learning algorithms are then well placed to build from observed recessions to identify the characteristics indicative of a future one. Berge (2015) early work on machine learning identified the challenge that classifications will always perform well when they predict no recession and there are very few recessions. Giusto and Piger (2017) argues the lack of a known data generating process in the yield spread means that traditional parametric modelling is not optimal. Natural disagreement between contemporary approaches and the NBER decisions are to be expected, as is the ability of data driven classification to determine faster than the NBER committee. However, on this Giusto and Piger (2017) remarks that the NBER focus is accuracy and that it is outperformance of alternative classification models that provides support for machine learning. Davig and Hall (2019) combines the classification and turning point literature to demonstrate how naive Bayesian techniques improve prediction of recession starts. In their work Berge (2015) argue that heterogeneity in agents will further muddy the delineation of periods, Davig and Hall (2019) responds to such with the suggestion that stronger penalties be applied if recessions are predicted far from actual dates. In what follows we note the propensity of all models to create small false positives of the type that could be penalised more strongly.

Yield spreads are inherently an indicator of investor expectations about the long-term state of the economy. Their subdivision into inflation expectations and associated real interest rate risk premia have offer more insight into why. Duffee (2018) and Benzoni et al. (2018) both identify investors worries about future inflation as being an important risk to which longer term interest rates must offer compensation. Through this channel those variables that would typically impact inflation will influence the interest rate expectations of traders. As the economy moves towards recession so the expectation for growth, and hence inflation, drops. A circle back to the Fisher equation and the literature that first prompted the use of the yield spread iteslf as a predictor Mishkin (1990),Estrella and Hardouvelis (1991) etc. is formed.

Contemporary work thus falls into three categories. Firstly, those which seek to augment the prediction set to obtain better fit through improved information. Secondly, those that adopt new methodologies for prediction and classification have demonstrated how insights from new approaches can improve forecasting. Finally, there is the work that nests the yield spread in the wider economic model to identify turning points and hence predict recession in this way. As reviewed each approach has its concerns and to date there is little to signal the end from the simple inverted yield curve rule of thumb.

2.2 Periodicity and Interest Rates

Log-periodicity as a forewarning of crashes is increasingly seen as of value to investors in stock markets (Sornette et al., 2001; Clark, 2004; Chang and Feigenbaum, 2006; Matsushita et al., 2006; Chang and Feigenbaum, 2008). This literature typically studies the build up to known crashes and then imposes a log-periodic wave upon the observed time series. Fitting such curves identifies an increasing period as the market moves into a period of "chaos" ahead of the crash[1]. By looking in isolation at particular crashes the work avoids the challenge of fitting a longer term measure, or understanding periodicity across the whole time series being studied. There is however a unanymity in the identification of periodic behaviour prior to crashes. Whether in the US markets as studied in early works (Sornette et al., 2001; Clark, 2004, for example) or the contemporary studies of Asian markets (Li, 2017; Ko et al., 2018), there is a consistency in the demonstration of increased periodicity ahead of crashes.

Observations of log-periodicity within the study of yield spreads have found its presence in the exit from, rather than entry to, recession. For this reason it is more likely that an absence of periodicity will be linked to recession forecasting. However, Benhabib et al. (2002) identifies a number of channels through which yield spread behaviour may become chaotic, or periodic, as a result of prevailing economic conditions. Zhou and Sornette (2004) notes a feedback from the log-periodicity of the stock markets going into the 2001 crash. In that case a link to interest rate dynamics is made to say that there was more periodic behaviour in 2002 precisely because of the 2001 dot-com crash. Both Wosnitza and Leker (2014) and Wosnitza and Sornette (2015) find log periodic behaviour within the credit spread, critically dismissing the possibility that these patterns could be random. We do identify periodic behaviour on the short scale in the work that follows in the same vein.

A feature in the periodicity literature is that the behaviour of the market changes in the build up to an event. In both the stock market and credit spread cases the periodicity is at its most intense close to the event being forecast; before in the case of market crashes and immediately afterwards in the case of yield spreads. However recession forecasting takes a longer term perspective, typically a horizon of 12 months (Estrella and Trubin, 2006). Consequently in looking to periodicity we are asking for longer term evidence within the series, seeking to identify the first seeds of behaviour rather than being free to watch the period all the way to the event.

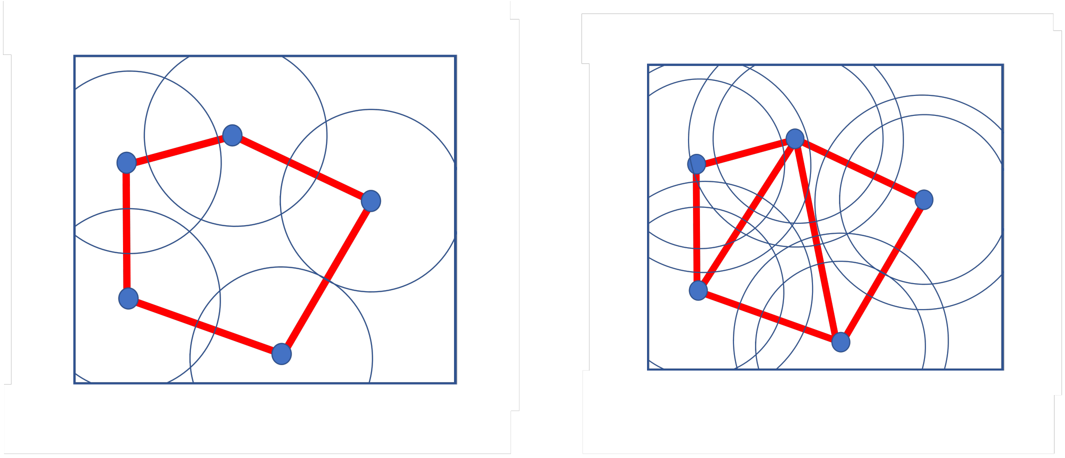
To date there is an identification of potential periodicity in parts of the yield spread series, an acknowledgement that peridoic behaviour in financial markets may be a pre-cursor to crashes, but no solid attempts to bring the two together. An identified temporal challenge is undoubtedly part of the motivation there for. This paper acts on that bridge between periodicity and recession forecasting as a preliminary exploration of the potential for periodicity to forecast further ahead.

3 Topology of Time Series

Gidea and Katz (2018) assertion that the persistence landscape can be used as a useful tool for crash prediction represents a natural launch point for reviewing links between yield spread topology and the prediction of recessions. In this section we define the persistence landscape and explain how that may then give rise to the norms. Secondly we explore how persistent homology can help detect periodicity following closely work by Dłotko et al. (2019). So doing we provide a toolkit to analyse time series. Finally we use empirical examples to show the robustness of this new addition to understanding. In so doing foundations for our analysis of the yield spread are laid.

[1] It must be noted here that we do not refer to "chaos" under its formal mathematical definition but as an intuitive nod to seeming irrationality in the behaviour of the series.

Figure 1: Persistent Homology Illustration



Filtration level b Filtration level d

Notes: Point cloud in two dimensions for illustration only. Data shown as large dots. All circles are of radius ϵ from the datapoint at their centre. Panel (a) shows a low filtration level at which the final edge forms between the two points on towards the lower right of the rectangle. This leaves an uncovered area in the middle of the shape where balls do not overlap but there are no points. This is referred to as a “hole”. Panel (b) shows a higher filtration level $\epsilon = d$ at which all the larger balls overlap. In persistent homology b represents the birth of the “hole” and d the death.

3.1 Persistent Homology and Landscapes

Data packages caring a number of independent information can be summarized as so called point clouds in a high dimensional space. As such they are carrying certain geometrical information which may not be apparent upon immediate observation[2].Persistent homology provides an understanding that point cloud through its shape. In what follows examples will be presented in two dimensions to facilitate discussion; such representations seem trivial in the sense that we can visualise without any TDA tools but do offer the intuition necessary to make the leap into multiple dimensional point clouds.

Intuitively let us imagine drawing ball of radius r centered on each point of the considered dataset. Where the balls overlap then we may regard those points as connected and we will denote it by plotting and edge between those points. As we increase r the more and more points will be considered similar, and therefore connected by an edge. Inherently the shape that is created will get larger and larger. Persistent homology provides us with the tools to formalise the process of identification delivering metrics with practical interpretation.

Consider the artificial example of Figure 1. In this case there are five points in the space. A collection of edges for the initial choice of the radius r is plotted in panel (a). This level is selected as that at which the final two points, those to the lower right of the plot, are connected so that the collection of edges closes up into a cycle. It can be seen that there is an area at the centre of the figure where the balls do not overlap. At this stage the persistent homology is able to identify this area and hence we term it a “hole”. The presented hole will be assigned a dimension 1 as it can be surrounded by a collection of (one dimensional) edges. Features in dimension 0 correspond to the connected components of the set. Had we plotted the step by step expansion of the circles the number of features (connected components) in dimension 0 would have gone from 5, the points, to 4 as the first pair connected, to 3, to 2, and finally to 1 as the full connections were made. Because

[2]In this regard we are familiar with visualising data in two, or three, dimensions but struggle with shape identification in more than that.

the hole forms at $\epsilon = b$ we term b as the birth of the feature. Panel (b) adds a second set of larger circles with radius d. Here we can see that the balls now overlap, the hole has disappeared. There are also two new edges formed as more of the balls overlap to create edges; the longer of these edges is the one that closed the hole. We term this closure of the hole as its death. Hence b is the birth and d is the death, $d - b$ is the life of the hole.

In a real dataset the number of features in both in dimension 0 and 1 will be much greater; there will be many births and deaths. All those features can be assembled into a single large set of points. From this we construct a persistence landscape. The landscape plots an iscocoles triangle centered at $(b + d)/2$ and of height $(b + d)/2$. By building these on top of each other we obtain a full landscape that can be enumerated level by level. The norm of this landscape, denoted as L, is the sum of areas of all those triangles. The larger the value of L the more features there are in the point cloud. Robustness is obtained by considering the life of features and placing a minimum requirement thereupon. Short-lived features would suggest that there was little to be read into their existence, they may simply be noise. Larger L comes from longer lives.

In the context of multiple variables the notion of many axes (dimensions) is well understood. For a univariate time series there is a need to think further about the construction of a point cloud. TDA analysis of time series begins from the construction of a point cloud in d dimensional space. A sliding window embedding (SWE) is performed on the series, defined by the number and size of jumps that are used. In this paper we abstract from a discussion of jump size and include all observations within the window[3]. This is a different approach to the one adopted in Gidea and Katz (2018), where they construct their point cloud from four time series and hence have a cloud of dimension four without any embedding. Their sliding window is simply a movement through time to capture a subsample of 50 points from the longer set.

An SWE captured over a fixed number of time periods creates the point cloud that informs our periodicity estimation. Defining the length of the embedding as m then for a series x the first line of the embedding will be $x_1, x_2, ..., x_m$. Moving forwards one jump the next point will be $x_{1+j}, x_{2+j}, ..., x_{m+j}$, the third line will be $x_{1+2j}, x_{2+2j}, ..., x_{m+2j}$. We continue this until the length of the window, w, which has $x_{1+(w-1)j}, x_{2+(w-1)j}, ..., x_{m+(w-1)j}$. The resulting w point, j dimensional, point cloud is then analysed using persistent homology in the way described above.

In this paper we will use an observation that existence of a large cycle in dimension 1 is a consequence of the periodicity of the function and we will be searching for that.

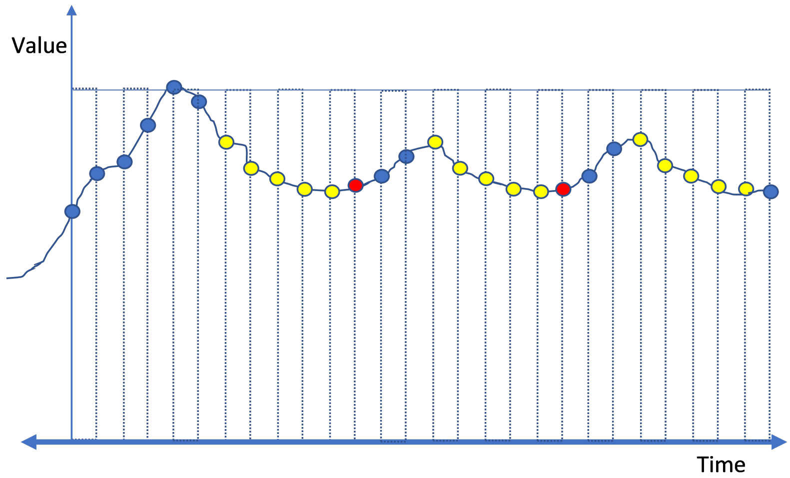
3.2 Periodicity

Within time series analysis periodicity, dynamics and persistence are all well defined concepts, as they are within the TDA literature. Here specific meaning is taken from the latter for mathematical consistency. Periodicity is thought of as the time between occurrences of the same motif within the time series no matter of the starting position. To observe the same shape once may be coincidence and so we require a second repetition to make any claim about (local) periodicity. Dynamics are understood through the movement of the time series through space; cycles for example form where a series moves away from a point and then returns. As the point in SWE is the representation of local motif in the time series, recurrent returning to the same elements of the point cloud is direct evidence of the periodicity of the input time series. Persistence is a measure of the extent to which observed topologies are robust to changes in the parameters informing measurement. Where the features are identified consistently then we can regard the patterns as persistent, whilst rarely observed phenomena are not persistent. All of these concepts are illustrated within Figure 2

As demonstrated in Figure 2 a recurring signature in the SWE would be identified as points $y_t, y_{t+1}, y_{t+2}, ...$ recur at $y_{t+h}, y_{t+h+1}, Y_{t+h+2}, ...$ and again at $y_{t+2h}, y_{t+2h+1}, ...$ Through identifi-

[3]However, we are able to prove that using jumps does not invalidate the approach. Using more points simply means more computational time is required. In this paper the additional processing of using a unitary jump size is insignificant deterrent from maintaining all of the information in the series.

Figure 2: Time Series Properties



Notes: Segment of an artificial time series sketched to illustrate key elements as understood in this paper. Lighter (yellow) shaded points represent a recurring motif used to extract period. Darker shaded (red) points illustrate a motif within the series which only recurs once and is regarded as not being persistent. Taking the lines as each representing one unit of time, the period of this series would be estimated as 8 units.

cation of this we identify h^4. In most cases the time series will contain noise, meaning that the identification of periodicity is not as simple as pattern matching. Our process for identification therefore takes the SWE and overlays a series of landmark points. For a cycle with some noise an SWE with cyclicity would be an annulus, the time series going round the circle to return periodically approximately to its original value. Intuitively therefore if we take landmark points around the annulus it will pass these in order and at the same period. Landmark points are created by taking the start point of the series then locating the point furthest therefrom. Next we identify the point furthest from these two landmarks and keep going until a pre-defined number of landmarks are found. Figure 3 illustrates this process. Note that the labelling is done once points are identified such that the sequence is intuitive; the actual labelling of points is not important as the algorithm recognises the points locations within the cloud.

3.3 Periodicity and Artificial Series

To demonstrate the robustness of the method we will employ benchmark functions consisting of four instances of sinusoidal functions with an increasing level of noise. Noise is simply a random sample from a uniform distribution $[0, no]$ in which no is equal to 0,1,2 and 3. We choose the uniform distribution as this has a high tail probability, ensuring that the observed values are away from the underlying function more than would be the case for a mean centered distribution. Evaluating the period for the sinus function Table 1 presents the results for a series of coefficients on the number of jumps and the number of points in the subsequence, parameters j and w of the SWE respectively[5]. For this table we use 10 landmark points.

In this section we use $sinx$ sampled from 0 with the step size 0.1 such that the period of the function is $10 \times 2\pi \approx 63$. Onto the function we add a uniform noise term of magnitude 0, 1, 2 and 3.

[4] A full exposition of the underlying mathematics of the process adopted here can be found in [DLOTKO2019].

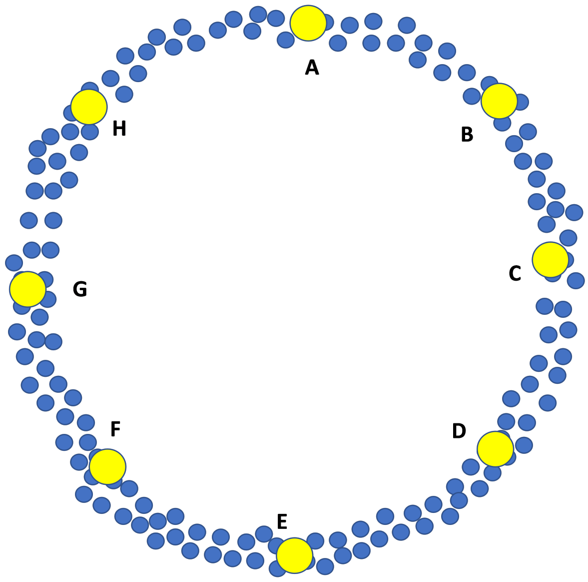
[5] A full table for each noise level, and each number of landscapes is available upon request. We do not include these here for brevity.

Table 1: Period Estimates for Sinus Function ($sinx$) with Noise

Noise	Jumps (j)	Length of Sliding Window (w) 120	160	170	180	190	200	250	300
Noise 0	50	58.92	62.94	62.98	63.00	63.00	63.00	63.00	63.00
		(15.02)	(0.23)	(0.13)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	60	55.17	62.95	63.00	63.00	63.00	63.00	63.00	63.00
		(19.87)	(0.21)	(0.07)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	70	53.71	62.91	63.00	63.00	63.00	63.00	63.00	63.00
		(21.65)	(0.45)	(0.07)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	100	47.26	55.53	57.95	60.58	62.90	62.96	63.00	63.00
		(26.82)	(20.23)	(16.93)	(11.80)	(0.39)	(0.20)	(0.00)	(0.00)
	150		41.25	43.04	45.00	47.14	49.49	62.85	63.00
			(29.79)	(29.13)	(28.26)	(27.12)	(25.60)	(0.36)	(0.00)
	200							46.62	62.61
								(27.26)	(0.49)
Noise 1	50	58.75	63.27	63.38	63.54	63.64	63.64	63.78	64.28
		(14.30)	(0.85)	(0.81)	(0.77)	(0.70)	(0.63)	(0.48)	(0.45)
	60	56.43	63.05	63.30	63.42	63.53	63.56	63.80	63.99
		(18.15)	(0.76)	(0.65)	(0.60)	(0.51)	(0.51)	(0.40)	(0.10)
	70	54.21	63.00	63.22	63.40	63.50	63.58	63.83	63.98
		(21.21)	(0.75)	(0.63)	(0.53)	(0.51)	(0.50)	(0.37)	(0.14)
	100	47.27	55.80	58.32	61.03	63.12	63.28	63.93	64.00
		(26.58)	(19.95)	(16.56)	(11.17)	(1.34)	(1.25)	(0.25)	(0.00)
	150		43.34	45.29	47.40	49.70	52.24	63.83	64.00
			(29.29)	(28.50)	(27.44)	(26.04)	(24.14)	(0.41)	(0.00)
	200							45.00	63.18
								(28.63)	(0.39)
Noise 2	50	55.89	60.50	60.73	61.01	61.33	61.52	62.20	63.15
		(13.48)	(1.81)	(1.79)	(1.73)	(1.66)	(1.58)	(1.13)	(0.76)
	60	53.76	61.39	61.76	62.25	62.39	62.62	63.01	63.41
		(17.54)	(1.82)	(1.48)	(1.10)	(0.95)	(0.87)	(0.69)	(0.67)
	70	53.23	61.83	62.23	62.55	62.83	63.00	63.30	63.07
		(19.04)	(1.95)	(1.44)	(0.99)	(0.69)	(0.68)	(0.59)	(0.29)
	100	47.73	56.49	59.06	61.80	62.62	62.86	63.31	63.67
		(24.99)	(18.05)	(13.98)	(6.15)	(1.05)	(0.63)	(0.48)	(0.47)
	150		42.52	44.47	46.52	48.70	51.02	62.28	63.87
			(28.05)	(27.26)	(26.20)	(24.86)	(23.21)	(2.79)	(0.34)
	200							46.69	63.37
								(26.94)	(0.65)
Noise 3	50	52.15	53.93	54.25	54.38	54.59	54.77	55.96	56.86
		(4.40)	(1.41)	(1.34)	(1.31)	(1.32)	(1.18)	(1.09)	(0.65)
	60	50.10	54.37	54.74	54.88	54.92	54.99	55.73	56.42
		(11.00)	(1.55)	(1.17)	(1.11)	(1.06)	(1.14)	(1.27)	(1.15)
	70	47.75	53.90	54.39	54.74	54.90	55.23	56.14	56.94
		(14.72)	(1.89)	(1.83)	(1.66)	(1.55)	(1.34)	(1.00)	(0.66)
	100	43.78	51.51	53.69	55.20	55.45	55.59	56.25	56.21
		(19.60)	(12.53)	(8.11)	(1.29)	(1.41)	(1.45)	(1.02)	(0.57)
	150		38.72	40.38	42.07	43.82	45.70	55.33	56.00
			(23.66)	(23.06)	(22.34)	(21.29)	(19.89)	(1.44)	(0.00)
	200							42.64	56.15
								(22.14)	(0.61)

Notes: Time series estimated as $sinx$ for $x \in [1, 40]$ at increments of 0.01. True period is $10 \times 2\pi$ or 63 to nearest whole number. Values report average of period estimates for observations between w and the length of the time series T inclusive. Standard deviations of the estimates over the range $[w, T]$ are included in parentheses. All jumps are of size 1. 10 Landmarks are used in the estimation of the periodicity.

Figure 3: Landmark Construction



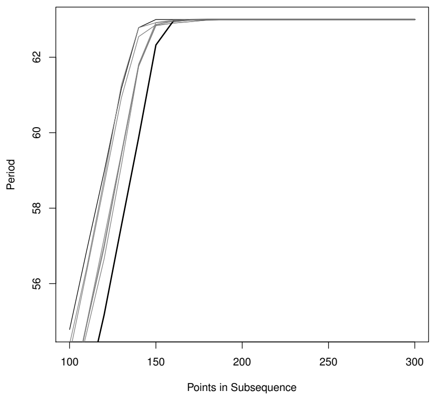
Notes: Plot shows eight points around a sliding window embedding of a time series. A periodic function would move around the circle through all landmark points, repeating with every period. A sequence constructed from the landmarks would be expected as A, B, C, D, E, F, G, H, A, B, C, D, E, F ,G ,H, A, ... , H, A... The time between A's, or B's, or C's etc. should then inform on the period.

Table 1 shows that estimates converge towards the true periodicity as the numbers of points in the subsequence increase. Formally to identify the period we need two periods to be completed on top of the number of jumps The intuitive reason for that is because to be able to claim that the considered time series repeat itself, at least one full repetition is required. As expected the period is correctly identified for $j = 180$ when $j = 50$ and noise is 0. Approaching this we see correct identification of the period with standard deviations. As the noise level increases the estimation is harder, although we see estimation close to the underlying period for both noise 1 and 2. As the noise reaches 3 the estimates are lower, this is because the measure is how quickly the function returns to the original landmark; noise means this will happen faster than it would have done otherwise.

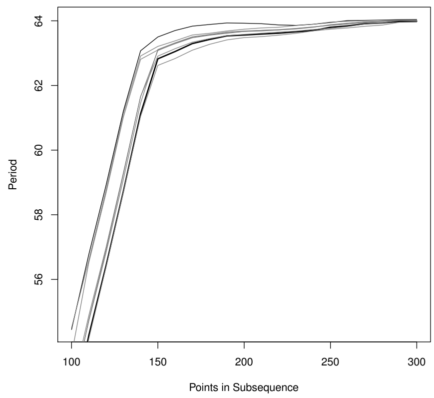
Landmarks are pivotal to the identification of the period and are chosen according to the preferences of the user. However, we can demonstrate a robustness to the selection provided that the number does not get too close to the number of points in the subsequence. Recognising the correct identification of the period when $j = 60$ we plot the estimates of the period for each noise level when there are between 5 and 12 landmarks. Figure 4 provides the results with the horizontal axis representing the number of points in the subsequence w. Observed tightness of the lines shows the robustness of the estimates, with the lower numbers of landmarks coming closer to estimating the known true period. Figure 4 also demonstrates the effect of noise through the comparison across panels (a) to (d), lower period estimates coming from the potential to return to landmarks much quicker in the presence of noise.

Robustness to noise, and the consistency of estimates across parameters, thus commend the TDA approach for the estimation of periodicity. However, there is no evidence of a continuous period within yield spread data, and hence it is the speed at which estimates are formed that is the main advantage demonstrated in this example.

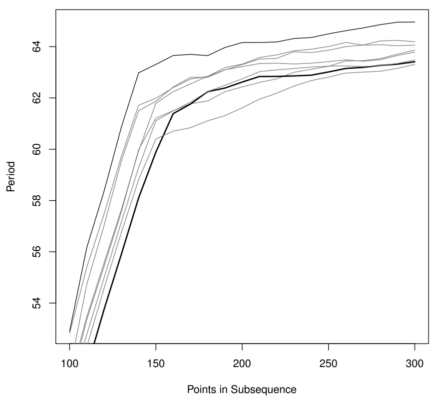
Figure 4: Estimation of Period for Sinus Function with Noise



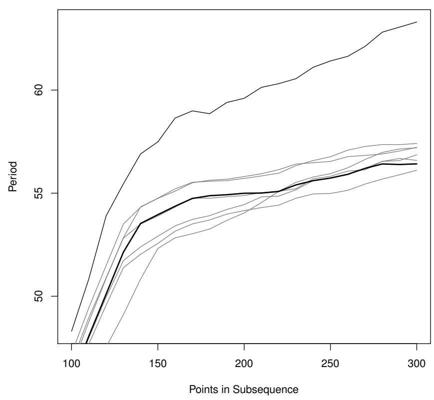
Panel (a): Noise 0



Panel (b): Noise 1



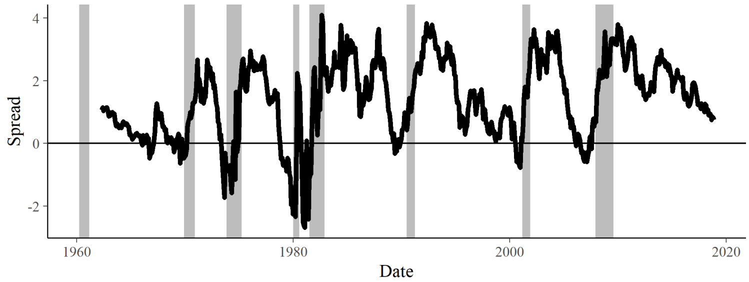
Panel (c): Noise 2



Panel (d): Noise 3

Notes: Lines plot the estimated period for the function *sinx* constructed at intervals of 0.01 such that the true period is $10 \times 2\pi \approxeq 63$. Thin black lines show $l = 5$, thick black lines show $l = 10$, with grey lines added for $l = 6, 7, 8, 9, 11, 12$. Horizontal axes show number of points in subsequence, w.

Figure 5: US Yield Spread (1961-2019)



Notes: Thick line denotes yield spread calculated according to equation 2, with adjustment to the three month rate following equation 1. Grey areas indicate the NBER recession dates.

3.4 Selecting Parameters

In the context of series with known period it is easy to parameterise the model. When the true period is unknown, or is varying, it is required that the researcher select appropriate numbers of points. The choice of w will impact upon the ability of the detection to locate longer patterns within the data. In selecting a window size of 100 and a number of jumps of 50 this paper is seeking to identify patterns of up to one month, one month being approximately 22 trading days. As specified we are not able to recover the economic cycle as the overriding period; this could be achieved by changing to a longer window and adding larger jumps to maintain a manageable point cloud. Finally a sensitivity analysis should be conducted to ensure robustness.

4 Data

Recognising the longevity of the U.S. data as critical to the understanding of investors renders the Treasury Yields for the U.S. the optimal test bed for our analysis. What follows would readily extend to other countries and represent an interesting extension. A benchmark for all recession forecasts remains the yield curve inversion, captured in the yield spread. Here we follow Estrella and Trubin (2006) to adjust the three month rate. Equation (1) defines the adjustment.

$$r_{tB3m} = 100 \times \frac{365(r_{tT23m}/100)}{360 - 91(r_{tT23m}/100)} \tag{1}$$

where, at time t, r_{tT23m} is the 3-month treasury rate from the secondary market rather than the primary rate quoted in the H.15 release. In this way the r_{tB3m} is bond-equivalent and offers "maximum robustness in predicting US recessions" (Estrella and Trubin, 2006, p.3). The yield spread is then calculated from the 10 year tresury, r_{tT10y} as:

$$Spread_t = r_{tT10y} - r_{tB3m} \tag{2}$$

Our resulting yield spread series is plotted in Figure 5. Grey boxes shade the recessions.Recession timing is taken from the National Bureau for Economic Research (NBER), with 7 recessions falling within the time frame of our data coverage. These recessions are visible in Figure 5. As established there are a series of inversions prior to recessions, with a single false positive where the yield curve inverted in the 1960s but there was no recession.

Table 2: TDA Monthly Measure Summary

Measure	Aggregation	Mean	s.d.	Min	Max
Panel (a): Spread topology:					
L^1 norm	Max	0.066	0.097	0	1
	Avg	0.030	0.079	0	0.948
Period	Max	2.269	9.427	0	72
	Avg	0.497	0.273	0	35.10
Panel (a): Demeaned spread topology:					
L^1 norm	Max	0.065	0.102	0	1
	Avg	0.050	0.086	0	0.956
Period	Max	2.094	9.236	0	74
	Avg	0.439	2.391	0	29.67

Notes: TDA analysis performed using 50 jumps with a subsequence of 100 points and 10 landmark points ($j = 50, w = 100, l = 10$). Summary statistics reported based on monthly compilation from daily data. Aggregation is either Max (maximum for the month) or Avg (average for measures obtained in that month).

5 Topology of US Yield Spreads

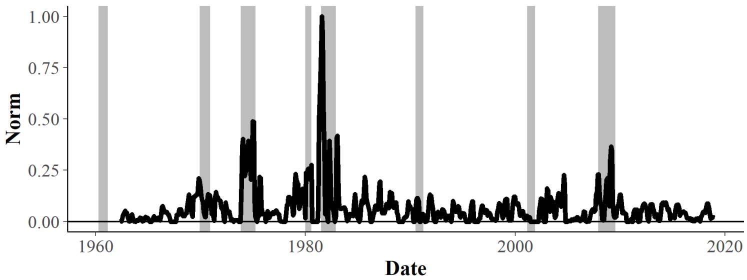
To demonstrate the theory of Section 3 for yield spread data we provide a more detailed look at the topology of the US yield spread. Table 2 provides summary statistics of the series revelaling slight differentials between the use of the spread, or the demeaned spread, for constructing the persistent homology. We see that there are a number of features forming through the time series, with the L^1 norm in panel (a) of Figure 6 showing a high degree of volatility throughout. The largest peaks in the range occur during the recessions of the mid 1970s, early 1980s and the global financial crisis of 2007-09. Interestingly we do not see any movement of the norm during the 2001 recession, whilst the recession of the early 1990s does not influence the norm series by much. Periodicity estimates likewise peak near the recessions, with some large peaks occurring in the time prior to the 2001 recession. There are a number of peaks in more recent years; these may be signals of impending recession but as the yield spread remains positive it is questionable whether this is the case.

From the period plot limited evidence of the post recession periodicity is found. Arguably it is only the recession of the early 1980s that led to any form of periodicity in the aftermath. As captured by the TDA code there is limited support for the log periodicity discussed in Clark (2004), Wosnitza and Sornette (2015) and others. However, to capture this better we can subtract a moving average of the spread from the series. In this way we can identify any periodic behaviour around the moving average. Figure 7 results.

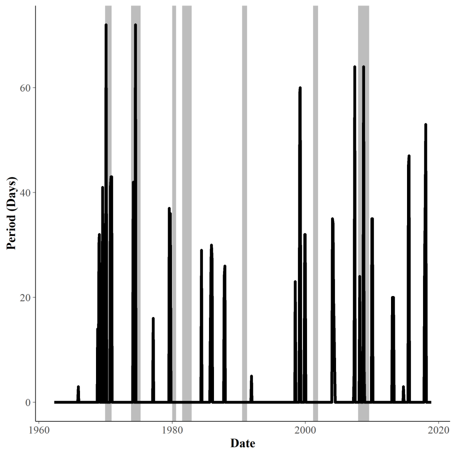
In this case the norm spikes are very similar and again concentrate around the recessions. Unlike the spread homology, which did not peak in the 2001 recession, we note that the de-meaned spread does have a small peak within the NBER recession period. Of more interest from Figure 7 is the periodicity in panel (b). Here there are notable peaks in the post recession period as foretold by the log-periodic literature. However there is also evidence of periodicity on the way down towards the crashes;note particular spikes ahead of the 2008 global financial crisis and the 1990s recession. If a criticism of the spread is that it did not predict the 1990 recession then the addition of periodicity has the potential to address that.

From the illustration there is potential for the TDA L^1 values to provide forecasting ability for the continuation of recessions, whilst the periodicity is unlikely to contribute to the overall improvement of recession prediction. When we subtract the one year moving average of the spread periodicity does look like a potentially useful measure, particularly for those recessions where the spread alone has been noted to forecast poorly.

Figure 6: Topology of the US Yield Spread (1961-2019)



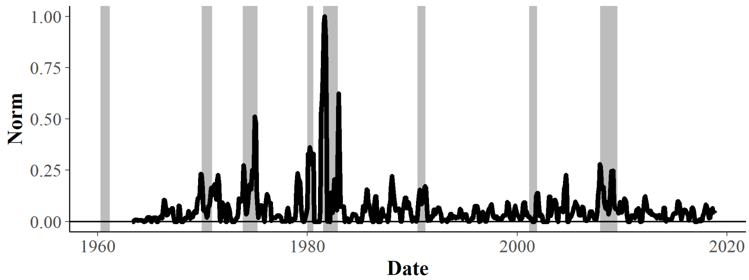
Panel (a): Dominant Interval Length (L^1 Norm)



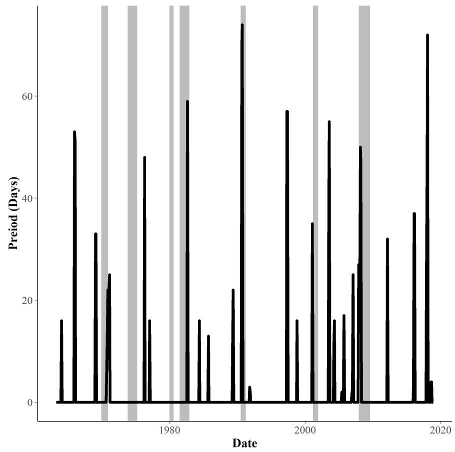
Panel (b): Periodicity estimate

Notes: TDA analysis performed using 50 jumps with a subsequence of 100 points and 10 landmark points ($j = 50, w = 100, l = 10$). Grey areas denote the NBER recession dates.

Figure 7: Topology of the De-meaned US Yield Spread (1961-2019)



Panel (a): Dominant Interval Length (L^1 Norm)



Panel (b): Periodicity estimate

Notes: TDA analysis performed using 50 jumps with a subsequence of 100 points and 10 landmark points ($j = 50, w = 100, l = 10$). Grey areas denote the NBER recession dates.

6 Forecasting Recessions

To maintain comparability with the extant literature we adopt a probit approach to the modelling of recession probability, this follows directly the work of Estrella and Mishkin (1996, 1998). A binary outcome R is defined which takes the value 1 in any month which is classified as a recession by the NBER. As we seek to forecast we define R_{t+p} as the recession dummy p periods forward from time t. Taking the TDA variables $DINT_t$ and PER_t as being the values of $DINT$ and PER in month t, as augmentation to the spread, $spread_t$, we fit:

$$R_{t+p} = F(\alpha + \beta_1 spread_t + \beta_2 DINT_t + \beta_3 PER_t) \quad (3)$$

with the function $F()$ being the cumulative normal distribution function. In this paper we specify three versions of the model, first with only the yield spread ($\beta_2 = \beta_3 = 0$). Next, inspired by the work of Gidea and Katz (2018) we introduce only the L^1 norm of the persistence landscape, $DINT$. This second specification imposes $\beta_3 = 0$. Finally we estimate a model with both TDA variables. Models are estimated for both the topology of the spread and the de-meaned spread. Likewise we fit both the maximum and average values for the TDA variables within the month. Necessarily the paper only features a subset of the results with full tables available in a supplementary appendix.

6.1 Spread Topology: Maximum Monthly TDA Values

As our first case we consider the topology of the yield spread as plotted in section 5. Table 3 contains estimates for a selected p range. Across four panels we report the three specifications of 3 and finally a series of residual deviance tests to aid model selection. An immediate message from Table 3 is the almost complete significance of the yield spread as a predictor, only at the 18-month horizon is the L^1 norm not significant. A strong significance is assigned to the constant term in the model in all cases recognising there would be scope to add more explanatory variables. Once introduced in the third model we see that the period estimate provided by our novel approach does not produce significance. This would be consistent with the literature identifying log periodicity as something that occurrs post recession rather than in the years preceding. Where significant high norms are associated with high recession probabilty; this is in line with the "chaotic" association between similar increases and crash probability in cryptocurrencies and stock markets (Gidea et al., 2018; Gidea and Katz, 2018).

Comparisons of model fit are provided through an exploration of the residual deviance of the models. Both the L^1 norm augmented model, and the full model with both norms and periods, outperform the basic spread model. The former holds throughout, whilst the latter is true at shorter horizons. Given the lack of significance inclusion in the model will naturally bring down the prediction accuracy.

Although the coefficients from the table give some impression of the way that the variables impact upon the probability of recession it is more instructive to plot that predicted probability as a visualisation of model fit. In Figure 8 we use a black line to plot Model 1, a red line for Model 2 and a blue line for Model 3.

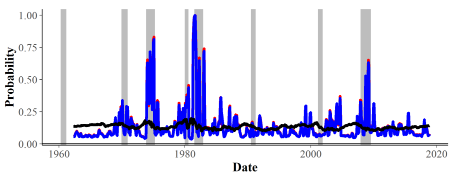
Figure 8 involves taking the fitted probabilities from the probit models and aligning them with the period for which they are designed to predict recession. We plot the predicted probability from model 1 in black, model 2 in red and model 3 in blue. Across the four panels we can see that the models usually spike at the right places. Perhaps unsurprisingly it is the double-dip recession of the early 1980s where the models do best; here they have the first recession in the dataset for predicting the second. However, almost all of the NBER recessions are predicted at a probability of 50% or higher. For the longer term forecasts all recessions are identified with a probability of 25% or higher, but this level is only marginal than some false positives. All models suggest there could have been a recession in the years leading up to the actual 2001 downturn, but the strength of the suggestion does increase relative to actual recessions as the time horizon for prediction extends.

Table 3: Recession Forecast Models: Spread Topology (Maximum)

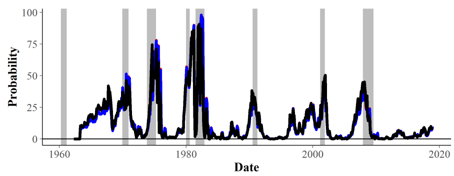
	Months ahead to forecast (p) 0	1	2	3	6	12	18	24
Panel (a): Model 1 spread only								
Constant	-1.023***	-0.927***	-0.846***	-0.76***	-0.629***	-0.523***	-0.564***	-0.710***
	(0.089)	(0.086)	(0.083)	(0.08)	(0.078)	(0.086)	(0.082)	(0.080)
Spread	-0.063	-0.142**	-0.216***	-0.304***	-0.473***	-0.684***	-0.596***	-0.367***
	(0.050)	(0.051)	(0.053)	(0.055)	(0.064)	(0.079)	(0.059)	(0.046)
Panel (b): Model 2 with $L^1 norm$								
Constant	-1.61***	-1.525***	-1.465***	-1.338***	-1.055***	-0.689***	-0.503***	-0.591***
	(0.115)	(0.118)	(0.119)	(0.119)	(0.107)	(0.099)	(0.097)	(0.098)
Spread	0.052	-0.039	-0.131*	-0.246***	-0.43***	-0.643***	-0.612***	-0.392***
	(0.047)	(0.052)	(0.052)	(0.053)	(0.064)	(0.078)	(0.067)	(0.052)
L^1 norm	5.096***	5.644***	6.386***	6.421***	4.878***	1.822**	-0.668	-1.452*
	(1.238)	(1.135)	(0.987)	(0.941)	(0.891)	(0.642)	(0.583)	(0.600)
Panel (c): Model 3 with $L^1 norm$ and period								
Constant	-1.629***	-1.534***	-1.467***	-1.334***	-1.066***	-0.693***	-0.499***	-0.596***
	(0.114)	(0.117)	(0.119)	(0.119)	(0.108)	(0.099)	(0.097)	(0.098)
Spread	0.053	-0.039	-0.131*	-0.246***	-0.43***	-0.642***	-0.613***	-0.391***
	(0.047)	(0.052)	(0.052)	(0.053)	(0.063)	(0.078)	(0.067)	(0.052)
L^1 norm	4.986***	5.562***	6.366***	6.467***	4.746***	1.79**	-0.650	-1.501*
	(1.210)	(1.132)	(1.002)	(0.967)	(0.88)	(0.649)	(0.589)	(0.623)
Period	0.009	0.005	0.001	-0.002	0.007	0.002	-0.002	0.003
	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.006)	(0.007)	(0.006)
Panel (d): Model comparison (residual deviance)								
Model 1	533.455	526.047	514.067	494.697	448.108	389.032	418.086	481.823
Model 2	460.895	446.333	425.155	410.261	399.043	381.598	416.933	477.217
Model 3	458.726	445.636	425.129	410.148	397.661	381.459	416.848	477.021
2 vs 1	-72.559***	-79.714***	-88.912***	-84.436***	-49.065***	-7.433**	-1.154	-4.605*
3 vs 1	-74.729***	-80.411***	-88.937***	-84.549***	-50.447***	-0.14	-0.085	-0.196
3 vs 2	-2.169	-0.697	-0.025	-0.113	-1.382	-0.14	-0.085	-0.196

Notes: TDA calculated using 50 jumps of size 1 with a sliding window embedding of 100 observations and 10 landmark points. Spread is constructed on a bond equivalent basis following Estrella and Trubin (2006) and Demeaned by subtracting the one year moving average. Significance denoted by * - 5%, ** - 1% and *** - 0.1%.

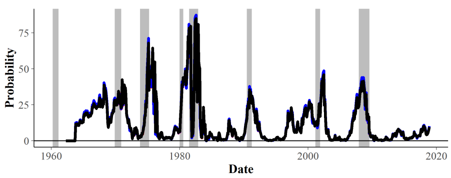
Figure 8: Predicted Recession Probability: Model with Spread and Monthly Maximum Topology



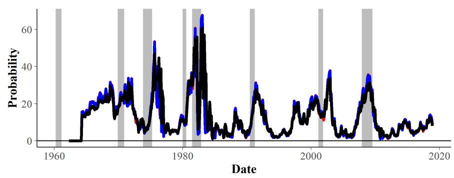
Panel (a): 0 Months



Panel (b): 12 Months ahead



Panel (c): 18 Months ahead



Panel (d): 24 Months ahead

Notes: TDA calculated using 50 jumps of size 1 with a sliding window embedding of 100 observations and 10 landmark points. Spread is constructed on a bond equivalent basis following Estrella and Trubin (2006). Probability reports the predicted probability of recession in given month. Predicted probabilities are moved forward to relate to the period $t + p$ being predicted and do not correspond to the time of the data used to form the prediction. Black lines denote Model 1, red lines Model 2, and blue lines plot the predicted probability from Model 3.

Table 4: Recession Forecast Models: Spread Topology (Average)

	Months ahead to forecast (p) 0	1	2	3	6	12	18	24
Panel (a): Model 1 spread only								
Constant	-1.023***	-0.927***	-0.846***	-0.76***	-0.629***	-0.523***	-0.564***	-0.710***
	(0.089)	(0.086)	(0.083)	(0.08)	(0.078)	(0.086)	(0.082)	(0.080)
Spread	-0.063	-0.142**	-0.216***	-0.304***	-0.473***	-0.684***	-0.596***	-0.367***
	(0.050)	(0.051)	(0.053)	(0.055)	(0.064)	(0.079)	(0.059)	(0.046)
Panel (b): Model 2 with $L^1 norm$								
Constant	-1.572***	-1.491***	-1.45***	-1.323***	-1.028***	-0.666***	-0.485***	-0.586***
	(0.114)	(0.12)	(0.118)	(0.115)	(0.104)	(0.097)	(0.094)	(0.096)
Spread	0.051	-0.037	-0.13*	-0.237***	-0.42***	-0.643***	-0.62***	-0.396***
	(0.046)	(0.051)	(0.052)	(0.052)	(0.064)	(0.077)	(0.067)	(0.052)
L^1 norm	6.027***	6.741***	8.051***	7.945***	5.72***	1.954*	-1.093	-1.910**
	(1.516)	(1.49)	(0.973)	(1.011)	(1.066)	(0.775)	(0.664)	(0.702)
Panel (c): Model 3 with $L^1 norm$ and period								
Constant	-1.573***	-1.491***	-1.451***	-1.324***	-1.027***	-0.665***	-0.484***	-0.588***
	(0.113)	(0.12)	(0.119)	(0.115)	(0.104)	(0.097)	(0.094)	(0.096)
Spread	0.049	-0.039	-0.129*	-0.236***	-0.426***	-0.642***	-0.62***	-0.397***
	(0.046)	(0.051)	(0.052)	(0.052)	(0.063)	(0.078)	(0.067)	(0.052)
L^1 norm	5.922***	6.676***	8.13***	8.033***	5.494***	1.988*	-1.082	-2.021**
	(1.520)	(1.519)	(0.997)	(1.039)	(1.056)	(0.796)	(0.669)	(0.737)
Period	0.016	0.008	-0.008	-0.008	0.026	-0.007	-0.003	0.017
	(0.017)	(0.018)	(0.022)	(0.026)	(0.019)	(0.028)	(0.029)	(0.019)
Panel (d): Model comparison (residual deviance)								
Model 1	533.455	526.047	514.067	494.697	448.108	389.032	418.086	481.823
Model 2	468.524	453.294	429.707	415.479	405.113	383.288	416.07	476.716
Model 3	467.912	453.136	429.581	415.324	403.605	383.232	416.059	476.302
2 vs 1	-64.93***	-72.752***	-84.36***	-79.219***	-42.996***	-5.744*	-2.016	-5.106*
3 vs 1		-65.543***	-72.911***	-84.485***	-79.373***	-44.503***	-0.056	-0.011
-0.414								
3 vs 2	-0.613	-0.159	-0.126	-0.154	-1.508	-0.056	-0.011	-0.414

Notes: TDA calculated using 50 jumps of size 1 with a sliding window embedding of 100 observations and 10 landmark points. Spread is constructed on a bond equivalent basis following Estrella and Trubin (2006) and Demeaned by subtracting the one year moving average. Significance denoted by * - 5%, ** - 1% and *** - 0.1%.

There is a suggestion from panel (a) of Figure 8 that the properties of the actual time periods of recession are very different from those one year earlier. In this way we see the TDA based models spike in the recessions in panel (a), but do not see any recession prediction from a spread only model. Understanding such differences is a motivation for the multivariate modelling that brings in additional macroeconomic and financial variables. That TDA is capturing much of this is an encouragement for its wider applicability.

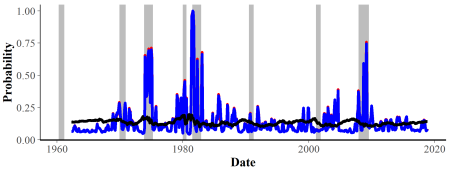
The evidence from the first augmentation of the yield spread with its TDA properties is that there are significant improvements to be made to overall forecasting. Most of the horizons studied here show significant reduced residual deviance arising from the extra variables. A similar position is found over the horizons not featured herein.

6.2 Spread Topology: Average Monthly TDA Values

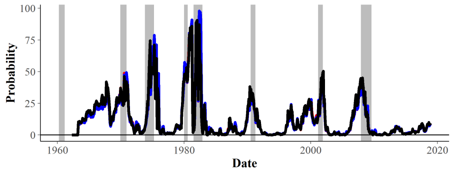
Table 4 uses the average values for the two TDA measure in the place of the maximums used in Table 3. Note that model 1 does not contain any TDA variables and so panel (a) is identical in both tables. We include panel(a) here for easier reference. Using the average values of the L^1 norm results in higher β_2 estimates in absolute terms. This is partially to be expected since the average values for any month are necessarily lower than the maximum. However, the continued significance is a sign that the maxima were not simply the result of one particular days SWE. Panel (d) shows that Models 2 and 3 continue to be regarded as better fits under the residual deviance test than the basic, spread only, Model 1. Critically at the 12-month ahead forecast range the average L^1 norm helps Model 2 to be preferred to Model 1 at the 12-month horizon. Since the acid test is performance 1 year ahead the improvement offered by TDA is something of large interest.

Figure 9 is similar to Figure 8 as might be expected from the table comparison. We see again that at the 0-month ahead level the code which embeds TDA identifies the recession much more accurately, the blue spike in panel (a) being testament there to. Likewise when we move to 24-months ahead there is a large amount of blue above the black line. Again the 18-month ahead shows a number of times for which Model 3 is predicting a positive recession probability in a subset of

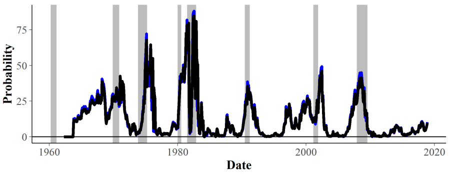
Figure 9: Predicted Recession Probability: Model with Spread and Monthly Average Topology



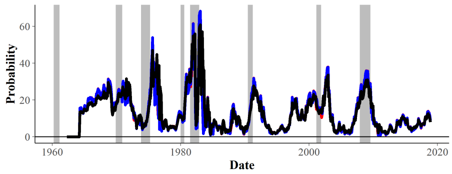
Panel (a): 0 Months



Panel (b): 12 Months ahead



Panel (c): 18 Months ahead



Panel (d): 24 Months ahead

Notes: TDA calculated using 50 jumps of size 1 with a sliding window embedding of 100 observations and 10 landmark points. Spread is constructed on a bond equivalent basis following Estrella and Trubin (2006). Probability reports the predicted probability of recession in given month. Predicted probabilities are moved forward to relate to the period $t + p$ being predicted and do not correspond to the time of the data used to form the prediction. Black lines denote Model 1, red lines Model 2, and blue lines plot the predicted probability from Model 3.

periods. The message from the average topology is thus in agreement with the maximal approach.

6.3 Demeaned Spread Topology: Average Monthly TDA Values

In the log-periodicity literature a wave of reducing period is placed over the time series this is done in a way which removes the trend from the series. Hence as a second consideration we look at a time series where the average spread for the past years trading has been subtracted[6]. We have already seen that there are differences between the de-meaned series and the original spread, particularly between the 2001 and 2008 recessions. We now consider how that has transferred into recesion prediction.

Table 5 shows a key difference with the corresponding Table 3 shows the biggest changes come in model 3, where the estimated period gains significance in the 6-month, 12-month and 24-month ahead forecast models. The coefficients to be broadly similar in models 1 and 2, with the TDA augmented models again providing a significantly lower residual deviance. From the increased significance of period comes an improvement to the model fit, and hence reduction in the residual deviance, of model 3.

Figure 10 also shares many similarities with it's spread topology counterpart (Figure 8). The 1990 recession is predicted with a higher probability when the de-meaned spread topoglogy is used, this following from the identification of periodic behavior being later than they were for the spread. Recession in the early 1990s is the one for which the least accurate forecasting has been found Estrella and Trubin (2006) and hence this marginal improvement is of interest. Mid 1960s false positive remains with all models predicting a recession in 1968.

6.4 Demeaned Spread Topology: Average Monthly TDA Values

As a final example we take the average monthly TDA values for the regessors, constructing them from the demeaned spread once more. Motivation for so doing is identical to our reasoning for

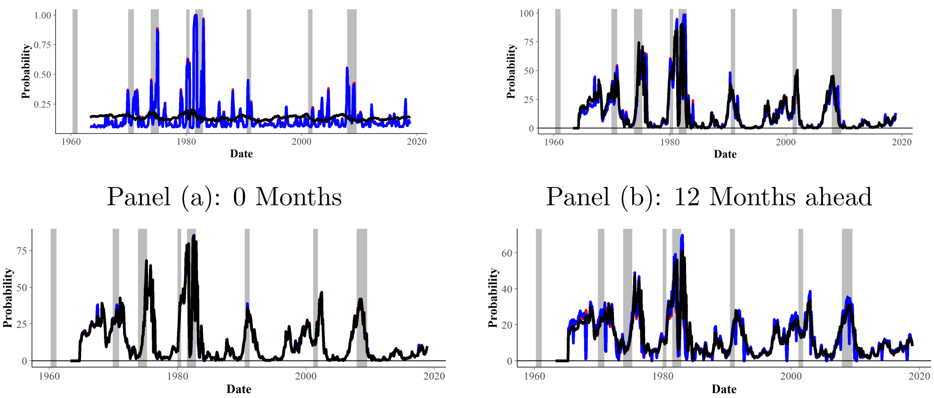
[6]The choice of 1 year preserves any annual seasonality, but alternatives would have served a similar purpose.

Table 5: Recession Forecast Models: Demeaned Spread Topology (Maximum)

	Months ahead to forecast (p)							
	0	1	2	3	6	12	18	24
Panel (a): Model 1 spread only								
Constant	-1.009***	-0.914***	-0.833***	-0.748***	-0.618***	-0.513***	-0.553***	-0.699***
	(0.089)	(0.086)	(0.083)	(0.08)	(0.078)	(0.086)	(0.082)	(0.080)
Spread	-0.065	-0.143**	-0.216***	-0.304***	-0.471***	-0.679***	-0.591***	-0.366***
	(0.049)	(0.05)	(0.052)	(0.054)	(0.063)	(0.078)	(0.058)	(0.046)
Panel (b): Model 2 with $L^1 norm$								
Constant	-1.582***	-1.462***	-1.362***	-1.23***	-0.903***	-0.704***	-0.539***	-0.561***
	(0.134)	(0.125)	(0.124)	(0.118)	(0.094)	(0.098)	(0.099)	(0.105)
Spread	0.021	-0.071	-0.164**	-0.269***	-0.436***	-0.646***	-0.595***	-0.393***
	(0.053)	(0.054)	(0.053)	(0.053)	(0.065)	(0.079)	(0.063)	(0.053)
L^1 norm	5.457***	5.602***	5.868***	5.719***	3.342***	2.244**	-0.174	-1.832*
	(1.548)	(1.425)	(1.357)	(1.207)	(0.782)	(0.71)	(0.618)	(0.891)”
Panel (c): Model 3 with $L^1 norm$ and period								
Constant	-1.592***	-1.472***	-1.377***	-1.245***	-0.921***	-0.719***	-0.545***	-0.534***
	(0.132)	(0.123)	(0.121)	(0.115)	(0.096)	(0.1)	(0.1)	(0.104)
Spread	0.017	-0.076	-0.17**	-0.277***	-0.445***	-0.656***	-0.597***	-0.394***
	(0.053)	(0.054)	(0.053)	(0.054)	(0.067)	(0.08)	(0.064)	(0.053)
L^1 norm	5.367***	5.496***	5.728***	5.563***	3.225***	2.164**	-0.194	-1.741*
	(1.504)	(1.38)	(1.296)	(1.137)	(0.737)	(0.688)	(0.615)	(0.853)
Period	0.008	0.009	0.011*	0.011*	0.011	0.01	0.004	-0.032*
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.015)
Panel (d): Model comparison (residual deviance)								
Model 1	529.85	522.348	510.338	491.004	444.676	386.211	415	478.225
Model 2	445.157	440.44	429.013	419.393	416.556	373.619	414.911	471.287
Model 3	443.214	438.248	425.738	415.766	413.105	371.61	414.609	465.498
2 vs 1	-84.693***	-81.909***	-81.326***	-71.61***	-28.12***	-12.592***	-0.088	-6.938**
3 vs 1	-86.636***	-84.1***	-84.6***	-75.237***	-31.572***	-2.008	-0.302	-5.790*
3 vs 2	-1.943	-2.191	-3.275	-3.627	-3.452	-2.008	-0.302	-5.79*”

Notes: TDA calculated using 50 jumps of size 1 with a sliding window embedding of 100 observations and 10 landmark points. Spread is constructed on a bond equivalent basis following Estrella and Trubin (2006) and Demeaned by subtracting the one year moving average. Significance denoted by * - 5%, ** - 1% and *** - 0.1%.

Figure 10: Predicted Recession Probability: Model with Demeaned Spread and Monthly Maximum Topology



Panel (a): 0 Months

Panel (b): 12 Months ahead

Panel (c): 18 Months ahead

Panel (d): 24 Months ahead

Notes: TDA calculated using 50 jumps of size 1 with a sliding window embedding of 100 observations and 10 landmark points. Spread is constructed on a bond equivalent basis following Estrella and Trubin (2006) and Demeaned by subtracting the one year moving average. Probability reports the predicted probability of recession in given month. Predicted probabilities are moved forward to relate to the period $t + p$ being predicted and do not correspond to the time of the data used to form the prediction. Black lines denote Model 1, red lines Model 2, and blue lines plot the predicted probability from Model 3.

Table 6: Recession Forecast Models: Demeaned Spread Topology (Average)

	Months ahead to forecast (p)							
	0	1	2	3	6	12	18	24
Panel (a): Model 1 spread only								
Constant	-1.009***	-0.914***	-0.833***	-0.748***	-0.618***	-0.513***	-0.553***	-0.699***
	(0.089)	(0.086)	(0.083)	(0.08)	(0.078)	(0.086)	(0.082)	(0.080)
Period	-0.065	-0.143**	-0.216***	-0.304***	-0.471***	-0.679***	-0.591***	-0.366***
	(0.049)	(0.050)	(0.052)	(0.054)	(0.063)	(0.078)	(0.058)	(0.046)
Panel (b): Model 2 with L^1 norms								
Constant	-1.542***	-1.425***	-1.342***	-1.212***	-0.87***	-0.683***	-0.53***	-0.564***
	(0.138)	(0.132)	(0.129)	(0.122)	(0.094)	(0.095)	(0.096)	(0.101)
Spread	0.018	-0.073	-0.169**	-0.269***	-0.434***	-0.644***	-0.597***	-0.395***
	(0.053)	(0.054)	(0.052)	(0.053)	(0.065)	(0.078)	(0.063)	(0.052)
L^1 norm	6.566**	6.76***	7.481***	7.161***	3.748***	2.525**	-0.339	-2.287*
	(2.023)	(1.949)	(1.657)	(1.608)	(1.086)	(0.919)	(0.662)	(1.078)
Panel (c): Model 3 with L^1 norms and period								
Constant	-1.543***	-1.426***	-1.343***	-1.213***	-0.873***	-0.689***	-0.534***	-0.542***
	(0.137)	(0.131)	(0.129)	(0.121)	(0.094)	(0.096)	(0.096)	(0.100)
Spread	0.016	-0.075	-0.17**	-0.271***	-0.438***	-0.655***	-0.6***	-0.395***
	(0.052)	(0.054)	(0.052)	(0.053)	(0.065)	(0.08)	(0.064)	(0.052)
L^1 norm	6.514**	6.703***	7.428***	7.082***	3.641***	2.378**	-0.376	-2.076*
	(2.042)	(1.966)	(1.674)	(1.605)	(1.041)	(0.859)	(0.654)	(1.004)
Period	0.011	0.011	0.008	0.012	0.021	0.036	0.017	-0.391*
	(0.019)	(0.019)	(0.018)	(0.019)	(0.02)	(0.021)	(0.02)	(0.194)
Panel (d): Model comparison (residual deviance)								
Model 1	529.85	522.348	510.338	491.004	444.676	386.211	415	478.225
Model 2	453.235	447.794	433.458	423.488	421.664	375.564	414.759	471.155
Model 3	453.023	447.6	433.345	423.251	420.94	373.662	414.42	464.065
2 vs 1	-76.615***	-74.555***	-76.88***	-67.516***	-23.012***	-10.648**	-0.24	-7.07**
3 vs 1	-76.827***	-74.748***	-76.993***	-67.752***	-23.736***	-1.901	-0.34	-7.09**
3 vs 2	-0.212	-0.193	-0.113	-0.237	-0.724	-1.901	-0.34	-7.09**

Notes: TDA calculated using 50 jumps of size 1 with a sliding window embedding of 100 observations and 10 landmark points. Spread is constructed on a bond equivalent basis following Estrella and Trubin (2006) and Demeaned by subtracting the one year moving average. Significance denoted by * - 5%, ** - 1% and *** - 0.1%.

considering both maximum and average for the spread topology. Results presented in Table 6 are again very similar to the other cases. As with the spread switching to average values yields much larger coefficients on the periodicity but very similar coefficients on the L^1 norm. In comparison to the maximum monthly TDA case period is no longer significant in the 2 month and 3 month ahead models. There are also reductions in the residual deviance gain at the shorter forecast horizons sitting alongside increased margins at the 24 month ahead horizon.

Figure 11 shows strong consistency with the messages from Figures 8 to 10 in that the best fits come for the 1980s recession, the TDA augmented model does classify recessions at the 0-month ahead range and the most accurate fits come from the 12-month ahead forecasts of panel (b). In Figure 11 this is at its most pronounced, but the tendancy to predict recessions will ocurr later than they did using longer horizons is evident in all. In the consistency of the message across the four cases discussed herein there is a robustness of conclusion offered.

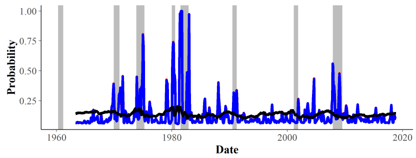
6.5 Summary

These four cases have shared many of the same conclusions. TDA works well in the same period, with a particularly good fit to the 1980s recession. Moving to further ahead time steps there is always a slight improvement of fit but this is usually insignificant once we move past the 6-month ahead forecasting. In the widely reported year ahead the L^1 norm does become significant and this is the first evidence that TDA can offer an improvement to the standard, spread only, model. Residual deviance improvement return in Model 3 at the 24-month ahead range. At times TDA can be an aid in predicting recessions, but at others it is merely a marginal improvement. As presented the results represent a first step in understanding.

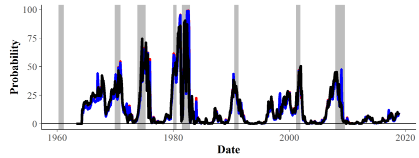
7 Discussion

Uncovering the topology of the time series has potential to help forecast recessions. This opening was highlighted in the financial literature on crash prediction. For that literature last minute predictions

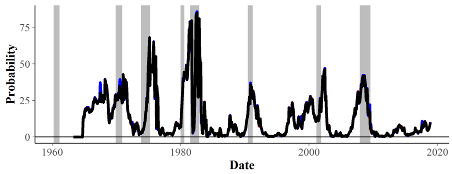
Figure 11: Predicted Recession Probability: Model 4 with Demeaned Spread and Average Topology



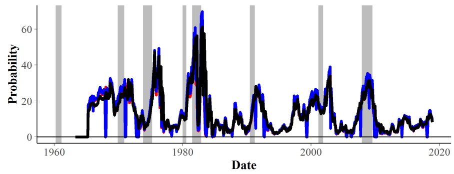
Panel (a): 0 Months



Panel (b): 12 Months ahead



Panel (c): 18 Months ahead



Panel (d): 24 Months ahead

Notes: TDA calculated using 50 jumps of size 1 with a sliding window embedding of 100 observations and 10 landmark points. Spread is constructed on a bond equivalent basis following Estrella and Trubin (2006) and Demeaned by subtracting the one year moving average. Probability reports the predicted probability of recession in given month. Predicted probabilities are moved forward to relate to the period $t + p$ being predicted and do not correspond to the time of the data used to form the prediction. Black lines denote Model 1, red lines Model 2, and blue lines plot the predicted probability from Model 3.

have more value; investors wish to take advantage of the growing market and only switch to short positions when they are certain that they will not make a loss so doing. Work by Gidea and Katz (2018) and Gidea et al. (2018) has shown that TDA is associated with the prediction of crashes in financial time series. In this there is motivation for considering TDA in the yield spread context.

Periodicity as currently understood is associated with the exit from recessions, it is therefore not a signal of impending recession. However, our novel approach to capturing periodic behaviour in noisy time series reveals that there are also episodes of cyclicity in the run up to some recessions. In encountering results like this we caution making too strong inferences; we do not know what is creating the periodicity and whether it is something that was uniquely present in those particular recession build ups. Our results confirmed that periodicity is observed on the exit and if longer periods are observed the economy is expected to be in a strong state; Clark (2004) suggesting the length of periods declines through recovery.

In recession prediction we seek to forecast much further ahead than the asset pricing literature to which Gidea and Katz (2018) speaks. Here we want to know whether there will be a recession a year from now, investor behaviour has not yet changed because there will still be growth up that recession point. This is then an exploratory study to see whether there is behaviour within the yield spread that would predict recessions at the longer time scale. We found that at many time periods the L^1 norms had significant coefficients, and the periodicity also had some significance. Residual deviance measures showed that the persistence landscapes do contain information about forthcoming recessions at a longer ahead time period.

For extending the understanding of yield curves over time, and appreciating the way that translates into recessions, TDA has a role. TDA helps practitioners and interest rate setters alike see what is happening in the market. There is little to choose between using either the spread or the version with the 12-month rolling average subtracted. The former offers simplicity whilst the latter better represents the log-periodicity literature. Using the yield spread is not without question, we used the bond-equivalence on the three month rate because of its high forecasting behaviour (Estrella and Trubin, 2006), but there are others who do not make such adjustments, or who consider the one year spread instead. For the demeaned version there are also potential questions on the using one year, further work in this direction would cast light on the importance of that decision.

Additional insights into the economy can be made from the TDA of the yield spread time series however. That Model 3 is putting forward signals of recession in current months is another voice in the debate about whether the US will enter a recession soon. It is early days and because of the time taken by the NBER to make decisions on classifying the time as a recession it may be some time before we will know if these signs are false positives. It should be noted the yield spread only model has not been so quick to suggest recession soon.

This paper is very much a first draft on the topic. It highlights potential but leaves much to be explored. The promise is good and the precedent is set.

8 Conclusions

Demonstrating a novel way of recovering the properties of time series with data topology we have shown how robust periodicity estimates can help understand the probability of recession. At almost all time frames we evidence significant reductions in residual deviance, with equally highly significant coefficients on the TDA measures in our probit regressions. Using a sliding window embedding on the bond-equivalent yield spread we showed that there are clear episodes of "chaotic" behaviour of the type seen in asset pricing markets. Such in turn makes classification of periods as recessions much easier. We evidence mild improvements to the predictions, but there is still room for much greater improvement to match the levels seen from TDA in other financial time series.

A number of questions emerge, including the parameters of the persistent homology, the choice of series upon which to perform the TDA and the potential for robustness across countries. There is also scope for future research to understand what precisely the topology is picking up. However, perhaps the most pertinent question to arise from the work presented is whether there is evidence in the topology that would inform on the yield curve inverting. If there is something within that, and given the rule of thumb link between inversion and recession, there may be a further productive line for recession prediction there.

Notwithstanding the questions raised for future research, and the potential for further robustness work, the results of this paper do point to the behaviour of the spread, as well as the spread itself, as being important in recession forecasting. We evidenced the existence of periodicity in the months following recessions, but have taken the understanding through the full US yield spread time series. Marginal improvements have been made with scope for more.

References

An, Z., Jalles, J. T., and Loungani, P. (2018). How well do economists forecast recessions? *International Finance*, 21(2):100–121.

Ang, A., Piazzesi, M., and Wei, M. (2006). What does the yield curve tell us about gdp growth? *Journal of econometrics*, 131(1-2):359–403.

Benhabib, J., Schmitt-Grohé, S., and Uribe, M. (2002). Chaotic interest-rate rules. *American Economic Review*, 92(2):72–78.

Benzoni, L., Chyruk, O., and Kelley, D. (2018). Why does the yield-curve slope predict recessions?

Berge, T. J. (2015). Predicting recessions with leading indicators: Model averaging and selection over the business cycle. *Journal of Forecasting*, 34(6):455–471.

Chang, G. and Feigenbaum, J. (2006). A bayesian analysis of log-periodic precursors to financial crashes. *Quantitative Finance*, 6(1):15–36.

Chang, G. and Feigenbaum, J. (2008). Detecting log-periodicity in a regime-switching model of stock returns. *Quantitative Finance*, 8(7):723–738.

Chauvet, M. and Potter, S. (2005). Forecasting recessions using the yield curve. *Journal of Forecasting*, 24(2):77–103.

Chen, N.-F. (1991). Financial investment opportunities and the macroeconomy. *The Journal of Finance*, 46(2):529–554.

Clark, A. (2004). Evidence of log-periodicity in corporate bond spreads. *Physica A: Statistical Mechanics and its Applications*, 338(3-4):585–595.

Davig, T. and Hall, A. S. (2019). Recession forecasting using bayesian classification. *International Journal of Forecasting*, 35(3):848–867.

Dłotko, P., Qiu, W., and Rudkin, S. (2019). Cyclicality, periodicity and the topology of time series. *arXiv preprint arXiv:1905.12118v1*.

Duffee, G. R. (2018). Expected inflation and other determinants of treasury yields. *The Journal of Finance*, 73(5):2139–2180.

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

Estrella, A. and Mishkin, F. S. (1996). The yield curve as a predictor of us recessions. *Current Issues in Economics and Finance*, 2(7).

Estrella, A. and Mishkin, F. S. (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(7):1375–1401.

Estrella, A. and Mishkin, F. S. (1998). Predicting US recessions: Financial variables as leading indicators. *Review of Economics and Statistics*, 80(1):45–61.

Estrella, A. and Trubin, M. (2006). The yield curve as a leading indicator: Some practical issues.

Gidea, M., Goldsmith, D., Katz, Y., Roldan, P., and Shmalo, Y. (2018). Topological recognition of critical transitions in time series of cryptocurrencies. *arXiv preprint arXiv:1809.00695*.

Gidea, M. and Katz, Y. (2018). Topological data analysis of financial time series: Landscapes of crashes. *Physica A: Statistical Mechanics and its Applications*, 491:820–834.

Giusto, A. and Piger, J. (2017). Identifying business cycle turning points in real time with vector quantization. *International Journal of Forecasting*, 33(1):174–184.

Harvey, C. R. (1988). The real term structure and consumption growth. *Journal of Financial Economics*, 22(2):305–333.

Ko, B., Song, J. W., and Chang, W. (2018). Crash forecasting in the korean stock market based on the log-periodic structure and pattern recognition. *Physica A: Statistical Mechanics and its Applications*, 492:308–323.

Kozlowski, S. E. and Sim, T. (2019). Predicting recessions using trends in the yield spread. *Journal of Applied Statistics*, 46(7):1323–1335.

Laurent, R. D. et al. (1989). Testing the" spread". *Economic Perspectives*, (Jul):22–34.

Li, C. (2017). Log-periodic view on critical dates of the chinese stock market bubbles. *Physica A: Statistical Mechanics and its Applications*, 465:305–311.

Matsushita, R., Da Silva, S., Figueiredo, A., and Gleria, I. (2006). Log-periodic crashes revisited. *Physica A: Statistical Mechanics and its Applications*, 364:331–335.

Mishkin, F. S. (1990). What does the term structure tell us about future inflation? *Journal of monetary economics*, 25(1):77–95.

Mishkin, F. S. et al. (1990). The information in the longer maturity term structure about future inflation. *The Quarterly Journal of Economics*, 105(3):815–828.

Perea, J. (2019). Topological time series analysis. *NOTICES OF THE AMERICAN MATHEMATICAL SOCIETY*, 66(5).

Sornette, D., Johansen, A., et al. (2001). Significance of log-periodic precursors to financial crashes. *Quantitative Finance*, 1(4):452–471.

Stock, J. H. and Watson, M. W. (2014). Estimating turning points using large data sets. *Journal of Econometrics*, 178:368–381.

Wosnitza, J. H. and Leker, J. (2014). Can log-periodic power law structures arise from random fluctuations? *Physica A: Statistical Mechanics and its Applications*, 401:228–250.

Wosnitza, J. H. and Sornette, D. (2015). Analysis of log-periodic power law singularity patterns in time series related to credit risk. *The European Physical Journal B*, 88(4):97.

Zhou, W.-X. and Sornette, D. (2004). Causal slaving of the us treasury bond yield antibubble by the stock market antibubble of august 2000. *Physica A: Statistical Mechanics and its Applications*, 337(3-4):586–608.