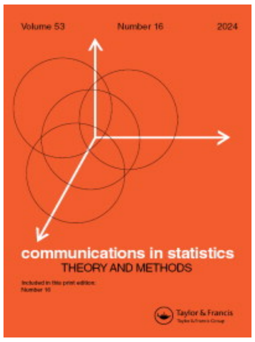

Communications in Statistics - Theory and Methods

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/lsta20

How many times until a coincidence becomes a pattern? The case of yield curve inversions preceding recessions and the magical number 7

Ned Kock & Augustine Tarkom

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
2024, VOL. 53, NO. 16, 5785–5792
https://doi.org/10.1080/03610926.2023.2232908

How many times until a coincidence becomes a pattern? The case of yield curve inversions preceding recessions and the magical number 7

Ned Kock[a] and Augustine Tarkom[b]

[a]Division of International Business and Technology Studies, Texas A&M International University, Laredo, Texas, USA; [b]Division of International Banking and Finance Studies, Texas A&M International University, Laredo, Texas, USA

ABSTRACT
Let us say that a coincidence involving two events, where one seems to predict the other, happens a number of times. How many times until it can be considered not only a coincidence, but a statistically significant pattern? We propose a framework to answer this question. Using the framework, we find that the number of times required is 7. We illustrate the practical application of our framework in the context of a very important phenomenon: When the percentage difference between 10-year and 3-month U.S. Treasury yields falls below zero, a U.S. recession appears to occur within the next 18 months.

ARTICLE HISTORY
31 January 2023
29 June 2023

KEYWORDS
Yield curve inversion; U.S. recession; magical number seven; confidence level; significance level; statistical power

JEL CLASSIFICATION:
C4; E4; G12

1. Introduction

Often what appear to be rare coincidences involving two events keep on happening, where the events seem to always happen one after the other, giving rise to the speculation that they are not mere coincidences, but manifestations of a pattern with a fundamental underlying reason. In these cases, and important question arises: What is the minimum number of times needed, for what appears to be a predictive coincidence involving two events to occur, before it can be considered not only a coincidence but a statistically significant pattern?

We propose a framework in this article to answer the above question. Using the framework, we find that the number of times required is 7. We illustrate the practical application of our framework in the context of a very important phenomenon, which we call the inversion-recession phenomenon, which is characterized by the following: When the percentage difference between 10-year and 3-month U.S. Treasury yields falls below zero, or becomes negative, a U.S. recession appears to occur within the next 18 months.

Our discussion and key conclusion apply to many predictive phenomena, not only the inversion-recession phenomenon, even though this is a particularly important phenomenon for governments, organizations, and individuals. U.S. recessions affect many domestic and global affairs. It is interesting to note that our discussion and key conclusion also appear to be yet another indication of the pervasiveoccurrence of the “magical number 7, plus or

CONTACT Augustine Tarkom augustinetarkom@dusty.tamiu.edu Division of International Banking and Finance Studies, Texas A&M International University, Laredo, TX, USA.

minus 2", often referred to as Miller's Law (Miller 1956; Richardson and Reischman 2011). It should be noted that this is rather a curiosity than a clear endorsement of Miller's Law, because the inversion-recession phenomenon is unrelated to the working memory limitation phenomenon discussed by Miller (1956). Moreover, valid criticisms of Miller's Law have been proposed in the past (see, e.g., Doumont 2002).

2. Yield curve inversions and recessions

Buyers of U.S. Treasury bonds of different maturities receive a yield from the U.S. Government, and at maturity they receive back their initial investment. Under favorable economic conditions, there is more interest from buyers in short maturity than long maturity U.S. Treasury bonds. Therefore, yields tend to be higher as maturity increases. If a recession is expected in the near future, this situation is reversed. That is, there is more interest in long maturity than short maturity U.S. Treasury bonds, leading to higher yields as maturity decreases. This phenomenon is generally referred to as a yield curve inversion (Estrella 2005; Evgenidis et al. 2020; Gogas et al. 2015).

A particular type of inversion is that involving 10-year and 3-month U.S. Treasury yields. This inversion has been shown to be a particularly good predictor of U.S. recessions (Chauvet and Potter 2005; Harvey 1989). When the percentage difference between 10-year and 3-month U.S. Treasury yields falls below zero, or becomes negative, a recession appears to occur within the next 18 months. When this inversion happens, buyers of 10-year Treasury bonds are essentially willing to receive a lower yield than buyers of 3-month U.S. Treasury bonds during a period of economic turmoil, possibly because that ensures that they will be safely receiving payments throughout a recessionary period.

Figure 1 shows the percentage difference between 10-year and 3-month U.S. Treasury yields for the period going from January 1985 to May 2022. Instances where the yield curve inverts are indicated by that percentage difference falling below zero or becoming negative. U.S. recessions are indicated by shaded areas. Four first inversions and four recessions are shown. The recessions occur with a period of time after the first inversions. While typically the first inversions are followed by one or more additional inversions, it is the first inversions that matter most.

As we can see, in the 37-year period depicted, only four times we see instances in which the difference between 10-year and 3-month U.S. Treasury yields falls below zero and is followed by a U.S. recession. So, this inversion-recession phenomenon is rather rare. Because the phenomenon is so rare, one could argue that it has not happened enough times to be deemed as statistically significant, often leading to the speculation that "this time is different" when inversions occur (Bauer and Mertens 2018, p. 4). In other words, the inversion-recession phenomenon may not have happened enough times to be upgraded from coincidence to statistically significant pattern.

Even though the inversion associated with 10-year and 3-month U.S. Treasury yields is the one that led to the coining of the "yield curve inversion" term, originally by Harvey (1989), other types of inversion that present similar patterns are often discussed in financial news media outlets (see: Liberto 2023). Notable in this respect is the inversion associated with 10-year and 2-year U.S. Treasury yields, which tends to be seen as a "yellow warning light" and as an indication that one should watch for the possible upcoming inversion associated with 10-year and 3-month U.S. Treasury yields. The reason for this is that

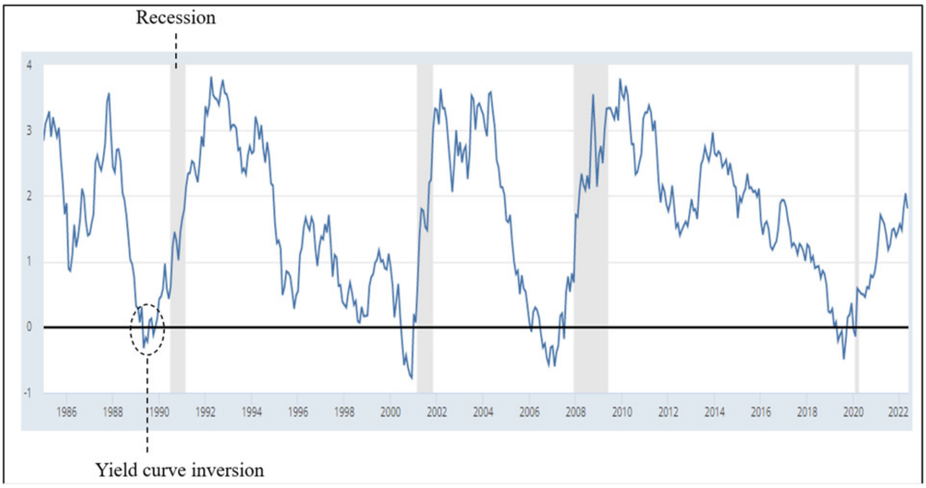

Figure 1. Yield curve inversions and recessions.
Notes. Source: Federal Reserve Bank of St. Louis; X axis: time; Y axis: percentage difference between 10-year and 3-month Treasury yields; shaded areas: recessions.

2-year U.S. Treasury yields tend to predict the future values of 3-month U.S. Treasury yields.

3. How many times until a coincidence becomes a pattern?

Let us say that we have two events E_1 and E_2, were E_1 seem to always temporally precede E_2, and where there is a plausible underlying reason why E_1 would be followed by E_2. For the inversion-recession phenomenon, E_1 is the difference between 10-year and 3-month U.S. Treasury yields falling below zero, and E_2 is a U.S. recession occurring within 18 months of E_1 first taking placing. A possible underlying reason why E_1 would be followed by E_2 is that E_1 signals an expectation by U.S. Treasury buyers that E_2 would happen.

Let us also assume that we have access to data on samples of E_1 and E_2 occurrences, where the samples are taken from a large set of all such occurrences (including future occurrences). For each sample, we can calculate a correlation, reflecting the extent to which E_1 and E_2 happen in sequence. If they always happen in sequence, then the correlation will be 1. If not, the correlation will fall between 1 and 0. If we randomly draw many samples of a given size, the correlations will be distributed around a mean correlation (r), with a standard deviation (S). This standard deviation is generally known as the standard error associated with r. The distribution of the ratio r/S will have the general shape indicated in Figure 2 (Kock 2015; Kock 2016; Weakliem 2016), where the location of a critical T ratio is indicated.

The critical T ratio depends on a specific significance level chosen, which is often .05 or 5% in standard statistical tests, with this ratio being indicated as $T_{.05}$. This significance level means that, in an assessment of whether a coincidence is in fact a statistically significant pattern, one does not want the probability of a false positive (or type I error) occurring to be more than 5%. This significance level is the complement of the confidence level chosen, which is often .95 or 95% in standard statistical tests. For each instance, where the ratio r/S is found to be

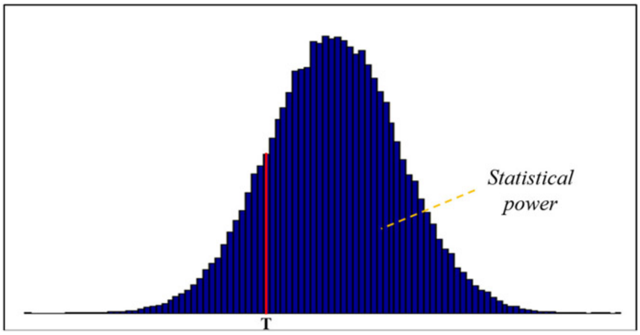

Figure 2. Distribution of the ratio. **r/S**
Notes. T: critical T ratio for a specific significance level chosen; area under the curve to the right: statistical power.

greater than the critical T ratio, the correlation r will be deemed to be statistically significant at the .05 level. That is, the correlation r will be suggestive of a statistically significant pattern, and will allow us to rule out a mere coincidence.

For each sample size, the magnitude of the ratio r/S will increase with an increase in the correlation r, and with a decrease in the standard error S. The standard error S decreases with increases in sample size. The statistical power (Scherbaum 2009) of this correlation test is the probability that the ratio r/S will be greater than the critical T ratio. This is expressed in (1), where the significance level chosen is noted as .05.

$$W = P\left(\frac{r}{S} > \mathrm{T}_{.05}\right). \tag{1}$$

As noted above, to uncover a pattern that is statistically significant we set an acceptable cap for the probability that a false positive (or type I error) will be committed. We do the same for the probability that a false negative (or type II error) will be committed. Typically, this is set at .20 or 20% in standard statistical tests. The complement of this probability, namely .80 or 80%, is the statistical power of our test that the correlation r will be suggestive of a statistically significant pattern, allowing us to rule out a mere coincidence.

The above can be expressed employing a cumulative probability function $\Phi(\bullet)$ for the standard normal distribution. Given our earlier assumption of random sampling, and the fact that coefficients calculated based on sample sets taken randomly from a population tend to be distributed in conformity with the central limit theorem (Kipnis and Varadhan 1986; Miller and Wichern 1977), we can also assume that the values for the ratio r/S are normally distributed. Therefore, we can say that statistical power of our test will be greater than .8 when the cumulative distribution function for the standard normal distribution indicated in (2) is greater than .8.

$$\Phi\left(\frac{r}{S} - \mathrm{T}_{.05}\right) > .8. \tag{2}$$

Since the values for the ratio r/S are normally distributed, the standardized score associated with the value .8 of the cumulative distribution function can be denoted as $z_{.8}$. For the same reason, the value of the critical T ratio at the .05 level ($\mathrm{T}_{.05}$)equals the standardized score $z_{.95}$

associated with the corresponding confidence level of .95. Combining these two properties takes us to the inequality expressed through (3).

$$\begin{aligned} &\frac{r}{S} - \mathrm{T}_{.05} > z_{.8} \rightarrow \\ &\frac{r}{S} > \mathrm{T}_{.05} + z_{.8} \rightarrow \\ &\frac{r}{S} > z_{.95} + z_{.8}. \end{aligned} \tag{3}$$

The true standard error (S) can be estimated through (4). The resulting estimate $\hat{S}$ is slightly biased (Gurland and Tripathi 1971; Kock and Hadaya 2018), with the bias leading to a conservative final estimate of the sample size required for the inversion-recession phenomenon that we are considering to be deemed a statistically significant pattern (and not a mere coincidence). We provide a further discussion of this later in this article.

$$\hat{S} = \frac{1}{\sqrt{N}}. \tag{4}$$

We can use the Excel function NORMSINV(x) to obtain the values for $z_{.95}$ and $z_{.8}$. The values for NORMSINV(.95) and NORMSINV(.8) are, respectively, 1.645 and 0.842. Thus, the sum $z_{.95} + z_{.8}$ is 2.486. With this, and combing Equations (3) and (4), we obtain the inequality expressed by (5).

$$\begin{aligned} &r\sqrt{\hat{N}} > z_{.95}+z_{.8} \rightarrow \\ &\hat{N} > \left(\frac{z_{.95}+z_{.8}}{r}\right)^2 \rightarrow \\ &\hat{N} > \left(\frac{2.486}{r}\right)^2. \end{aligned} \tag{5}$$

If the inversion-recession phenomenon events E_1 and E_2 always happen in sequence, then the correlation r will be 1. This leads us to the minimum number of times required for us to observe the inversion-recession phenomenon so that we can conclude that the phenomenon is not a coincidence and that it is in fact a statistically significant pattern. As we can see in (6), this number is 7. This is the smallest integer that satisfies the inequality.

$$\begin{aligned} &\hat{N} > \left(\frac{2.486}{1}\right)^2 \rightarrow \\ &\hat{N} > 6.180 \rightarrow \\ &\hat{N} = 7. \end{aligned} \tag{6}$$

Even though the discussion above uses the inversion-recession phenomenon for illustrative purposes, it is arguably generic enough to apply to any situation where we have two events E_1 and E_2, where E_1 seems to always temporally precede E_2, and where there a plausible underlying reason why E_1 would be followed by E_2. Note that we do not assume causality between the two events, only temporal precedence. In our example, E_1 (the difference between 10-year and 3-month U.S. Treasury yields falling below zero) does not cause but rather signals an expectation among U.S. Treasury buyers that E_2 (a U.S. recession occurring within 18 months) will happen.

Table 1. Yield curve inversions and recessions since 1970.

Recession	First inversion	Start of recession	Time between inversion-recession
1970 recession	December 1968	January 1970	13
1974 recession	June 1973	December 1973	6
1980 recession	November 1978	February 1980	15
1981–1982 recession	October 1980	August 1981	10
1990 recession	June 1989	August 1990	14
2001 recession	July 2000	April 2001	9
2008–2009 recession	August 2006	January 2008	17
COVID-19 recession	May 2019	March 2020	10

Note: time between inversion and recession provided in months.

4. Yield curve inversions and recessions: coincidence or pattern?

Even though the widely used database from the Federal Reserve Bank of St. Louis provides the percentage difference between 10-year and 3-month Treasury yields only since the 1980s, datasets that allow for this percentage difference calculation date back to 1970. These include datasets maintained by the National Bureau of Economic Research, which is the official source of data on U.S. recessions. Table 1 shows that all U.S. recessions since 1970 have been preceded by a 10-year to 3-month U.S. Treasury yield inversion within the past 18 months.

In the previous section, we have established the minimum number of times required for the inversion-recession phenomenon to be deemed more than a coincidence, and rather a statistically significant pattern. That number is 7. Therefore, given that since 1970 we have observed 8 instances of the inversion-recession phenomenon, we can conclude that this not a coincidence, and that it is in fact a statistically significant pattern.

5. Discussion and conclusion

We have proposed a generic framework for the estimation of the minimum number of times needed, for what appears to be a predictive coincidence involving two events to occur, before it can be considered not only a coincidence but a statistically significant pattern. This framework can be summarized through the inequality expressed in (6), where: $z_{.95}$ and $z_{.8}$ are the standardized scores, respectively, associated with the 95% confidence level and 80% statistical power level (levels that are widely used in statistical tests); and r is the correlation between variables measuring the joint occurrence of the two events.

$$\hat{N} > \left(\frac{z_{.95} + z_{.8}}{r}\right)^2. \tag{7}$$

We illustrated the practical application of our framework through the estimation of the minimum number of times required for us to observe a very important phenomenon, the inversion-recession phenomenon, so that we can conclude that it is not a mere coincidence but a statistically significant pattern. That minimum number of times is 7. The inversion-recession phenomenon has been characterized in this article by the following: When the percentage difference between 10-year and 3-month U.S. Treasury yields falls below zero, or becomes negative, a recession appears to occur within the next 18 months. In fact, given that since 1970 we have observed 8 instances of the inversion-recession phenomenon, we can say that the inversion-recession phenomenon is a statistically significant pattern with a 97% confidence level. This is because if $r = 1$ and $\hat{N} = 8$ in our inequality; then $z_{.97}$ will satisfy it.

As we noted earlier, we used a slightly biased estimate for the standard error $\hat{S}$. If we had used an unbiased estimate of $\hat{S}$ (Gurland and Tripathi 1971; Kock and Hadaya 2018), the minimum number of times required for us to observe the inversion-recession phenomenon number would be 6. We prefer the more conservative approach, whose estimate is 7, for three main reasons. The first reason is that the gamma-exponential equations necessary to correct bias (Gurland and Tripathi 1971; Kock and Hadaya 2018) do not have analytic solutions, only numeric solutions, which could create a new possible source of bias. The second reason is that the differences in the estimates produced based on the conservative and unbiased approaches are very small. For examples in difference contexts, see: Kock and Hadaya (2018).

The third reason why we prefer the more conservative approach is probably the most important of the three. It is related to the fact that our confidence level is 95%, and not 100%. Given this, as more instances of the inversion-recession phenomenon are assessed in the future, we would still expect 1 in 20 to possibly be a false positive in our significance assessment. If this were to occur, each mismatch would be associated with a decrease in the correlation among the inversion and recession events, thus bringing the minimum number of times required for statistical significance to a value that is slightly higher than it would be if the correlation was 1.

Our estimation framework applies to many predictive phenomena, not only the inversion-recession phenomenon, although not many predictive phenomena rise to the importance of the one chosen by us for illustration purposes. Being able to predict U.S. recessions is arguably of high importance and value in domestic and global affairs; for governments, organizations, and even individuals (e.g., retail investors). Interestingly, our discussion and key conclusion could also be seen as yet another indication of the pervasive occurrence of the "magical number 7, plus or minus 2" (Miller 1956; Grissom 1996; Richardson and Reischman 2011), also known as Miller's Law. On the other hand, our discussion could be viewed as a curiosity associated with the number 7, as opposed to a clear endorsement of Miller's Law. The reason for this is that the inversion-recession phenomenon is unrelated to the working memory limitation phenomenon discussed by Miller (1956). Additionally, Miller's Law has been the target of compelling criticisms in the past (see, e.g., Doumont 2002).

Finally, we thank an expert reviewer for bringing to our attention a connection between the number 7 and the Fibonacci sequence (Falcon and Plaza 2007; Horadam 1961), in terms of patterns and repetitions. If one takes the Fibonacci numbers modulo 7, one can also observe a cyclical pattern, which is that the results repeat themselves in groups of 16 numbers in sequence. This essentially means that if one takes any Fibonacci number and divide it by 7, the remainder will be part of a repeating pattern of 16 numbers, which is known as a Pisano period (Falcon and Plaza 2009).

References

Bauer, M. D, and T. M. Mertens. 2018. Information in the yield curve about future recessions. *FRBSF Economic Letter* 20:1–5. S3.E11.

Chauvet, M, and S. Potter. 2005. Forecasting recessions using the yield curve. *Journal of Forecasting* 24 (2):77–103. doi:10.1002/for.932.

Doumont, J.-L. 2002. Magical numbers: the seven-plus-or-minus-two myth. *IEEE Transactions on Professional Communication* 45 (2):123–7. doi:10.1109/TPC.2002.1003695.

Estrella, A. 2005. The yield curve and recessions. *The International Economy* 19 (S3):8–9.

Evgenidis, A., S. Papadamou, and C. Siriopoulos. 2020. The yield spread's ability to forecast economic activity: What have we learned after 30 years of studies?. *Journal of Business Research* 106:221–32. doi:10.1016/j.jbusres.2018.08.041.
Falcón, S, and Á. Plaza. 2007. The k-Fibonacci sequence and the Pascal 2-triangle. *Chaos, Solitons & Fractals* 33 (1):38–49. doi:10.1016/j.chaos.2006.10.022.
Falcon, S, and Á. Plaza. 2009. k-Fibonacci sequences modulo m. *Chaos, Solitons & Fractals* 41 (1):497–504. doi:10.1016/j.chaos.2008.02.014.
Gogas, P., T. Papadimitriou, M. Matthaiou, and E. Chrysanthidou. 2015. Yield Curve and Recession Forecasting in a Machine Learning Framework. *Computational Economics* 45 (4):635–45. doi:10.1007/s10614-014-9432-0.
Grissom, R. J. 1996. The magical number .7 + / − .2: meta-meta-analysis of the probability of superior outcome in comparisons involving therapy, placebo, and control. *Journal of Consulting and Clinical Psychology* 64 (5):973–82. doi:10.1037//0022-006x.64.5.973. 8916626
Gurland, J, and R. C. Tripathi. 1971. A Simple Approximation for Unbiased Estimation of the Standard Deviation. *The American Statistician* 25 (4):30–2. doi:10.1080/00031305.1971.10477279.
Harvey, C. R. 1989. Forecasts of Economic Growth from the Bond and Stock Markets. *Financial Analysts Journal* 45 (5):38–45. doi:10.2469/faj.v45.n5.38.
Horadam, A. F. 1961. A Generalized Fibonacci Sequence. *The American Mathematical Monthly* 68 (5):455 doi:10.2307/2311099.
Kipnis, C, and S. R. S. Varadhan. 1986. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Communications in Mathematical Physics* 104 (1):1–19. doi:10.1007/BF01210789.
Kock, N. 2015. One-Tailed or Two-Tailed P Values in PLS-SEM?. *International Journal of e-Collaboration* 11 (2):1–7. doi:10.4018/ijec.2015040101.
Kock, N. 2016. Non-Normality Propagation among Latent Variables and Indicators in PLS-SEM Simulations. *Journal of Modern Applied Statistical Methods* 15 (1):299–315. doi:10.22237/jmasm/1462076100.
Kock, N, and P. Hadaya. 2018. Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal* 28 (1):227–61. doi:10.1111/isj.12131.
Liberto, D. 2023. *Inverted yield curve: Definition, what it can tell investors, and examples*, https://www.investopedia.com/terms/i/invertedyieldcurve.asp.
Miller, G. A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63 (2):81–97. doi:10.1037/h0043158.
Miller, R. B. and D. W. Wichern, 1977. *Intermediate business statistics: Analysis of variance, regression and time series*. New York, NY: Holt, Rihehart and Winston.
Richardson, M, and D. Reischman. 2011. The Magical Number 7. *Teaching Statistics* 33 (1):17–9. doi:10.1111/j.1467-9639.2009.00406.x.
Scherbaum, C. A, and J. M. Ferreter. 2009. Estimating Statistical Power and Required Sample Sizes for Organizational Research Using Multilevel Modeling. *Organizational Research Methods* 12 (2):347–67. doi:10.1177/1094428107308906.
Weakliem, D. L. 2016. *Hypothesis testing and model selection in the social sciences*. New York, NY: Guilford Publications.