Modelling the Trend: The Historical Origins of Some Modern Methods and Ideas

by
Terence Mills
and Kerry Patterson

2013
100

Department of Economics
University of Reading
Whiteknights
Reading
RG6 6AA
United Kingdom

www.reading.ac.uk
Modelling the Trend: the Historical Origins of Some Modern Methods and Ideas

by

T. C. Mills, Loughborough University and K. D. Patterson, Reading University

“Trend-lines are intended to describe some imaginary state of ‘normalcy’ wherewith to compare the actual conditions depicted by the curve.” (Karl Karsten, 1926, p. 31.)

“The significance (and perhaps the usefulness) of secular trend, considered merely as a problem in curve fitting, is uncertain, not only because long-term undulations which we know to have no real significance appear in random series, but because the trend is largely indeterminate.” (Victor von Széligski, 1929, p. 245.)

Abstract

The development of the methods of correlation and regression analysis at the turn of the 20th century, led to their use in attempting to identify relationships between economic variables. However, caution was soon expressed that correlating series with ‘secular’ trends was likely to be misleading. After some discussion of methods, linear detrending by least squares estimation became the default method. By the 1920s, however, some voices of dissent expressed the view that linear detrending was likely to be inappropriate in some, even many, cases. This led to a number of innovative methodological developments, including rolling window estimation, moving integration, nonlinear trends, structural breaks, sigmoid-type smooth adjustment functions, the beginnings of stochastic trends and the construction of ‘smoothers’ and filters. Although generally failing to have an impact at the time substantially predate their current use in econometrics. This article establishes precedence for these ideas and recreates some of the empirical examples and early simulations.

Keywords: trend, detrending, decomposition, nonlinear trends, structural breaks, stochastic trends, filtering.

JEL: B1, B2, B4, C1

(Word count: 9,234 + 12 figures = 12,234)
I. Introduction

Following the development of correlation and regression analysis in the late 19th century (see, especially, Edgeworth, 1892, 1893, 1894; Galton, 1888, 1900; Pearson, 1896; and Yule, 1897, 1907), there was considerable interest in applying these then novel techniques to uncovering economic relationships and, especially, to using them in the construction of charts or ‘barometers’ of economic activity with a view to forecasting key variables, generally in the short to medium term. The interest in forecasting was both academic and commercial, with many attempts at constructing indices, barometers and ‘dials’ of economic activity: see Knauth (1923) and Friedman (2009). There was also a competitive edge to producing forecasts: for example, Smith (1929) noted that, however the forecast series (or index thereof) was obtained, the moment of comparison of the actual with the forecast was always eagerly awaited:

“After the forecasting index has been constructed in one fashion or another, there comes that breathless moment of suspense in the statistician's life when he hopefully computes the coefficient of correlation between his new-born forecasting index and the series which is the subject of forecast. One thing is pretty sure: the higher the correlation coefficient, the faster the statistician's pulse.” (ibid., p. 94)

In the early stages of calculating correlation coefficients and estimating regressions, there was a concern about separating out the apparently secular movements in time series from fluctuations of higher frequency. There was a commonly held view that regressing one trended series on another might mislead as to the strength of association between the two variables; this, of course, was a concern that resurfaced much later (most notably by Granger and Newbold, 1974), and led to the development of cointegration (Engle and Granger, 1987). A typical practice was to first remove the secular movements in some way, usually based on a deterministic view of the trend, and then correlate the resulting detrended series, often referred to as the cycles. A less frequently used alternative was to difference each series first and then correlate the differences, as in Hooker (1905) and Student (1914).

In essence, the underlying problem to be solved was what Yule (1921) described as the ‘time-correlation problem’:

“the essential difficulty of the time correlation problem is the difficulty of isolating for study different components in the total movement of each variable.” (ibid., p. 501).
This study takes as its starting point the development of the idea of a ‘secular’ trend, the methods for dealing with it, and the associated developments in concept and methodology. The modelling of trends is one of the most important issues in modern time series econometrics (for example see Phillips, 2005, and White and Granger, 2011), but such concerns have been around for a century or more. While some of the early efforts at modelling trends are well known to modern researchers, others are only just being rediscovered: what they typically have is a remarkable prescience in suggesting methods and pointing out phenomena that are at the core of modern econometrics.

The primary period of focus for this study is the decade or so dating from the early 1920s, which was a very productive period for concepts and methodology related to the trend. These developments included some that are now treated as commonplace: rolling window estimation; moving integration; nonlinear trends and sigmoid-type smooth adjustment to multiple structural breaks; quadrature; an early understanding of stochastic trends; and time series filters. All these remarkable developments are, in one form or another, part of the conceptual basis of modern econometrics, but are little referenced, if at all.

This paper is structured as follows. Section 2 outlines the dominant view prevailing after the initial development and modelling of trends at the beginning of the 19th century, which was aimed at removing the ‘secular’ element (generically referred to as the trend); the ‘points winner’ in this development being detrending by a simple polynomial, usually a linear trend. Section 3 continues by noting that, after this initial development, there were several critical suggestions along the lines that detrending in this way may not be adequate for some, if not many, economic time series. This led to some methodological developments, such as rolling window estimation, the concept of integration and the possibility of nonlinear trends, which, whilst not taken up at the time, are now an important part of the economist’s ‘toolkit’ for modelling time series. The questioning of the adequacy of polynomial detrending also led to first thoughts about stochastic trends, which is the topic of section 4. In Section 5 we note that detrending is, in essence, just a filtering of the original data, leading to a series that is ‘smoother’ in some sense than the original series; we outline in this section a number of important developments that underlie much of the current view on smoothing time series (for example, the influential Hodrick-Prescott filter). Section 6 concludes.

II. Decomposition of a time series
By the turn of the 20th century, applications of the theory of correlation were becoming popular using economic and social data, but there was an awareness of the potential problems in applying correlation and regression methods to the ‘raw’ time series. For example, Hooker (1901), in examining the correlation between the marriage rate and trade (taken to be the value of exports per capita) over the period 1857–1899, raised an important difficulty with correlation analysis when applied to time series data. As he saw it, if the movements in the two series that are being correlated are produced by a combination of slow secular movements and more rapid, say year to year, changes, then the latter may be highly correlated while the former may be unrelated, so that the overall correlation between the two series may turn out to be small. This is exactly what appeared to be happening with the marriage rate and trade which, over the period, exhibited declining and increasing secular movements, respectively, thus producing a calculated correlation of just 0.18, with a probable error of 0.09. Arguably, what was really of interest was the correlation between the minor oscillations – the short run movements in the series. What was necessary, so it was argued, was to remove the secular component from each series and then correlate the residuals. The practical question, a problem often seen as ‘curve fitting’, was how to remove the secular movements. (For an historical perspective on this issue and its relation to spurious correlation see Aldrich, 1995.)

II.i. Moving Average and Linear Trends

In contrast to removing the sample mean from the series to be correlated, Hooker (1901) proposed a moving average (MA) trend:

“To correlate the oscillations of two curves, I propose that all deviations should be reckoned, not from the average of the whole period, but from the instantaneous average at the moment. The curve or line representing the successive instantaneous averages I propose to call the trend.”, (ibid., p. 486).

Hooker proposed a moving average trend in which, having identified a period of, say, $p$ years, the ‘instantaneous average’ for year $t$ is the average over the period of length $p$, where $t$ is the middle year. If $p$ is odd, then $t$ is exactly the median year of the period.

Linear detrending based on a prior least squares regression was also popular and, indeed, became the default practice. The example in Figure 1 is from Persons et al. (1922, Charts 16 and 17), who constructed an Index of British Economic Conditions using the methodology established by Persons (1916) for the construction of a ‘business barometer’ for the United States. The series illustrated here is for the freight
receipts of sixteen British railways, where the linearly detrended series is shown in the lower chart.

“Figure 1 here”

**II.ii. Hooker: differencing**

Hooker (1905) returned to the issue of ‘detrending’, but suggested an alternative method for removing secular trends, that of differencing, which “consists simply in calculating the correlation coefficients of the differences between successive values of two variables” (ibid., p. 697). This correlation coefficient was applied to daily changes of corn prices, finding, unlike in the marriage and trade example, that the absolute sizes of the correlation coefficients so obtained were considerably smaller than (often less than half the size of) the corresponding correlation coefficients calculated from the levels. The conclusion drawn by Hooker (1905) from this analysis seems particularly prescient when viewed from a modern perspective:

“Correlation of the deviations from an instantaneous average (or trend) may be adopted to test the similarity of more or less marked periodic influences. Correlation of the difference between successive values will probably prove most useful in cases where the similarity of the shorter rapid changes (with no apparent periodicity) are the subject of investigation, or where the normal level of one or both series does not remain constant. It may even, in certain cases, be desirable to combine the two methods, and to correlate the deviations from the mean in the one series with the successive changes of the other.” (ibid., p. 703, italics in original)

Hooker was thus clearly aware of the distinction between what are now called integrated processes and of the difficulties inherent in modelling the relationships between series of different orders of integration.

**III. Developments of concepts and methods**

The debate on detrending versus differencing is reasonably well known, with significant contributions being Student (1914), Cave and Pearson (1914), Elderton and Pearson (1915), Persons (1917), Yule (1921) and Pearson and Elderton (1923). The ‘points’ winner, in practical terms, of this debate was detrending, although, in an important and largely unnoticed contribution, Smith (1926) suggested a methodology that combined both methods (see Mills, 2011). In contrast to this well-known debate, this section outlines some variations to fitting linear trends either to the original data
or to their logarithms, thus indicating the breadth of thinking at the time, even though in large part these developments were generally not taken up either at all or until much later on.

**III.i. Rolling window estimation**

Hall (1925) suggested what we would now recognise as a rolling window method of estimation rather than simple least squares (LS). He motivated his concerns as follows:

“the secular trend may be a continuous function the form of which we do not know and which can only be represented by a group of known functions.” (ibid., p. 15)

The idea is to choose a window length, which Hall refers to as a ‘period’, say \( p \), which is fixed; the LS estimate of the trend is then calculated for each window, which successively increment by one period and, therefore, lose one observation from the beginning of the previous window. Hall used \( p = 4 \) to illustrate his method, with \( T = 24 \) observations; starting at \( t = 1 \), there are then 21 windows of length 4: in general, there will be \( M \) such windows in a sample. In discussing the choice of \( p \), Hall suggests using the period of the cycle (and thus agrees with Hooker, 1901, who uses a moving average); too short a period relative to the cycle may confuse the trend with the cycle, but too long a period does not achieve the aim of inducing some sensitivity into the trend estimates. Evidently in such a window, which adopts an ‘add one, drop one’ strategy, the change in the LS estimates depends on the difference between observations \( (m - 1)p + 1 \) and \( mp + 1 \), \( m = 1, \ldots, M \); if these are equal, the trend estimate will be unchanged.

Rolling window estimation is popular in a number of present day contexts, for example in unit root tests (Banerjee et al., 1992), in estimation in the presence of structural breaks and in forecasting (for example, Pesaran and Timmermann, 2007), and is available as an estimation option in some commercial software packages, for example STATA. However, Hall’s (1925) contribution in this development seems to have been overlooked.

Hall illustrated his approach using Bradstreet’s Index of Prices, a quarterly series available from 1900q1 to 1924q2, and we use these data to revisit Hall’s results and to demonstrate the different trends and cycles that result\(^1\). He noted a possible break in

---

\(^1\) All programming herein uses MATLAB.
the price series effective from 1915q1, so he estimated one linear trend for 1903q1 to 1914q4 and one for 1915q1 to 1921q4 (although the available data extended to 1924q2, and curtailed the sample for the simple trend as “the fixed trend is not adaptable to approximately current data” (ibid., p. 17)). The resulting trends and corresponding cycles are shown in Figures 2 and 3; Hall’s preference was to use the residuals relative to the trend, which we follow here and below (as % of trend). Evidently, the rolling window trend, referred to by Hall as the ‘moving secular trend’, tracks the data much more closely than a standard linear trend, even one with a break.

“Figure 2 here”

“Figure 3 here”

In a subsequent article, Hall (1926) suggested a method that was less computationally intensive than computing the LS estimates for M windows. Instead of computing M individual LS trend regressions and taking the ordinates at the sequentially fitted values \( p, p+1, \ldots, T \), Hall (ibid.,) proposed estimating the linear trends at four observations apart (assuming quarterly data for illustrative purposes), so that the sequence of windows ends at \( p, p+4, \ldots, T \) (for simplicity \( T \) is assumed to be an integer multiple of \( p \)). Then, to complete the missing quarters, we interpolate linearly between \( p \) and \( p+4 \), between \( p+4 \) and \( p+8 \), and so on. If \( T \) is not an integer multiple of \( p \) (for example, suppose that two quarters are needed to complete the period), then a rolling regression of period \( p \) is fitted in the usual way to temporarily obtain the missing trend ordinates, these being adjusted retrospectively once the 3rd and 4th quarter observations become available.

III.ii. Moving Integration

Hall (1925) also has some interesting comments on how to obtain the (relative) cycle (and so implicitly the trend) by a procedure described as moving integration. Let the original data as a function of time be denoted \( f(t) \); if the function is continuous, then the definite integral over a period of length \( a \) is \( \int_{-a}^{a} f(t) dt \) and, using discrete data, this can be approximated as \( \sum_{i=k-a}^{i=k} y_i \) (notationally we use \( y_i \), whereas Hall uses \( x_i \) and, implicitly, \( \Delta t = 1 \)). Decomposing \( y_i \) into the three components of trend \( u_i \), seasonal \( v_i \), and cycle \( w_i \), then \( y_i = u_i + v_i + w_i \). Illustrating with quarterly data over the period of a cycle of length 12, then:

\[
\sum_{i=k-1}^{i=k} y_i = \sum_{i=k-1}^{i=k} u_i + \sum_{i=k-1}^{i=k} v_i + \sum_{i=k-1}^{i=k} w_i,
\]
If the seasonal effects are (exactly) periodic (and hence cancel over the seasonal period), then \( \sum_{i=k-1}^{i=k} v_i = 0 \) and, if the period of the cycle is indeed 3 years, then \( \sum_{i=k-1}^{i=k} w_i = 0 \), so that \( \sum_{i=k-1}^{i=k} y_i = \sum_{i=k-1}^{i=k} u_i \), the equality holding over the period of the cycle. This last equality might be thought to be contentious, unlike the cancelling of the seasonal components, embodying as it does the assumption that the up and down phases of the cycle cancel. However, replacing the equality assumption with approximation does not damage Hall’s general argument.

Thus, assuming that this equality holds for any period coinciding with the length of the cycle, we have the sequence: 
\[
Y_k = \sum_{i=k-1}^{i=k} y_i, \quad Y_{k+1} = \sum_{i=k-1}^{i=k+1} y_i, \quad Y_{k+2} = \sum_{i=k-2}^{i=k+2} y_i, \ldots,
\]
which Hall describes as ‘moving integration’, in this case ‘three year moving integrals’ of the secular trend, referred to here as \( MI(12) \). Next, as the seasonal effects are eliminated over a year, 
\[
\sum_{i=k}^{i=k-3} y_i = \sum_{i=k}^{i=k-3} u_i + \sum_{i=k}^{i=k-3} w_i,
\]
and again a moving integral can be formed, this time over a year, 
\[
X_k = \sum_{i=k}^{i=k+1} y_i, \quad X_{k+1} = \sum_{i=k-1}^{i=k+2} y_i, \quad X_{k+2} = \sum_{i=k-2}^{i=k+1} y_i, \ldots;
\]
these are the one-year moving integrals of the secular trend plus cycle.

To obtain a relative cycle two chain indices are constructed: first, the secular trend plus cycle based on one year of integration: \( X_k / X_{k-1} \); second, the secular trend based on three years of integration: \( Y_k / Y_{k-1} \). The difference between the two indices is the relative cycle (note here as residuals relative to the trend):
\[
C_k = X_k / X_{k-1} - Y_k / Y_{k-1}
\]
Intuitively this difference is due to the cycle being the difference of the trend from the trend plus cycle. Equally of interest, although not emphasised directly by Hall, is the relative trend \( Y_k / Y_{k-1} \). As with the LS trends, he illustrated this moving integration method using Bradstreet’s Index of Prices. The resulting \( MI(12) \) relative trend and chained actuals, \( y_k / y_{k-1} \), are shown in Figure 4; whereas Figure 5 shows the relative \( MI \) cycle and, for comparison, the cycle derived from the moving secular trend. The correlation coefficient between the two cycles is 0.9, which supports Hall’s view “that the results from the moving secular trend and moving integration methods are in close agreement.” (ibid., p. 22)

Although innovative, Hall’s ideas, both on rolling window estimation and (moving) integration and summation, were not picked up until very much later and, as far as we can tell, without acknowledgement to his work.
III.iii. The arctan trend

Carmichael (1928) observed that a number of time series did not exhibit uniformly smooth growth about which there was cyclical movement. This led him to consider a nonlinear trend that would be suitable in three circumstances: (i) inappropriate projection of a negative linear trend, leading, for example, to “negative or ridiculously small positive values when comparatively large positive values only are possible” (ibid, p. 13); (ii) approximately linear growth that is resumed after interruption by an abrupt change in level; (iii) as (ii) but with a first interruption, for example a sharp drop, followed by another abrupt change in level, before resumption of the previous growth.

“Figure 4 here”

“Figure 5 here”

Carmichael addressed the question of what form could capture such movements and noted that a central criterion for a trend is “to emphasize the importance of reasonable projection into the future” (ibid, p. 13, fn 2). His suggested form uses the arctangent, the properties of which we briefly review. Throughout Carmichael worked in degrees rather than radians. The tangent of an angle, $z$, that varies from $-90^\circ$ to $90^\circ$, varies itself from $-\infty$ to $+\infty$; in general, if $x = \tan(z)$ then $z = \tan^{-1}(x)$, where $\tan^{-1}$ is the inverse function referred to as $\arctan(z)$; $z$ is in degrees and the conversion to radians, which are now used more frequently, is $y = z(\pi/180)$. Thus, in radians, the limits of $\arctan(z)$ are $\pm(\pi/2) = \pm 1.5708$ and, therefore, by using a suitable scaling constant, say $c$, the limits can be adjusted: for example, setting $c = \pi^{-1}$ results in limits of $\pm 1/2$.

Carmichael’s suggestion was to use terms of the form $c \arctan(x)$ for the nonlinear trend. As we illustrate below, under an appropriate choice of parameters, the arctan function is close to the logistic function $\gamma/(1 + \exp(\psi x))$, favoured at the time in population studies (for example, Pearl and Reed, 1923). Carmichael’s contribution, however, was more than to suggest an alternative nonlinear trend, for he showed that, in combination with a linear trend, the arctan function can be used to model what we would now regard as structural breaks, an idea that was taken up much later in the work of Teräsvirta using smooth transition functions (see for example Teräsvirta, 1994).
The arctan function is illustrated in the top panel of Figure 6, along with the centred logistic function \((1 + \exp(yx))^{-1} - 0.5\), where, for comparability, the arctan function is in radians and the calibration is \(\psi = -4c\), \(c = \pi^{-1}\). The derivatives of the arctan and logistic functions are shown in the bottom panel of Figure 6; both approach zero as \(x \to \pm\infty\) and are clearly similar elsewhere.

Carmichael suggested the following functional relationships to capture his three previously delineated cases:

(i) \(y = a + c\ \text{arctan}(x)\)
(ii) \(y = a + bx + c\ \text{arctan}(x)\)
(iii) \(y = a + bx + c\ \text{arctan}(x) + d\ \text{arctan}(\alpha x + \beta)\)

We would now conventionally add time subscripts to the variables \(y\) and \(x\) and a random term, \(\varepsilon\), say.

“Figure 6 here”

In contrast, the standard linear trend model is \(y = a + bx\). The variable \(y\) may either be the original data or their logarithms (note that Carmichael used logs to the base 10, rather than to the base \(e\)). The variable \(x\) is a ‘distance’ measure relative to an origin \(t_0 = 0\). Case (i) is a simple arctan trend, which allows for a positive asymptote; case (ii) allows for one ‘interruption’ to the linear trend; and case (iii) allows two ‘interruptions’. In case (iii), the first \text{arctan} component allows what Carmichael (ibid, p. 14) describes as a rapid change in the “level as \(x\) increases through zero” and the second component allows a second change through the time-shifted point \(\alpha x + \beta\). What Carmichael is effectively describing, in current terminology, is a switching model with a single or double smooth transition adjustment function.

We use one of Carmichael’s examples to illustrate the use of the arctan trend. The time series is of annual data for Anaconda Copper common stock prices over the period 1908 to 1926. Whilst the linear trend’s fitted values are invariant to the origin of the time variable (although the estimate of the intercept will change), that is not the case with the arctan trend, where it is necessary to choose the origin \(x = 0\) for the time variable and a scale for yearly changes in \(x\). In this case, the problem addressed by Carmichael is that of a fairly rapid decline in the series approaching 1924, the projection of which would lead to negative prices. Carmichael acknowledged that his method required matters of judgement as to the timing of the switch point and the
speed of adjustment, as represented in the choice of \( x \). On these he wrote: “The necessity for arbitrary choice of origin and unit on the \( x \)-axis is admittedly objectionable. Could it [the arbitrary choices of scale and origin] be obviated without loss of flexibility of the curve, the technique would be vastly improved.” (Clarification in [.] brackets.) These are relatively simple matters for current econometric techniques to determine.

The estimated linear trend, taking \( x = 0 \) for 1917, was \( y = 65.18 - 2.870x \); the arctan trend, taking \( x = 0 \) for 1900 and \( \Delta x = 0.9 \), was \( y = 863.29 - 9.304\arctan(x) \) (other variations were tried by Carmichael). The series and the estimated trends are shown in Figure 7; the trends are projected through to 1935 to show the effect of the different trend assumptions, in particular the LS trend projects negative stock prices, whereas there is a positive limit to the arctan projections, with a limiting value of 26.02.

The corresponding LS and arctan cycles are shown in Figure 8. In determining \( \Delta x \) and \( t_0 \), based on a graphical inspection of the time series, Carmichael tried two possible origins and five values of \( \Delta x \), not all of which led to an asymptote that was positive. Whilst the arctan model is nonlinear, once the dependence on \( \Delta x \) and the switch point are taken into account, it is simple enough to extend the conditional least squares search over a two-dimensional grid of possible values of \( \Delta x \) and \( t_0 \), say \( S = X \times T \), which will result in a consistent estimation method provided that the population values of \( \Delta x \) and \( t_0 \) are included in \( S \).

**IV. The beginning of stochastic trends**

Hall’s (1926) idea of using rolling window estimation, in effect to approximate an unknown and flexible form for the trend, resulted from the view that “the secular trend may be a continuous function the form of which we do not know and which can only be represented by a group of known functions” (ibid., p. 15). Whilst this was a development of the linear trend assumption, it was still framed in the context of a (stationary) decomposition of a series into “its secular, seasonal and cyclical components” (ibid, p. 14). Hall’s moving integration method, although implicitly quite ‘revolutionary’, was seen as part of the same development. However, if we add to this Carmichael’s (1928) development of smooth transition adjustment functions, together with contributions to be outlined in this section by Karsten (1924, 1926) and von Szeliski (1929), we see in the 1920s the beginning of the development of nonstationarity and stochastic trends as applied to economic time series.

“Figure 7 here”
IV.i. Quadrature
Although received with some criticism, Karsten published two articles, in 1924 and 1926, which showed a detailed understanding of how some economic time series could be generated as partial sums of others and that there could, therefore, be a pair of series where one was the cumulation of the other. Underlying Karsten’s ideas was the concept of quadrature, due to Charles Edge of the ‘Quadrature Bureau’.

“Two forces are said by Mr. Edge to be in quadrature when they trace curves such that the fluctuations of one of the curves correspond to the fluctuations of a curve of the integration or cumulation of the data of the other curve. When a cause-and-effect relation exists between such two phenomena, the first may then be said to be cumulatively affected by the second.” (Karsten, 1924, p. 14).

Moreover on the detail:

“The quadrature method therefore applies cumulation to obtain a series representing the integral function, and differencing to obtain a series of data representing the derivative.” (ibid., p. 15)

A pair of series (or curves), x and y, are said to be in quadrature if they are related such that if x crosses the zero axis then y has a turning point. The simplest (but not the only) cases are those that have a stock-flow relationship, and Karsten motivates quadrature with two such examples, one of a bank account that cumulates the net deposits (series x) into a running balance (series y) and the other the net flow into a reservoir (series x) and the level of the reservoir (series y). Karsten put it thus: “the

2 Karsten (1924, p. 15) gives a footnote of definitions of quadrature, showing that the term has several meanings in use at the time, for example as a method of numerical integration and the alignment of planets; its use here concerns the phase difference between two signals (series), a quarter cycle (90°) difference leading to the use of the term quadrature. For example, if a time series is generated by a sine curve, cumulating the terms results in a cosine curve that is displaced by a quarter of a period relative to the sine curve. However, Karsten recognised that economic time series are not as simple as sine curves: “That economic forces usually fluctuate or pulsate, so that the curves of economic statistics are periodic or cyclic, is a fact with which everyone is familiar. Sometimes the waves are very smooth, sometimes they seem cut up with short minor waves, presenting a very choppy appearance. Sometimes the primary or long-time waves are fairly regular, reminding one of sine curves; sometimes they are very irregular, large waves and small waves, short and long, all following close upon each other. But whatever their form, these wave-like irregularities are present in almost all economic data.” (ibid., p. 15)
peak and valley of one curve appearing simultaneously with the intersection of the zero-line of the other” (ibid., p. 24).

A contemporary example of what Karsten meant by quadrature can be illustrated by the now familiar case of a (pure) random walk. Consider a white noise sequence \( \{ \varepsilon_t \}_{t=1}^T \) and the partial sum sequence \( \{ y_t \}_{t=1}^T = \{ \sum_{s=1}^t \varepsilon_s \}_{t=1}^T \). By construction these paired series are in quadrature: whenever \( \varepsilon_t = 0 \) or changes sign, then there is a local peak or trough in \( y_t \).

The context of Karsten’s (1926) article are the three curves, A, B and C, of the Harvard Business Indexes: the A curve is the index of speculation conditions (on the New York Stock Exchange); the B curve that of business conditions which, though variably defined, were in essence based on bank clearings; and the C curve is the index of money and banking conditions, essentially an interest rate measure. (The indexes are based on seasonally adjusted and deterministically detrended series, so their value of ‘normalcy’ is zero.) Karsten suggested that the series A and B were a pair of series in quadrature, that is with A being the integral and B the integrand, so that local peaks and troughs in stock market prices were paired with business conditions crossing the ‘normal trend or zero line’.

In Karsten’s view, the mistake made by the Harvard Committee was ‘which of the two really moves first’ (ibid., p. 405), since they viewed the A curve as preceding the B curve but, in quadrature, series A is the cumulative series and cannot precede the B series. He commented:

“Perhaps it escaped the attention of the Harvard Committee that the curve of the level of prices in the stock market, Harvard Curve A, has a top or bottom, a high or low-water mark, whenever the curve of the flow of money into or out of the stock market from or into the business world - represented by Harvard Curve B, the curve of business conditions - crosses its normal trend or zero line.” (ibid., p. 404)

To examine Karsten’s assertion from a modern perspective, we consider the time series properties of the Harvard A and B series. A necessary condition for Karsten to be correct is that if B is an I(0) series, then A should be I(1). The data are monthly for the period 1919m1 to 1929m12, and the A, B and C series are graphed in Figure 9.

“Figure 9 here”
A visual inspection suggests that series A and B do seem to have different characteristics. As Karsten only had information through to the end of 1925 available to him, the first sample period is taken to end in 1925m12. The Dickey-Fuller (DF) type unit root test statistics are reported in Table 1 (the C series is referred to below), and these confirm Karsten’s assertion: the null hypothesis of a unit root is clearly not rejected for series A, whereas there is a rejection in the case of series B, which is the case for both sample periods.

### Table 1 ADF tests for Harvard Indexes A, B and C

<table>
<thead>
<tr>
<th></th>
<th>Series A:</th>
<th>Series B:</th>
<th>Series C:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$I(1)$, $\hat{\beta}$</td>
<td>$I(1)$, $\hat{\mu}$</td>
<td>$I(1)$, $\hat{\beta}$; $I(2)$, $\hat{\mu}$</td>
</tr>
<tr>
<td>1919:4 – 1929:12</td>
<td>$-1.638$</td>
<td>$-3.244$</td>
<td>$-3.222$ $-12.983$</td>
</tr>
</tbody>
</table>

Notes: the maintained regression for the test statistic $\hat{\beta}$ includes an intercept and trend, $\hat{\mu}$ includes an intercept; the lag lengths were chosen by AIC. 5\% critical values are approximately $-3.47$ and $-3.45$ for $\hat{\beta}$ and $-2.88$ and $-2.90$ for $\hat{\mu}$ (smaller sample size first). The null of a second unit root for each series was uniformly rejected.
Karsten also considers the relationship between the C series and, jointly, the A and B series. He views these as being in ‘double quadrature’: “the C Curve is to a great extent the cumulative of both the A and B Curves”. The C series relates to the price of money, which “is largely determined by the demand for money either for business purposes or for investment purposes, that is, either in the business world or in the security markets” (ibid., p. 415). Thus, in Karsten’s view, the driving force of the A and C curves is the B curve, which is not itself explained: “To explain Curve B we must have recourse to other business statistics” (ibid., p. 417).

Schematically, Karsten’s implicit model may be expressed as:

\[ C_t = \alpha_1 \sum_{t=1}^{t'} A_t + \alpha_2 \sum_{t=1}^{t'} B_t \]

\[ = \alpha_1 \sum_{t=1}^{t'} \sum_{j=1}^{t'} B_j + \alpha_2 \sum_{t=1}^{t'} B_t \]

In time series terms, if the B series is \( I(0) \), then the C series is \( I(2) \). The ADF tests for series C (see Table 1) point to a firm rejection of the null of two unit roots in favour of one unit root. However, Karsten’s arguments show an implicit awareness of the operations of single and double partial summation that underlie the generation of stochastic, as opposed to deterministic, trends.

**IV.ii. The response to quadrature**

Karsten’s (1926) criticism of the interpretation of the Harvard ABC curves was not generally well received, prompting responses from the Harvard group, Bullock, Persons and Crum (BPC, 1927), and Hansen (1927)\(^3\); von Szelsiki (1929) was critical of the conceptual basis of the method and (much) later Sasuly (1947, p. 267) referred to quadrature as the ‘correlation hoax’.\(^4\)

---

\(^3\) Hansen (1927) was not critical of the idea that one series (the A curve) could be the cumulation of another (the B curve), but that even in that case one may have forecasting power for the other: “ Mr. Karsten’s own article gives an excellent illustration of the great value of the Harvard A Curve as a forecaster. He concludes that because of the interrelation of the A Curve and the B Curve, the A Curve begins to turn down (or up, as the case may be) at the point when the B Curve is crossing the normal line. From this he establishes his trend for the B Curve. He uses the A Curve to determine the exact point at which the B Curve crosses the normal line. Why? Because it is a simple matter to detect the high and low points on the A Curve, but it is not so simple to find the points at which B crosses the normal line. It is just for this reason that the A Curve is useful as a forecaster.” (ibid, p. 369)

\(^4\) The idea that Karsten’s results were a hoax related in part to the difficulty that BPC had in determining the method of detrending that was used by Karsten (see BPC, 1927, p. 88, for a summary of the correspondence on this issue). Referring to Karsten’s proof of quadrature for the Harvard A and B curves, BPC commented: “His "statistical proof" consists in premising the quadrature theory,
The Harvard reply, although critical of Karsten’s theory and methods, contained some important insights into how stochastic trends arise. BPC refer to the cumulation method as comprising two series: series I, the generating (or original) series, and series II, the cumulative series (of the items of series I). Series I may be a transformed series, for example, deviations from trend or ‘normal’ values, so as to result in a sequence comprising both positive and negative values. To illustrate what BPC saw as the erroneous and arbitrary nature of Karsten’s cumulation argument, they generated random numbers for series I and then cumulated these into series II. The data for this, one of the first Monte Carlo experiments, was generated as follows. BPC took the following sentence: "The cumulation process creates cyclical fluctuations out of random sequences quite common in numbers of all kinds," assigned to each letter its numerical place in the alphabet, with A assigned 0, and demeaned the resulting series. They found that the cumulative series (series II) is recognisable as having the characteristics of a random walk (ibid, p. Chart A).

BPC noted that:

“It should be remarked especially that the successive positive and negative items of the generating series need not be arranged by groups in a wave-like form in order that the resulting cumulative series shall exhibit wave-like fluctuations. On the contrary it is quite possible for a very disordered succession of positive and negative items when used as a generating series to yield a cumulative series in which the oscillatory fluctuations partake of the wave-like form encountered in these economic movements. … Despite the lack of relevancy of the quoted assertion to the problem of securing an index of business fluctuations and the wholly accidental sequence of numbers used as a generating series, it is a notable fact that the resulting cumulative series is itself a not unacceptable picture of the fluctuations in general business conditions since the war”. (ibid, p. 82)

These observations were not, however, regarded constructively. BPC observed that the cumulation of random inputs led to the creation of cycles, an exaggeration of cyclical fluctuations, ‘pronounced undulating movements’, and what we would term as ‘spurious’ correlation. On the latter they noted:

______________________________

deliberately manufacturing a high coefficient of correlation to support it, and offering the coefficient thus secured as proof of the validity of the quadrature theory.” (ibid, p. 88)
“Consequently, when the cumulation process has been used, high values of coefficients of correlation between any two series (or similarity between the curves) are very doubtful evidence that the series are related or even vary together.” (ibid, p. 83)

**IV.iii. Further experimental evidence**

BPC had, with their simple simulations, highlighted a number of the key characteristics of $I(1)$ processes and the (potentially spurious) correlation of two $I(1)$ series. These implications were not developed constructively, but von Szeliski (1929) took the BPC analysis further at an experimental level:

“The purpose of this paper is to show that correlations between time sequences are of little significance in themselves, and in particular that the process of cumulation or quadrature described by Karl G. Karsten may introduce spurious correlation and pseudo-cycles.” (ibid., p. 241)

Szeliski undertook a number of simulations and, to illustrate his method and arguments, we recreate his Experiment 1, which was to take ‘random’ draws from slips of paper numbered from −4 to +4, with replacement. There were $n = 110$ such draws, resulting in the sequence $\{x_t\}_{t=1}^{t=110}$; he then created the partial sum of the $x_t$, so that $y_t = \sum_{s=1}^{t} x_s$. A trend was then fitted by least squares to $y_t$, say $\tilde{y}_t$, with Szeliski’s charts suggesting that a 2nd or 3rd order polynomial was used depending on the visual properties of the cumulated series, and, finally, the deviations from the fitted trend, $y_t - \tilde{y}_t$, were cumulated into $z_t = \sum_{s=1}^{t} (y_s - \tilde{y}_s)$.

The recreated results for one such experiment are shown in Figures 10 and 11. Szeliski’s rationale for this experiment was to show that the partial sum has trend-like behaviour (see Figure 10), which we would now recognise as a stochastic trend, and thus, in the methodology of the time, would first be deterministically detrended using a low order polynomial fitted by least squares. The result is a series that shows cyclical or pseudo-periodic behaviour (see Figure 11).

“Figure 10 here”

“Figure 11 here”
In von Szeliski’s original Experiment 1, he found the correlation between series A and B (the first two in his simulation) was 0.0105, whereas the correlation between the corresponding series as cumulated deviations from trend was –0.847. We can confirm the apparently spurious nature of the correlation in the cumulated series by replicating the experiment $10^6$ times; the bivariate correlations converge to 0.076, to 0.421 for the cumulative series and to 0.237 for the deviations from trend series.

Von Szeliski (1929) carried out a number of other experiments to illustrate the consequences of cumulating quasi-random series, such as coin tossing and dice rolling. As a result, he drew five conclusions (ibid., p.245):

1. Series comprising cumulated random items exhibit “every characteristic of economic time series, except seasonal variation.”
2. “The significance (and perhaps the usefulness) of secular trend, considered merely as a problem in curve fitting, is uncertain, not only because long-term undulations which we know to have no real significance appear in random series, but because the trend is largely indeterminate.”
3. “The process of cumulation increases the amplitude of the longer cycles.”
4. “The validity (or usefulness) of the deviations-from-trend method of discovering relationships between time series is open to question, because it may show a correlation where in reality none exists, or it may incorrectly measure a true relationship … .”
5. “The cumulation of time series tends to introduce spurious correlation. This has been abundantly illustrated above. It need occasion no surprise.”

Although the connections are not yet drawn together, one comment is worth recalling:

“Even a Ph.D from a school of business could not tell which of two graphs was, let us say, a wholesale price index, and which the cumulation of successive throws of a dice.” (ibid., p.)

---

5 Szeliski commented: “Chance sequences were obtained from a succession of dice throws. These were cumulated, and the results graphed. They had a striking resemblance to the “point charts” of a stock trader. ‘Chart fiend’ would doubtless pronounce them typical examples of market action, and point out the periods of accumulation, resistance points, double tops, and so on. The writer could not resist the temptation of these charts, and made quite a lot of paper money ‘playing’ the indications as the dice throws ran the ‘price’ up and down the scale. The possibilities of the cumulative method are remarkable.” (ibid., p. 245).
In retrospect it is the inverse question that is of interest: were trends and business cycles generated by the cumulation of random causes? Kuznets (1929) was wary of making the inverse inference that actually observed cycles were indeed caused by cumulating random shocks.

“can one invert the proposition and say that, therefore, cyclical oscillations may be conceived primarily as results of summations of random causes, and that the characteristics of some of these cyclical oscillations can best be grasped as a result of the underlying random events or of the process of cumulation?” (ibid., p.)

The properties of series that resulted from cumulating underlying random sequences were becoming known in the late 1920s and early 1930s: Slutsky’s famous paper on the generation of spurious cycles was published in Russian in 1927 and in English in 1937. Kuznets (1929) published an insightful paper on the generation of cyclical movements by cumulation or constructing a moving average, but he ruled out the possibility that time series could be generated by a simple cumulation of all past shocks:

‘(w)e shall omit the straight cumulation altogether, because it is too far removed from reality. In economic life, there is no perpetual influence of an event once occurred on all the subsequent events.” (ibid., p. 269).

IV.iv. Working

The essentially stationary view of the world expressed by BPC, Kuznets and others, was not shared by Working (1934)⁶, who was aware of a growing sense of the importance of what we now recognise as I(1) series, even though the standard practice at the time was very much within a trend stationary framework:

“It has several times been noted that time series commonly possess in many respects the characteristics of series of cumulated random numbers. The separate items in such time series are by no means random in character, but the

---

⁶ Working notes that “It has several times been noted that time series commonly possess in many respects the characteristics of series of cumulated random numbers. The separate items in such time series are by no means random in character, but the changes between successive items tend to be largely random. This characteristic has been noted conspicuously in sensitive commodity prices. On the basis of the differences between chain and fixed-base index numbers King has concluded that stock prices resemble cumulations of purely random changes even more strongly than do commodity prices.” (ibid., p. 11)
changes between successive items tend to be largely random. This characteristic has been noted conspicuously in sensitive commodity prices. …

The fact that series commonly used as indexes of business activity closely resemble series obtainable by cumulating random numbers has given support to the theory that so called business cycles result in large degree from cumulative effects of independent random influences bearing on the business situation – some favourably, some unfavourably.” (ibid., p. 11)

Working (ibid.,) noted that some of the practices in drawing ‘random’ numbers to illustrate the characteristics of cumulated series were unlikely to produce a random series; for example, assigning numbers to the letters of the alphabet and then taking particular sentences and constructing the sequence of associated numbers. Working drew his random numbers from tables of Tippet’s Random Sampling Numbers (Tippet, 1927), which were converted to form draws from a normal distribution with a known standard deviation. He then provided a sequence of 2,399 random numbers obtained in this way, which were cumulated into a partial sum sequence and, for presentational purposes this series was arranged as 46 years of weekly data. This was an example of what Working called a random-difference series. The resulting series showed a noticeable negative ‘trend’, which Working observed might ‘excite suspicion’ that the original series had a negative mean which was not, however, supported by a test to that effect. In fact, the series so created was an example of a largely negative random walk.

This artificially created series was then used as a basis of comparison with actual time series and Working (as previous authors had done) noted the similarity with actual time series. But, unlike BPC (1927) and von Szeliski (1929), Working did take the view that partial sum (or random-difference) series were relevant to economic analysis and not an irrelevant construct. He took the view that actual series, with the characteristics of trend and cycle, could have been generated by the cumulation of shocks.

Working compared his artificial time series to a series on wheat prices, which had been a subject of interest to him. Whilst he noted some similar characteristics, some differed, particularly in the bimodal nature of changes in wheat prices, and he offered what transpired to be a prescient comment:
“I find that to the important extent that wheat prices resemble a random-difference series, they resemble most closely one that might be derived by cumulating random numbers drawn from a slightly skewed population of standard deviation varying rather systematically through time.”, (ibid., p. 24).

From a modern perspective, Working is clearly suggesting a martingale difference sequence having time varying variances, which may be interpreted as measuring volatility.

V. Time series smoothers and filters

Although moving averages had been introduced by Hooker (1901) for the purposes of smoothing a time series and extracting a trend, more sophisticated applications of the method were developed in the 1920s. These had their genesis in actuarial science, where the procedure of ‘graduation’, by which crude mortality rates are transformed to produce smoother estimates of mortality, uses iterated moving averages, as originally suggested by Spencer (1904, 1907) and outlined in Whittaker and Robinson (1924, chapter XI).

Suppose that the primary series to be graduated is \( u_t \). Whittaker and Robinson introduced the notation

\[
[2m+1]u_t = \sum_{j=-m}^{m} u_{t+j}
\]

to denote the sum of \( 2m+1 \) \( u \)'s centered on \( u_t \), and the following generalization to weighted moving sums,

\[
[w_{-m}, \ldots, w_0, \ldots, w_m]u_t = \sum_{j=-m}^{m} w_j u_{t+j}
\]

Spencer’s 21-term moving average can then be written as the iterated moving average

\[
v_i = \left[ \frac{5}{5} \right]\left[ \frac{7}{5} \right]\left[ \frac{4,9,4}{7} \right] u_t
\]

This may be expanded to obtain

\[
v_i = \frac{1}{355}\{60u_t + 57(u_{t-1} + u_{t+1}) + 47(u_{t-2} + u_{t+2}) + 33(u_{t-3} + u_{t+3})
+ 18(u_{t-4} + u_{t+4}) + 6(u_{t-5} + u_{t+5}) - 2(u_{t-6} + u_{t+6}) - 5(u_{t-7} + u_{t+7})
- 5(u_{t-8} + u_{t+8}) - 3(u_{t-9} + u_{t+9}) - (u_{t-10} + u_{t+10}) \}
\]

or
\[ v_t = 0.171u_t + 0.163(u_{t-1} + u_{t+1}) + 0.134(u_{t-2} + u_{t+2}) + 0.094(u_{t-3} + u_{t+3}) \\
+ 0.051(u_{t-4} + u_{t+4}) + 0.017(u_{t-5} + u_{t+5}) - 0.006(u_{t-6} + u_{t+6}) - 0.014(u_{t-7} + u_{t+7}) \\
- 0.014(u_{t-8} + u_{t+8}) - 0.009(u_{t-9} + u_{t+9}) - 0.003(u_{t-10} + u_{t+10}) \]

which is a symmetric moving average containing ten leads and ten lags of the primary series \( u_t \) to obtain the graduation \( v_t \). A property of such a graduation is that it will reproduce a cubic polynomial in time without distortion.

In a sequence of papers, Henderson (1916, 1924) and Whittaker (1923, 1924) independently considered the problem of designing a smoother (later more commonly known as a filter) that, as well as reproducing a cubic polynomial trend without distortion, would also satisfy certain smoothness conditions. The primary condition was that the filter should minimize the variance of the third differences of the smoothed series, i.e., it should minimize \( \text{Var}(\Delta^3 v_t) \). Henderson’s 23-term moving average, for example, is then given by

\[ v_t = 0.148u_t + 0.138(u_{t-1} + u_{t+1}) + 0.122(u_{t-2} + u_{t+2}) + 0.097(u_{t-3} + u_{t+3}) \\
+ 0.068(u_{t-4} + u_{t+4}) + 0.039(u_{t-5} + u_{t+5}) + 0.013(u_{t-6} + u_{t+6}) - 0.005(u_{t-7} + u_{t+7}) \\
- 0.015(u_{t-8} + u_{t+8}) - 0.016(u_{t-9} + u_{t+9}) - 0.011(u_{t-10} + u_{t+10}) - 0.004(u_{t-11} + u_{t+11}) \]

These filters eventually found their way into use as smoothers of economic time series, with Henderson filters being adopted for trend estimation in the X-11 seasonal adjustment procedure, replacing the Spencer moving averages used in earlier versions of the U.S. Bureau of the Census seasonal adjustment procedure.

Almost forty years after their introduction, Leser (1961) revisited the Whittaker-Henderson approach and extended the methodology by deriving the weights using the principle of penalized least squares, in which a linear combination of two sums of squares is minimized. The first sum of squares contains the deviations of the observations \( u_t \) from the filter \( v_t \), the second contains the second differences of successive smoothed values \( \Delta^2 v_t \), with the linear combination of the two being defined by the weights of unity and \( \lambda \), i.e., for the observed sequence \( u_1, u_2, \ldots, u_T \), the minimand is

\[ \sum_{t=1}^{T} (u_t - v_t)^2 + \lambda \sum_{t=2}^{T-1} (\Delta^2 v_{t+1})^2 \]

The first term measures the goodness of fit of the filter, the second penalizes the departure from zero of the variance of the second differences of the filter, so that it is
a measure of smoothness: hence $\lambda$ is referred to as the smoothness parameter. Successive partial differentiation with respect to the sequence $v_i$ leads to the first order conditions

$$\Delta^2 v_{t+2} - 2\Delta^2 v_{t+1} + \Delta^2 v_t = \lambda (u_t - v_t)$$

so that, given $T$ and $\lambda$, $v_t$ is a linear function of $u_t$ with time varying weights:

$$v_t = \sum_{j=1}^T w_{t,j}u_j$$

Leser then developed an algebraic method of obtaining the coefficients $w_{t,j}$, providing a number of examples in which the solutions were obtained in laborious and excruciating detail. However, its historical importance lies in the fact that the method developed by Leser and based upon the Henderson-Whittaker approach was exactly that proposed some two decades later by Hodrick and Prescott (1997) and which has entered into macroeconomic modelling as the H-P filter.

As an example, Whittaker and Robinson (1924) considered graduating, or smoothing, the series generated by:

$$u_t = f(t) + \varepsilon_t = (t - 26) + \frac{1}{10} (t - 26)^3 + \frac{1}{100} (t - 26)^3 + \varepsilon_t$$

where the $\varepsilon_t$ are drawn from a uniform distribution ranging from $-149.5$ to $+149.5$. Figure 12 shows $u_t$ itself, Henderson’s 23-term moving average $v_t$, and the H-P filter using a standard setting of $\lambda = 100$: the latter two graduations are virtually indistinguishable and have a correlation coefficient in excess of 0.9999.

“Figure 12 here”

VI. Concluding Remarks

There is no doubt that one gains from the early literature on correlation and regression analysis a sense of great excitement, arising from the promise of using these ‘new’ tools of statistical analysis to ‘discover’ relationships between variables. However, it is also clear that even from early use of these techniques there was an awareness that a time series may have several co-existent characteristics, the key distinction being between the slowly moving (low frequency) component, usually identified as ‘the trend’, and the more rapid movements, usually referred to as the ‘cycle’. This is a distinction that is familiar today. In what may be seen as the first phase of the development of time series analysis, a fairly standard approach was to detrend the
time series of interest, usually adopting a simple linear trend function, in what we would now refer to as a stationary decomposition of a time series along the lines of an unobserved components model.

In broad terms, after the initial phase of interest, there were some quite fundamental implicit criticisms of this standard practice. This questioning phase we might refer to as phase two. It recognised that linear detrending was a practical (rather than conceptual) solution to the time-correlation problem. The ‘trend’ was in general ill-defined and, anyway, unlikely to be linear. Thus, for example, Hall (1925, p. 14) commented that: “It is improbable that the secular trend is strictly a straight line because it is unlikely that any economic phenomena will be encountered which is growing or increasing at a constant rate.” Hall, therefore, proposed a more flexible way of modelling the trend. In a parallel literature in actuarial science, a key issue was to define the ‘smoothness’ required when filtering a time series, of which a linear trend was a special case. This discussion necessarily involved a definition of the criterion of smoothness, which had been initiated by Sprague (1887) based on the second difference of a time series and was further developed in the 1920s by Henderson and Whittaker, and forms the basis of the well-known Hodrick-Prescott filter.

Also in the 1920s, the questioning of the suitability of the detrending approach, in fact the fundamental question of whether it was even sensible to conceptualise the trend as deterministic, had begun, if not by accident then with an element of unintended consequences in the discussion between Karsten (1926) and Bullock, Persons and Crum (1927). This discussion led to the conclusion that the cumulation of random shocks could lead to a time series that looked remarkably like some ‘real world’ series. Whilst this might lead, as in BPC, to a dismissal of the idea that cumulated or integrated random shocks could be the data generating process, the idea was taken up further by Szeliski (1929), who critically appraised the dominant detrending methodology; an important conclusion being that the ‘trend is largely indeterminate’,

---

7. The question of what constituted a ‘trend’ was still a matter of discussion in the 1930s. For example, Frickey (1934), in a paper entitled ‘The Problem of the Secular Trend’, started from the basis that “the secular trend has often been defined in some such language as the following: the gradual and persistent movement of the series over a period of time which, contrasted with the short run fluctuations of the series, is long.” (ibid., p. 199)

8. The dominance of the linear or loglinear trend was, however, itself persistent (indeed as it is today as the default choice). For example, Bernstein and Cowden (1937) proposed the use of a special paper which plotted the lines corresponding to different constant rates of growth over a particular length of time (known as isoropic lines). They noted that: “The type of paper we have developed for graphic presentation of trend data is designed to eliminate automatically regular geometric growth. It is probable that most important economic series in common use show this type of growth.” (ibid., p.446).
an observation that we now understand more fully in the context of integrated series (in which the trend is stochastic). Unfortunately, Szeliski’s promised theoretical analysis of his results did not materialise (or was not published), but we now recognise that his ‘experiments’ generated integrated series.

Phase two led to the promise of a new interpretation of the generation of economic time series, particularly insofar as the modelling of the trend was concerned\(^9\). It presaged a move away from the simple two-step procedure of fitting the trend by a simple polynomial, possibly allowing for breaks in the trend, using a moving average or a nonlinear trend function, then saving the residuals and using these in the subsequent analysis. However, this promise was not fulfilled until a number of years later.

\(^9\)Despite the promise of the development of what we now refer to as stochastic trends, the methodological paradigm remained broadly the same for a number of years. For example, Kendall (1941) commented: "In recent years the mathematics of trend fitting have been brought to an advanced stage of development. There are various methods advocated by different writers for different purposes; but, putting aside graphical methods, which involve personal judgment and cannot claim serious attention for theoretical work, they all depend on one of two processes: either (a) a polynomial of chosen degree is fitted to the whole series, almost invariably by least squares; or (b) a’ moving average of chosen extent and with chosen weights is used to determine trend values at different points of the series.”, (ibid., p. 43).
References


Knauth, O. W. (1923), Statistical Indexes of Business Conditions and Their Uses, chapter 7 in Committee of the President's Conference on Unemployment, and a Special Staff of the National Bureau, Business Cycles and Unemployment, NBER.


Figure 1 Linear detrending: freight receipts (Persons et. al., 1916)

Chart 10. — Freight Receipts of Sixteen British Railways with Line of Secular Trend

Chart 17. — Adjusted Figures (Cycles) for Freight Receipts of Sixteen British Railways
Figure 4 Bradstreet prices: chained actual and moving integration

Figure 5 Bradstreet prices: comparison of cycles
Figure 8: Anaconda stock prices: estimated cycles

- Dotted line: arctan cycle
- Dashed line: linear trend cycle

Yearly data from 1908 to 1926.