Journal of the History of Economic Thought,
Volume 37, Number 4, December 2015

CARMICHAEL'S ARCTAN TREND: PRECURSOR OF SMOOTH TRANSITION FUNCTIONS

BY

TERENCE C. MILLS AND KERRY PATTERSON

In an almost unreferenced article, Fitzhugh Carmichael (1928), writing of the period around the First World War, noted that "during the past twelve years many economic series have undergone what appears to be a permanent change in level." These are prescient words that are widely applicable today. Carmichael noted that the then-standard practice of linear detrending was inappropriate in the presence of what we would now call "structural breaks"; as a result he proposed a method that would not only model a nonlinear trend, but would be suitable for situations where the transition from one regime to another was smooth. This study establishes the precedence of Carmichael's ideas, re-examines his methods, and solves the problems that he thought would hinder wider applications of his approach, which has since become a central part of contemporary nonlinear econometric methods and for which Carmichael should be given credit.

I. INTRODUCTION

The purpose of this paper is to draw attention to, and re-examine, an important and rarely acknowledged paper by Fitzhugh Carmichael (1928),[1] who introduced the arc

Terence C. Mills, Loughborough University; Kerry Patterson, University of Reading. The authors gratefully acknowledge the constructive comments of two anonymous referees and the encouragement of the editor. We would also like to thank Amy Catlin of the Registrar's Office of the University of Denver for biographical information on Professor Carmichael.

[1]Professor Fitzhugh L. Carmichael was born on October 29, 1893, and died on June 11, 1990, aged ninety-six. He received the following degrees: AB from the University of Alabama, MA from Princeton University, and MS from the University of Denver. He was professor of statistics at the University of Denver from 1925 to 1962, director of the Bureau of Business and Social Research at the University of Denver, and a charter member of the Denver Chapter of the American Statistical Association (ASA), for which he was district secretary; later, he was one of three vice-presidents of the ASA. He was an army officer during World War II. The article referred to in this paper has just two citations in Google Scholar, neither of which is substantive. JSTOR lists six sole author articles by Carmichael between 1916 and 1934 and two jointly authored articles in 1934; there are no listings after 1934.

ISSN 1053-8372 print; ISSN 1469-9656 online/15/04000537-557
doi:10.1017/S1053837215000462

tangent (arctan) trend as an alternative to the linear trend, which was very much the standard of his day, as illustrated by Warren Persons (1916). In addition, Carmichael extended the arctan trend to include smooth transition adjustment to structural breaks, which might otherwise be modelled by discrete changes. Smooth transition is a feature we now recognize in a different form, largely in sigmoid-type adjustment functions, such as the logistic and exponential; see especially Timo Teräsvirta (1994). We note Carmichael's skill in solving, by reasoned argument and a careful graphical analysis of the data, what was then, and still is to some extent, a complex estimation problem.

Carmichael's paper is a prescient taxonomy of the problems of structural breaks, and how to model them, that continue to bedevil the fitting of time series models today. In his opening paragraph he noted that:

> The straight-line tendency commonly experienced prior to the World War was interrupted in many cases in 1915–1917 by a sharp rise, followed in 1920–1921 by an abrupt drop and subsequent resumption of a more regular movement at a level widely different from that of the pre-war period. (Carmichael 1928, p. 253)

These were, as Carmichael noted, "permanent changes in level," and economic time series for which this pattern was apparent included stock prices and the real price of agricultural and livestock products. Whereas the pre-war era was one in which a "straight-line tendency" was a reasonable assumption, thereafter that was, Carmichael argued, no longer the case. A practical question was, therefore, whether, in forming the trend of a series, earlier data should be ignored or whether it could be modelled in a more general and integrated approach, relative to the simple, but usual, straight line assumption. Moreover, in commenting on the impact of the First World War on trend fitting, Carmichael noted:

> Instead of ignoring the war period in the analysis of time series, as a result of trend difficulties, it is possible by their [arc tangent function] use (in many cases) to fit one continuous curve to the entire period of the data which will serve as an adequate measure of the trend. (Carmichael 1928, p. 254)

Carmichael's ideas anticipated two later concerns in econometrics—concerns that were to become of central importance in modern methods. The first is the problem of structural breaks that divide the sample into different regimes; in the quotation above Carmichael had in mind the effects of the First World War, but recent history is replete with similar examples, such as the two Gulf wars, '9/11,' and the credit crisis of 2007–08. Given the occurrence of potentially regime-changing events, how should these be dealt with in econometric modelling? A popular approach is to divide the observations into subsamples by means of discrete dummy variables or, more generally, by an indicator variable, where both imply a sudden switch between the two regimes. Carmichael's suggestion included this possibility, but was not limited by it; rather, he conceived a framework in which the transition phase could be smooth, as expressed in his phrase "one continuous curve."

Carmichael's practical analysis was, however, left with two problems, which he feared might hinder its use; namely, the selection of the break date, or dates, and the "strength" of the adjustment between regimes—the "smoothness" of the adjustment. Although Carmichael offered a graphical method of determining both of these parameters, it was clear that some skill was required in making an appropriate selection.

https://doi.org/10.1017/S1053837215000462 Published online by Cambridge University Press

It is probable that the importance of Carmichael's contributions was passed by partly, if not largely, because of the complex nature of the problem he had posed, even though he provided some informal guidance with which to aid an at least approximate solution.

From a modern perspective, Carmichael sought to solve a problem that involves a process that is changing in some form over time, specifically from one regime to another and so on; in current terminology the process is said to be nonstationary and contrasts with its opposite of a stationary process, where the underlying structure is constant. Thus, in Carmichael's case, the way that the observations are being generated over time is varying, possibly with an indicator of 'regime' change; he introduced a further complexity by specifying a nonlinear process of adjustment between regimes, due to the role of the strength of an adjustment parameter. At the time Carmichael was writing, practical regression models were linear and complexities such as nonlinearities and structural breaks were barely conceived; regression methods remained largely concerned with linear models for such a long period of time that Carmichael's paper was overlooked in the precedence of ideas.

The handmaiden of the practical means of implementation, although subordinate, has to accompany the idea before it can elicit its full impact. In the 1920s, the best means that were available to aid computations were relatively simple electromechanical calculators: see, for example, Arthur Norberg (1990). While, as we show below, Carmichael had the skill to come close to an optimal solution of what is, in effect, a highly nonlinear optimization problem, the lack of a 'ready-made' routine to solve the general problem is likely to have held back the use of his ideas. Even when, much later, the ideas of regime change were again considered, they were initially formulated within a linear (and quite limited) framework (for example, Richard Quandt 1958, 1960), which made their solution computationally feasible. Subsequently, there were developments in a nonlinear framework. See, for example, David Bacon and Donald Watts (1971) on smooth transition between two regimes, Howell Tong (1978, 1982, 1983) on threshold autoregressive models (TARs), and Kung-Sik Chan and Tong (1986), who combined the ideas of Bacon and Watts with the TAR; these developments accompanied improvements in computing power that made feasible the practical solution of such problems.

The remainder of this paper is organized as follows. Section II assesses the background of the trend/cycle decomposition that was prevalent at the time that Carmichael was writing, including some methods to deal with changing trends. Section III outlines the models suggested by Carmichael, with a stylized illustration of one of his models. Section IV considers the relationship of Carmichael's arctan function to other smooth sigmoid-type functions, while section V discusses estimation issues. Section VI re-examines two of Carmichael's original examples using both his and modern methods. Finally, section VII concludes.

II. THE BACKGROUND: DECOMPOSING A TIME SERIES; DEALING WITH STRUCTURAL BREAKS

The concept of a trend is one that has intrigued economists for the last century or so, notwithstanding considerable improvements in econometric and statistical methods.

But exactly what is a trend? Just prior to Carmichael's 1928 article, Karl Karsten wrote: "Trend-lines are intended to describe some imaginary state of 'normalcy' wherewith to compare the actual conditions depicted by the curve" (Karsten 1926, p. 31). Although still imprecise, perhaps deliberately so, this quotation serves to emphasize a trend as an underlying tendency ('normalcy') that is masked by periodic movements such as those due to cycles, seasonal effects, and random irregular events. Recently, Peter Phillips commented that a

> distinguishing characteristic of most economic time series is trending behavior. Such time series often behave either in a wandering manner with long and erratic cycles (as in the case of interest rates and stock prices) or as if they were influenced by some secular drift over time (like many national income components). … In practice, therefore, while many economists see trends in the data, the econometric modeling of such trends is a much more difficult task. (Phillips 2005, p. 402)

Early concerns about the presence of a trend, which was often referred to as being 'secular,' were twofold, broadly arising from (a) modelling and (b) forecasting.[2] In the first case, there was a concern about whether trends should be removed from time series data and, in the second, about the implications of trends for projecting data series forward in time.

Following developments in correlation and regression analysis in the late nineteenth and early twentieth centuries (for example, Karl Pearson 1896, and G. Udny Yule 1897), there was considerable interest in applying these new techniques of analysis to 'uncover' economic relationships.[3] However, there was also a widely expressed concern that short- and long-run movements would be confounded by regressing one possibly trended series on another;[4] see Reginald Hooker (1901) and the later seminal contribution of Yule (1926). Thus, in what soon became a fairly standard procedure, the data were first linearly detrended and the resulting residuals (adjusted if necessary for seasonal movements) were termed 'the cycle,' and it was these cycles rather than the 'raw' series that were the subject of correlation analysis; see, for example, Persons (1916). Wesley Mitchell summarized this approach succinctly, if somewhat harshly: "Secular trends of time series have been computed mainly by men who were concerned to get rid of them" (Mitchell 1927, p. 212).

Mitchell distinguishes the empirical method of determining trends from a more theoretical approach related to the interpretation and implications of particular forms of the trend. On the former, he provides a useful summary of empirical practice at the time as to how the trend may be modelled: fitting a mathematical curve to the series or its logarithm; fitting a moving average (mean or median); fitting a curve to the moving

---

[2]Simon Kuznets (1928, p. 403) suggested the difference as follows: "The purposes for which analysis of time series is undertaken may be divided into two groups. In the first, the interest is in measuring cyclical fluctuations themselves, and the secular movement is determined only to be eliminated and forgotten. In the second, the long-time changes are described in order themselves to be studied and made use of further. To the first group belong all the investigations of correlation; to the second, the studies in forecasting, the comparisons of past changes, the attempts to generalize as to the course of long-time movements."

[3]For more development of this theme, see John Aldrich (1995) and Terence Mills (2011) for a critical overview.

[4]These concerns were to become much more important later, as in Clive Granger and Paul Newbold (1974).

average; drawing a freehand curve; and using ratios between items paired in time where series are considered to have substantially the same secular trend. By contrast, motivating the form of a trend from possible underlying economic causes was relatively undeveloped.[5]

On this view the fitting of a trend is a rather practical matter, only very loosely, if at all, related to any theoretical considerations. This led to some rather ad hoc (but still used) ways of dealing with what we now recognize as nonstationarities in the underlying generating process. The influential work of Persons and his co-authors ('the Harvard group') illustrates the dominant line of thinking that had arisen at the time, especially in the construction of what were known as Business Barometers (for example, Persons 1916).[6] The underlying methodology was to fit a linear trend and measure the cycle as the deviation of the original observation from the trend—a methodology that is not unfamiliar now. The barometer then aggregated a number of component cycles into a single indicator. A key consideration in the often prior graphical examination of the component time series was to establish 'homogenous' periods over which to fit the linear trend. A good example of the issues that were considered is provided by Persons (1919). In examining various time series, ranging from the price of pig iron and general wholesale prices through to immigration and the number of shares sold, Persons observed that the

> study of seventeen series of annual items for the period 1879–1913 led to the conclusion that the secular trend of all the series was quite different during the subperiod 1879–96 from that during the subperiod 1897–1913. Straight lines fit the data very closely for the two subperiods but not for the whole period 1879–1913. … If, however, we had broken the period into two subperiods, the first ending in the 90's and the other beginning there, *two* straight lines would have shown the secular trend of each subperiod accurately. (Persons 1919, p. 38; italics in original)

Segmenting the trends implicitly assumes a discrete break, and hence a 'jump,' between adjacent points in time. For example, Persons (1919, Table H) reports the following

[5]Concerned about the implications for the mechanical projection of empirically determined trends, Mitchell poses the question: "What meanings have the secular trends fitted to time series by empirical methods? … To take the simplest example: a straight line sloping upward implies future increase without limit" (Mitchell 1927, p. 221). He refers approvingly to work by Simon Kuznets, who fitted trends by means of Gompertz and logistic curves, but this work does not appear to have been published; these curves are in the family of sigmoid curves, as is the arctan function chosen by Carmichael. The perceived economic advantages of a logistic type trend (shared by the 'sigmoid' family) are that its limits are finite, the approach to the upper limit is asymptotic and the rate of growth declines, especially as the upper limit is approached. Mitchell, again referring to unpublished work of Kuznets, notes: "These three characteristics, Dr. Kuznets supposes, appear in the history of the many economic processes, whose long-time statistical record is well described by a logistic curve" (Mitchell 1927, p. 225).

[6]Persons's opening comment is of interest: "A barometer showing the fluctuations of business and industrial activity may be put to many uses. Economists and sociologists need such a barometer when dealing with the phenomena of a dynamic society; government officials when handling the problem of unemployment or when considering the advisability of inaugurating large government undertakings; manufacturers and dealers when considering the desirability of making extensions to their plants or of contracting or expanding their purchases, sales or commitments; bankers need a business barometer to guide them in extending or calling their loans and discounts; and investors need one to direct their purchases and sales of securities" (Persons 1916, p. 739). Well-known business barometers of the time included those by Brookmire and Babson.

estimates for two trends of the production of pig iron, here denoted y_t, and regarded as an important leading indicator of economic activity:

1879–1896: $\hat{y}_t = 32.0 + 3.59t$

1897–1913: $\hat{y}_t = 107.0 + 11.9t$

These trend lines indicate substantial, discrete jumps in the intercept and slope of trend production of pig iron, which occur between 1896 and 1897.

Even if this approach is somewhat ad hoc, it predates later and rather more sophisticated models of segmenting trends. For example, it can be considered as part of the general problem of (linear) regressions that may be subject to two regimes, as in Richard Quandt's influential (1958) article, followed by Quandt (1960), who proposed a now widely used approach to test for two regimes against the null hypothesis of a single regime. Subsequently, D. E. Robison (1964) and Derek Hudson (1966) considered the problem of how to join the regression sub-models so as to minimize the discrete nature of the change in regimes.

It is perhaps too harsh to dismiss these early attempts to deal with nonstationarities, arising from structural breaks in the underlying process, as being purely 'empirical.' Modelling change in the form of a discrete jump between regimes is generally 'shorthand' for a more complex adjustment process. Apparently discrete changes, such as the start of the First World War, or more contemporaneously the 'credit crisis,' are often much more detailed than a simple 'off/on switch.' The interesting practical question is whether the data for such a period should be ignored or whether methods can be devised to either approximate a 'local' trend that adjusts to the changing economic circumstances or allows a non-instantaneous transition between regimes.

Some of the empirical approaches referred to by Mitchell can be viewed more positively in this light. Rather than assuming a linear trend subject to a discrete break, a method that adjusts the trend gradually, perhaps giving a greater weight to more recent observations, may be a better way of modelling what appears to be a break in a linear, or indeed a nonlinear, trend. Reginald Hooker (1901) suggested using a (simple) moving average (MA) to represent the trend, from which a cycle would be measured. The moving average was centered on the period t, taking a window of p (usually an odd number) observations to compute a simple arithmetic average, referred to as the 'instantaneous' average. In effect, each observation receives an equal weight in the window, but the window moves through the sample and so a 'local' trend is constructed. The choice of window length was guided by what was thought to be the length of the cycle. The moving average offered some flexibility in modelling the trend, being responsive to the evolution of a time series, and is still a widely applied technique, but it has the disadvantage that, given a sample of length T, the MA trend cannot be computed for the first and last $T - \frac{1}{2}(p-1)$ observations.

An alternative method that similarly constructed a local trend was due to Lincoln Hall (1925), who suggested a method that we now know as a form of recursive estimation, in which a moving (or rolling) window of observations of fixed length is used to estimate a local linear trend; as new observations become available, they replace the observations that were previously at the beginning of the sample. Hall's method of recursive estimation was, however, another idea well ahead of its time and it was not taken up in force until much later in the development of econometric methods. It is likely that the computational demands of Hall's method were too much for the pre-computer age in which his suggestion was made; to some extent he recognized this by offering a slightly simpler computational method in a later paper (see Hall 1926).

The conceptual framework underlying these various methodologies was in effect an embryonic unobserved components (UC) model. This is the decomposition of a series into trend, seasonal, cycle, and irregular components; these components are not directly observed, but constructed from a (particular) modelling approach. Later developments in econometrics saw an increasing sophistication in the nature of the underlying models for each component, but within the same underlying paradigm; see, for example, the developments due to Stephen Beveridge and Charles Nelson (1981), which are based on an underlying random walk framework, and the structural modelling approach due to Andrew Harvey (1985).

In the methodology of the 1900s, the removal of the 'secular' trend was regarded as the first of two stages in analyzing a time series or relationships between time series, where the trend, in keeping with the absence of an accepted framework for a probabilistic interpretation, was generally regarded as deterministic. (This approach was sometimes labelled the "deviations from trend" or "individual trend" method.) Attention could then be focused on the 'cycles' that remained, as these were thought to be of greater interest than the trend itself; see Mary Morgan (1990, chs. 1–3).

However, there were two far-reaching contributions that showed that, in a linear framework, this two-stage procedure was not necessary. The first was due to Bradford Smith (1926), who showed that the time trend could be included in the regression of direct interest; that is, suppose there was interest in the relationship between two variables, say Y and X, each of which was trended; then Smith (1926, equation (8)) showed that the time trend, t, could be included directly in the regression of Y on X and that the resulting coefficients would be simply related to those in the deviation from trend approach. Smith's approach was referred to by Ragnar Frisch and Frederick Waugh (1933) as the "partial trend regression," and they went on to correct Smith in some respects, although they agreed with the fundamental approach: "The partial trend regression method can never, indeed, achieve anything which the individual trend method cannot, because the two methods lead by definition to identically the same results. They differ only in the technique of computation used in order to arrive at the results" (Frisch and Waugh 1933, p. 388). Thus, the circle is 'squared': there is no need to regard trend removal as a prior and separate stage—nothing will be gained by such an approach.

III. CARMICHAELS'S ARCTAN TREND: MULTIPLE REGIMES AND SMOOTH TRANSITION

Neither a rolling window estimation of a linear trend nor a moving average trend would react instantaneously to a change in regime; both would have to 'sense' the change from new observations as they became available to their respective algorithms. Moreover, forecasting with Hall's rolling window trend would not differ in principle from projecting a linear trend; in his case projecting the last local trend that had been estimated. Carmichael's approach offered a change in method that would have implications for forecasting; indeed, one of the explicit rationalizations for his approach was expressed as follows: "While no attempt is made to summarize criteria for curve fitting, it is desired to emphasize the importance of reasonable projection into the

https://doi.org/10.1017/S1053837215000462 Published online by Cambridge University Press

future and the desirability of avoiding gaps in the analysis" (Carmichael 1928, p. 253). Carmichael's concerns find support in a recent assessment of modelling trending behavior: "if the trend mechanism is poorly captured in an empirical model, we can expect forecasts from the model to carry forward the poor approximation" (Phillips 2005, p. 402).

Rather than adopt a methodology that either smoothed the impact of a break, as in Hooker's MA approach or Hall's local trend method, or imposed a jump in the trend or foreshortened a sample, Carmichael suggested modelling the breaks by allowing a smooth transition between multiple regimes. Moreover, in his most general model, he allowed for the possibility of three regimes and so anticipated more recent developments that allow for multiple regime changes, on which there is now an extensive literature; see, for example, Jushan Bai (1997) and Bai and Pierre Perron (1998, 2003). Carmichael thus not only suggested an alternative and inherently nonlinear trend, he also showed how to model multiple structural breaks in the fashion of smooth transition adjustment functions, which have become an important component in modern econometric methods.

Carmichael was not only aware that events, such as the First World War, could lead to structural breaks in the underlying economic processes, but also that to remove such data directly or to model the distinct regimes by means of a simple split trend were not the only, nor necessarily the most desirable, options. Using a split trend in effect involves a step change from one regime to another, whereas the adjustment could be modelled so as to allow a smooth transition between regimes and, moreover, there could be more than one adjustment process in any historical period. While accommodating the impact of the First World War was Carmichael's impetus to nonlinear modelling, there is no shortage of recent events that have been considered as leading to structural breaks—hence the importance of Carmichael's framework.

Against the background of the decomposition of a time series outlined in section II and the predominant practice of linear detrending, Carmichael 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" (Carmichael 1928, p. 253); (ii) approximately linear growth that is resumed after interruption by an abrupt change in level; (iii) as in (ii) but with a first interruption, for example a sharp drop, followed by another abrupt change in level, before resumption of the previous growth.

Carmichael's suggested form, either to amend or supplement the linear trend, was the arctangent function. Three models were distinguished by Carmichael, corresponding to the cases discussed above, as follows:

$$y = a + c\arctan(x) \tag{1}$$

$$y = a + bx + c\arctan(x) \tag{2}$$

$$y = a + bx + c\arctan(x) + d\arctan(\alpha x + \beta) \tag{3}$$

We would now conventionally add time subscripts to the variables y and x and include a random term, ε_t. Carmichael suggested that the variable y may either be the level or

https://doi.org/10.1017/S1053837215000462 Published online by Cambridge University Press

the logarithm of the original data, noting that the difference was between modelling constant absolute changes or constant rates of change (Carmichael 1928, p. 255). The variable x is a 'distance' measure relative to an origin, for which $x = 0$; it is related, but is not necessarily equal, to the simple time trend t that increments by 1 (or some other constant) each time period. Carmichael worked in degrees rather than the now more usual radians. Consider an angle z that varies from –90° to 90°; its tangent then varies from $-\infty$ to $+\infty$. In general, $x = \tan(z)$, and conversely to obtain the angle from the tangent x, $z = \tan^{-1}(x)$, where $\tan^{-1}$ is the inverse function referred to as $\arctan(z)$; z is in degrees and the conversion to radians is $y = z(\pi/180)$. Thus, in radians, the limits of $\arctan(z)$ are $\pm(\pi/2) = \pm 1.5708$ and, by a suitable scaling constant, say c, the limits can be adjusted; for example, $c = \pi^{-1}$ results in limits of $\pm\frac{1}{2}$.

While the estimated trend from the standard linear model is invariant to the scaling of x, this is not the case for the arctan function, so Carmichael also considered how to choose the origin for x and the increment Δx, which varied in his examples from $\Delta x = 0.25$ to $\Delta x = 1$. Carmichael commented that "it is necessary to make [an] arbitrary choice, based upon graphical considerations, of the point of origin and scale of the x axis" (Carmichael 1928, p. 255), and continued in a footnote, "could it [the arbitrary choice] be obviated without loss of flexibility of the curve, the technique would be vastly improved." As we illustrate below, these unknowns can be estimated although, at the time (the 1920s and 1930s), it seems likely that the skill required in making appropriate choices may well have hindered the practical dissemination of Carmichael's ideas.

For our purposes, the arctan function can be reparameterized with $x = \delta(t - t_0)$, where t_0 is the (fixed) origin in time and, therefore, $\Delta x = \delta \Delta t \Rightarrow \Delta x/\Delta t = x/(t - t_0)$ and $\delta = \Delta x$ for $\Delta t = 1$; thus the origin and the increment become explicit parameters, so that the arctan function, as a function of t, becomes $\arctan\left(\delta\left(t - t_0\right)\right)$. This reparameterization allows the interpretation that the coefficient δ governs the speed of transition, whereas the coefficient c governs the overall impact of the change in regime. Also, a simple generalization of Carmichael's specification is to allow the arctan effect to operate on a function of x, $f(x)$, so that, for example, $y = a + cA[f(x)]$, where $A[f(x)] = \arctan\left(f\left(x\right)\right)$.

In a terminology that has now become almost universal, the function $A\left[\cdot\right]$ is referred to as the "transition function," so that it governs the path of adjustment, or transition, between regimes. Moreover, the transition between the regimes is, in Carmichael's model, smooth rather being a discrete step change. Thus, although writers before Carmichael had suggested regime changes by fitting two linear trends, for example to take account of the First World War, their implicit transition function simply switched between 0 and 1, corresponding to regimes 1 and 2, respectively. The arctan function is able to capture this case because, as $\delta \rightarrow \infty$, the transition function becomes a switching function. On the other hand, for 'small' values of δ the transition is smooth and gradual, a feature that Carmichael was evidently aware of when determining empirical values of δ.

Carmichael's typology of models first allows for a single regime governed by a nonlinear trend; the second model then allows for a change in regime, where the change between regimes is smooth; and the third model allows for two changes in regime, and so three regimes overall, where each change between regimes is smooth. It seems reasonable to presume that Carmichael had in mind a model that would allow multiple regimes but that it was sufficient to illustrate the general pattern by his model (3). Carmichael's ideas, therefore, clearly anticipated modern concerns about the existence of multiple regimes in economics, finance, and even (electrical) signal processing; to illustrate the breadth of applications, see, for example, Stephen Gray (1996), Suzanne Cooper (1998), and Matthieu Sanquer et al. (2013), respectively.

It may be helpful to stylize what Carmichael had in mind and, to that effect, we illustrate his model (2), the two-regime model, in figures 1a and 1b. In this case, the model taken to generate the data is (for clarity of illustration no random 'errors' are included):

$$y_t = 100 + 0.1t + 2.0\left(\arctan\left(\delta\left(t-50\right)\right)\right) \qquad t = 1, \ldots, 150.$$

This is a linear trend modified by a nonlinear regime change, with the 'origin' at $t = 50$. The speed of adjustment coefficient δ is varied, starting from slow adjustment, $\delta = 0.1$, to increasingly fast adjustment, with the final $\delta = 1.9$. The successive values of y_t are plotted in Figure 1a and the corresponding changes in y_t, Δy_t, are plotted in Figure 1b.

The variation in δ shows how adjustment can be relatively slow (the shallow adjustment path) or fast (the steep adjustment path). With the corresponding Δy_t small and increasing with δ then in the limit $\delta \rightarrow \infty$, the adjustment becomes a step change.

IV. THE ARCTAN AND OTHER TRANSITION FUNCTIONS

Carmichael wanted a transition function that allowed for smooth nonlinear adjustment, with a parameter that would govern both very fast and very slow adjustment. He chose the arctan function for that purpose; the choice of a trigonometric function seems natural in the context of his likely exposure to such concepts (on the history of trigonometric education in the United States in the relevant period, see Jenna Van Sickle 2011). However, whatever the precise choice of function, it is important to realize that the idea precedes the choice of function, although one can only speculate as to what other functions Carmichael might have considered off the record; we are not given any clues in the text save for the general comment: "The inverse trigonometric function known as the arc tangent appears to be adapted to measuring the trend of series behaving as above described" (Carmichael 1928, p. 253). In a footnote Carmichael commented: "Numerous examples are presented in subsequent paragraphs in which the arc tangent and logarithmic arc tangent curves appear to possess a decided advantage, so far as these requirements are concerned, over the equation types generally employed" (Carmichael 1928, pp. 253–254).

By the nature of his choice, Carmichael is clearly aware of the sigmoid nature of the arctan function. In an influential article of a much later period, David Bacon and Donald Watts mentioned the arctan function as a possible transition function but noted: "There are many transition functions which could be used: for example, the cumulative distribution function of any symmetric probability density function or the hyperbolic tangent" (Bacon and Watts 1971, p. 527). Bacon and Watts used

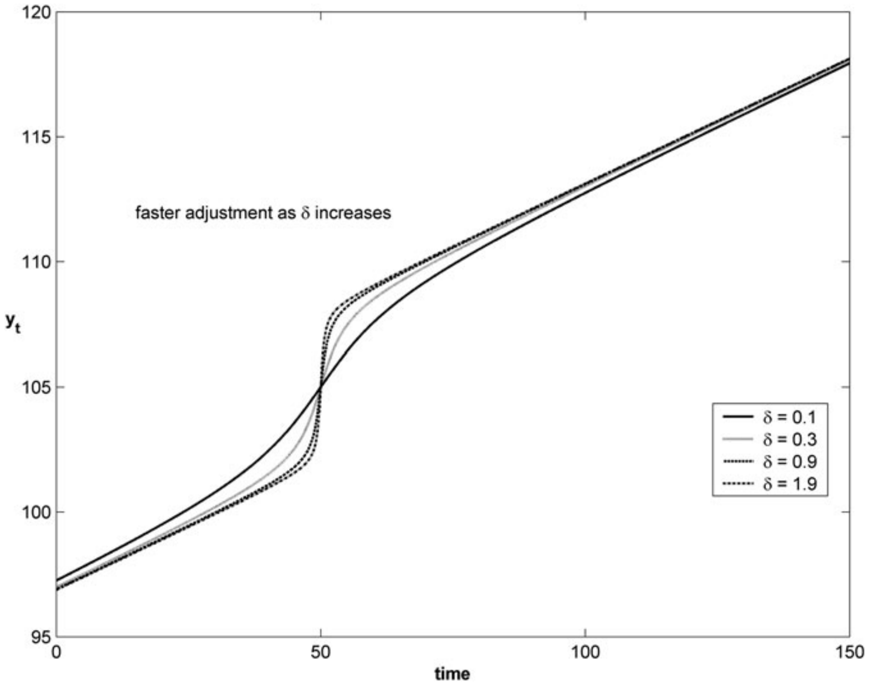

FIGURE 1a. Stylised Arctan transition function.

the hyperbolic tangent for purposes of illustrating the estimation of smooth transition models.[7]

Another possibility, although far less likely to be considered at the time that Carmichael was writing, is the logistic function, which (along with the exponential function) became the dominant choice in the (much) later work of Teräsvirta and his colleagues. Nevertheless, it is worth making a comparison between the arctan and logistic functions, given the present dominance of the latter as a transition function. The arctan function is close under some choice of parameters to the logistic function, $L(x) = \gamma/(1+\exp(-\psi x))$, where $L(x) \to \gamma$ as $x \to \infty$, for $\psi > 0$. The logistic function was favored at the time in population studies: see, for example, Raymond Pearl and Lowell Reed (1923). It was originally developed by Pierre-François Verhulst (1845) as an alternative to modelling the growth of a population as exponential, which was unsatisfactory because it was without a limit; in contrast the rate of growth of the logistic was not constant, but depended on how close the population was to its limiting value. Jan Cramer (2002) notes that the logistic curve was discovered 'anew' by Pearl and Reed.

The derivatives of the arctan function are well known: for example, the first derivative with respect to x is

[7]In the context of smooth TARs, Eleftherios Giovanis (2008) considers a number of alternative transition functions for the sigmoid family.

https://doi.org/10.1017/S1053837215000462 Published online by Cambridge University Press

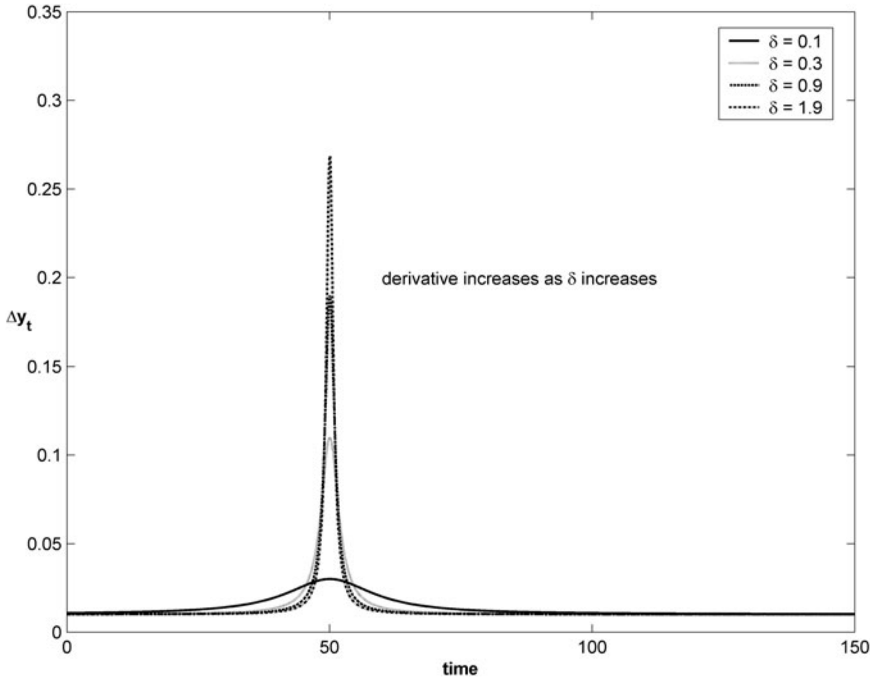

FIGURE 1b. Corresponding transition changes.

$$\frac{\partial A[.]}{\partial x} = \frac{\partial f(x)/\partial x}{(1 + f(x)^2)} \qquad \text{where } y = a + cA[f(x)]$$

$$\frac{\partial A[.]}{\partial x} = \frac{1}{(1 + x^2)} \qquad \text{for } f(x) = x.$$

The derivative of $A[\cdot]$ with respect to time (rather than the distance measure x) is

$$\frac{\partial A[.]}{\partial t} = \frac{(\partial f(x)/\partial x)(\partial x/\partial t)}{(1 + f(t)^2)}$$

$$\frac{\partial A[.]}{\partial t} = \frac{\delta}{(1 + f(t)^2)} \quad \text{for } f(t) = \delta(t - t_0).$$

A simple arctan function is illustrated in Figure 2 with its derivative shown in the lower figure; in both cases these are compared with the centred logistic function $L(x) = 1/[1 + \exp(-\psi x)] - 0.5]$, so that $L(x) \to \pm 1/2$ for $x \to \pm\infty$ (for comparability, the arctan function is in units of radians and the calibration is $\psi = -4c$, $c = \pi^{-1}$). Although $x = 0$ is referred to as the "origin," being the point at which t_0 is the origin on the time axis, it is perhaps better described as a switch point, or point of symmetry, where the sign of the first derivative with respect to x (and t) changes from positive to negative: see the lower panel in Figure 2.

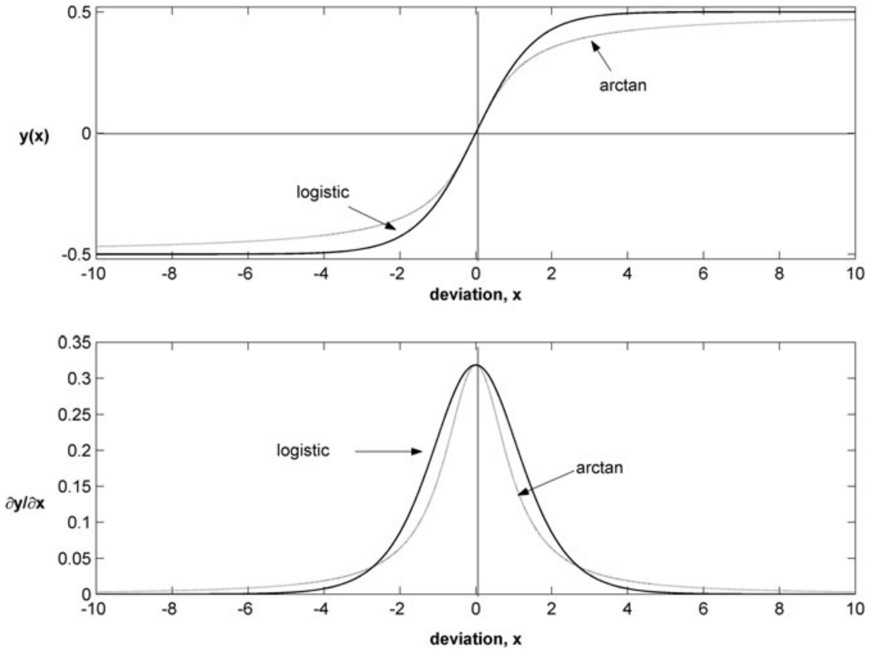

FIGURE 2. Arctan and logistic trends/transition functions.

A relatively recent but important development in time series modelling is the TAR model of Tong (1977, 1978, 1983). The basis of the TAR is the autoregressive (AR) model due to Yule (1927), in which the variable of interest, y_t, is regressed on, say, p lags of itself. A TAR model extends the AR model to allow for more than one regime or state. In the simplest extension the models are distinguished by state. For example, if, for some choice of the lag parameter d, $y_{t-d} < c$, then the model is generated in state 1; whereas if $y_{t-d} \geq c$, the model is generated in state 2. The models are distinguished by time if c is replaced by a time index, such that regimes 1 and 2 correspond to $yt - d < t_b$ and $yt - d \geq t_b$, respectively. The general idea is that there is a threshold, respectively c or t_b, the crossing of which triggers transition between the states/regimes.

The further development of the TAR to have smooth transition, as in Bacon and Watts (1971), is due to Chan and Tong (1986), who introduced the acronym STAR, the S denoting 'smooth'; their choice of transition function was the cumulative distribution function (cdf) of the normal, which, as the normal distribution is symmetric, gives the classic sigmoid shape.[8] Ritva Lukkonen et al. (1988) noted that using the logistic function, which has analytical derivatives, for the transition function was computationally easier than using the cdf of the normal and, moreover, that the logistic was a good approximation to the normal cdf. In subsequent applications of STAR models, the choice of the logistic function, leading to the acronym LSTAR (with the T sometimes

[8]The sigmoid shape is like the letter S, although in most cases a rather elongated S, which is stretched and pulled to the right. Sometimes the sigmoid shape is taken to refer to the logistic function, although the term is more general than relating to a single function.

now taken to refer to transition rather than threshold), has dominated the literature; for a review of the many variants of the TAR model, see Bruce Hansen (2007).

V. ESTIMATION

Carmichael's examples were estimated using graphical inspection for the origin of the switch point and for two or three likely values of Δx; however, the model is linear given $\delta = \Delta x$ and conditionally on t_0, so we first define two sets each containing the respective true values, $\delta \in \boldsymbol{B}$ and $t_0 \in \boldsymbol{T}$, and minimize the residual sum of squares $\sum_{t=1}^{T} \hat{\varepsilon}_t^2$ over $\boldsymbol{B}$ and $\boldsymbol{T}$: that is $\min_{t_0}(\min_{\delta} \sum_{t=1}^{T} \hat{\varepsilon}_t^2)$ over the two-dimensional grid formed by the Cartesian product $\boldsymbol{B} \times \boldsymbol{T}$. The resulting estimators are consistent under fairly standard conditions. (The procedure is an application of the principle in Dag Tjostheim [1986], which has been used in various structural break papers that consider all possible break points in a particular sample; see, for example, Donald Andrews [1993], and Hansen [2000].) Note that t_0 is not constrained to lie between 1 and T, and in one of Carmichael's examples the origin is taken to pre-date the beginning of the sample, in which case t_0 would be negative.

This method can be extended to a double break and, hence, to two transitions as in Carmichael's model (3). In this extension it is convenient to define $\delta_1 \equiv \Delta x_1$ and $\delta_2 \equiv \Delta x_2$, and the corresponding sets $\delta_1 \in \boldsymbol{B}_1$ and $\delta_2 \in \boldsymbol{B}_2$; similarly define two origins and corresponding sets for the centred break points, $t_{0,1} \in \boldsymbol{T}_1$ and $t_{0.2} \in \boldsymbol{T}_2$. The least squares solution then minimizes over $\boldsymbol{B}_1 \times \boldsymbol{B}_2 \times \boldsymbol{T}_1 \times \boldsymbol{T}_2$.

While this describes the modern approach, how did Carmichael solve the problem of obtaining estimates of the parameters? From his perspective he was faced with a problem without a practical solution unless it could be simplified. The simplifications came in the form of conditioning the least squares estimates on (a) the origin and (b) the speed of adjustment. In Carmichael's own words:

> When fitting curves of types (1), (2) and (3) by the method of least squares it is necessary to make arbitrary choices, based upon graphical considerations, of the point of origin and of the scale on the *x*-axis. If a downward tendency approximating the straight-line throughout the period is indicated, the origin is chosen at a point several years prior to the beginning of the data. In the other cases referred to, the origin is taken near a point of rapid change in level. (Carmichael 1928, p. 255)

Once these choices are made, and together with some consideration of whether the data is best modelled in terms of its original form or its logarithms, the estimation becomes simply that of a linear least squares model. What is evident from Carmichael's careful prior analysis is the importance of understanding the nuances of movements in the data based on a careful graphical analysis.[9] This supports the views expressed in Jeff Biddle (1999) and Mary Morgan (1990) on the use and importance of graphs in the statistical analysis of the period. Moreover, while

[9]To aid the reader, Carmichael notes: "Determination of the yearly change in *x* in the case of the logarithmic arc tangent curve may be made similarly by use of a graph of the logarithms of the actual data on simple arithmetic paper, or of the actual data on semi-logarithmic paper" (Carmichael 1928, p. 255).

Carmichael does not report formal measures of the goodness of fit of his proposed trends, he is aware of the need to 'fit the data well' in making his choice of functional form and conditioning parameters.

VI. EXAMPLES

We consider two of Carmichael's examples[10] to illustrate the application of the arctan function and how the selection of the choice parameters can be included in the process. In so doing we are able to see how well Carmichael did from a modern perspective.

Carmichael's first example concerns modelling the trend in the common stock price of International Paper, which poses a challenge because neither a linear nor an exponential trend would seem to capture the key trend features of the data, which embody a gradual, but sustained, increase about halfway through the sample period of 1900 to 1926. The data, which are annual and obtained as the yearly average of 'tenth-of-month' prices ($ per share), are plotted in Figure 3. Carmichael used logs of the data and obtained his results for choices of $\Delta x = 0.75$ and 0.6; he chose the origin of the switch point by a visual examination of a graph of the data and, by this method, chose the origin to be 25 December 1917 (with 25 June 1917 as a possible alternative). Estimation over the original sample period gives results very close to those reported by Carmichael; see Table 1, rows (1) and (2). The 'original' arctan trend is also plotted in Figure 3. To illustrate the difference with the then standard of the day, a linear (in logarithms) trend was also estimated and is also shown on Figure 3. Evidently the implied cycles (residuals) and even fairly short-term projections differ quite markedly, depending on the choice of trend.

Searching over $\delta = \Delta x$ and the origin t_0 locates the minimum at $\Delta x = 0.71$ and the switch point at t_0 = 25 December 1916, which is twelve months earlier than Carmichael's choice; see row (3) of Table 1. This earlier choice is evident in Figure 3, which also shows the revised arctan trend, and picks up the increase in the series noticeably earlier than the 'original' arctan trend.

How well did Carmichael do in his original choices? In broad terms, Carmichael came fairly close to the best fitting trend by using his graphical methods. Figure 4 shows the residual sums of squares as a function of the break date and Δx. Note that the objective function has a distinct valley approaching 1916/1917 (the projection from left to right on the figure), although a choice of 1917/1918 is not far away, and there is also a local minimum much earlier in the sample around 1908, which Carmichael avoids. The objective function is much flatter for the choice of Δx (see the projection from right to left), so selecting $\Delta x = 0.75$ is relatively close to the minimum that occurs at $\Delta x = 0.71$; nevertheless the residual sum of squares can be reduced by approximately 25% by moving to the minimum of the RSS function.

The second example illustrates Carmichael's case (3), where a first 'abrupt' change is followed by a further change. The series is the price of Central Leather common stock; data are annual from 1900 to 1926, centred on 25 June of each year, and are shown as

[10]Carmichael also illustrates his arctan model with data on the price of American Car and Foundry common stock and the US Index of Wholesale prices. The examples are chosen by Carmichael as they show evidence of rapid, large, and possibly multiple, changes that would not be well fitted by straight-line trends.

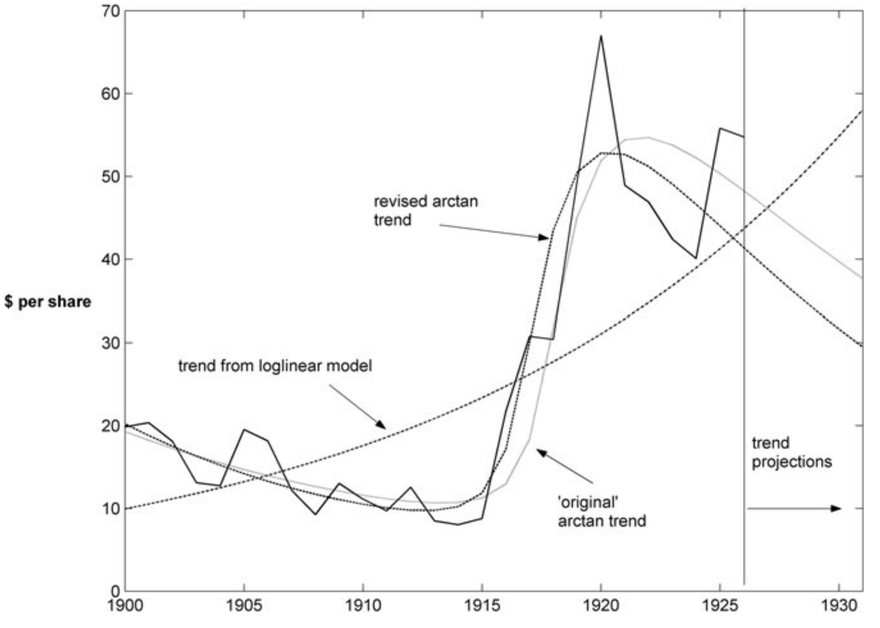

FIGURE 3. Internationl Paper Stock Prices.

Table 1. International Paper common stock: single break, arctan estimation

Dep. var: log(y)	constant	X	arctan(x)	RSS	Δx	origin, $x = 0$, t_0
	Carmichael's original estimates					
(1)	1.38295	–0.034338	0.0064119	n.a	0.75	25th Dec, 1917
	Re-estimation					
(2)	1.38274	–0.034328	0.0064082	0.2640	0.75	25th Dec, 1917
('t')	(58.02)	(–4.369)	(9.534)			
	Global minimum: Estimation over ***B*** and ***T***					
(3)	1.35593	–0.047389	0.0071171	0.1954	0.71	25th Dec, 1916
('t')	(69.46)	(–6.050)	(11.455)			

Notes: logs are to the base 10; the number of decimal places reported follows Carmichael; RSS = residual sum of squares; n.a indicates not available.

the unbroken line in Figure 5. In present terminology, the model embodies a double structural break, with smooth transition, the first captured by the arctan function arctan(x), with origin at $x = 0$, and the second by the arctan function arctan($\alpha x + \beta$), with origin at $\alpha x + \beta = 0$. Carmichael again selected the break points by graphical methods, which are worth briefly recounting: the point of most rapid rise was identified as 25 December 1915, and that of most rapid decline as 25 June 1920, a difference in time of 4.5 years; the origin of x was, therefore, taken to be the first of these dates, and α and β were then determined such that $\alpha x + \beta = 0$.

The parameter α governs the 'strength' of the second adjustment function relative to the first, which Carmichael judged to be equal, and on setting $\alpha = 1$ this implies that $\beta = -4.5$.

https://doi.org/10.1017/S1053837215000462 Published online by Cambridge University Press

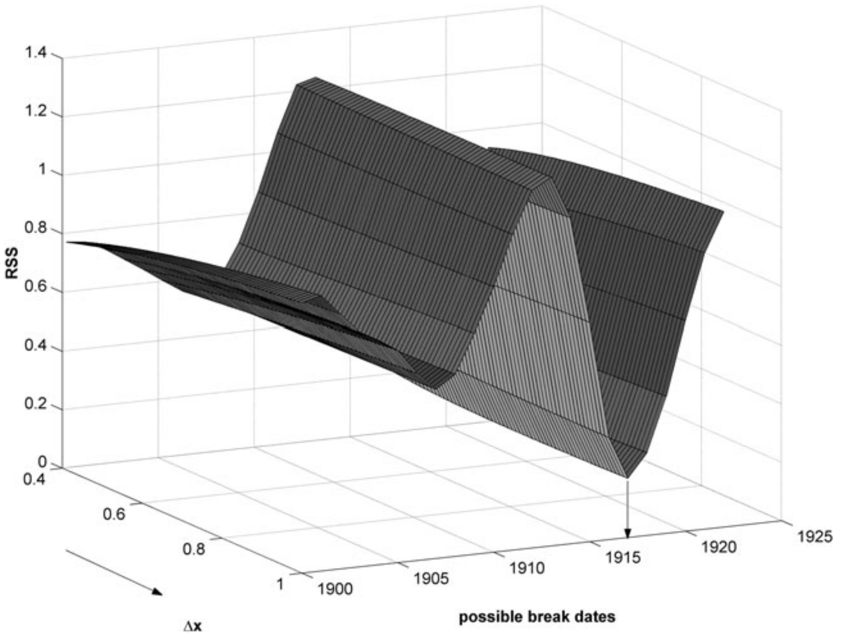

FIGURE 4. Minimising over the break and ΔX.

The double-switch model was estimated using Carmichael's parameters and also by extending the search procedure over α and the second switch point (which implicitly determines β). The results are reported in Table 2. Row (1) constrains the choice parameters as in Carmichael (his reported results are confirmed; see footnote to table). The restrictions are then relaxed in stages: i) keep $\Delta x = 1$, but search for the best double switch points, which results in row (2); ii) constrain the Δx to be equal but not equal to 1, which results in $\Delta x = 0.86$ and $\hat{\beta} = -4.5$; the estimated switch dates are, however, each one year later than Carmichael's estimates, although the difference between them is the same. Finally, iii) unrestricted estimation, that is, relax the restriction of equality of the Δx in the two arctan functions: see row (3), where Δx_1 and Δx_2 are now distinguished.

The results show that Carmichael was again 'out' by one year in his timing of the breaks, and relaxing the assumption that $\Delta x = 1$, and then distinguishing Δx_1 and Δx_2, shows that Carmichael overestimated Δx for the first break point ($\Delta x_1 = 0.81$), but was very close in his choice for the second break point ($\Delta x_2 = 1.03$). However, given the four-dimensional search undertaken across two possible switch points and the values of Δx_1 and Δx_2, Carmichael clearly showed great skill in coming as close as he did to the global minimum in the RSS function. The original and revised trends, including the simple (log)linear trend, are also shown in Figure 5.

VII. CONCLUDING REMARKS

Carmichael's opening sentence was prescient of present times: "During the past twelve years many economic series have undergone what appears to be a permanent change

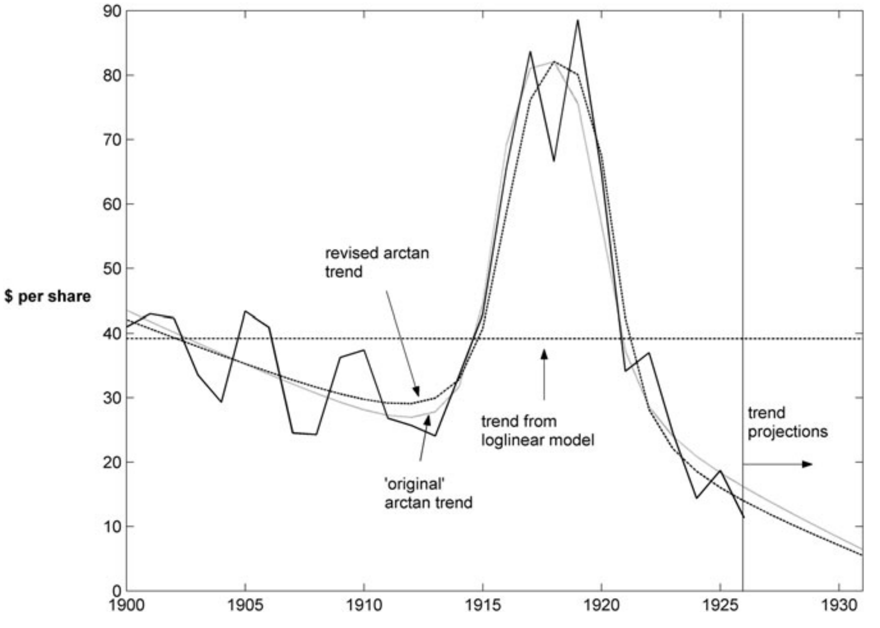

FIGURE 5. Central Leather Stock Prices.

in level" (Carmichael 1928, p. 253). As a result he recognized that fitting a simple linear trend to economic time series would often be inadequate and, instead, proposed a simple but effective method of modelling changes by means of smooth transition functions. While this is the language of the present, rather than of the 1920s, Carmichael was clear in his intentions and modelling strategy, of which there were three key principles: to allow continuity in the adjustment rather than a step jump from one regime to another; to use a sigmoid-type adjustment function in which the rate of adjustment varied depending on the stage in the path of adjustment; and to allow for multiple regimes, the number depending on the particular economic circumstances of the case at hand.

In proposing such a strategy Carmichael made a major, but previously unrecognized, contribution to time series analysis. The idea that economic processes may undergo shocks that result in permanent changes was taken up much later, especially in concerns about the implications for testing for unit roots; for example, Perron (1989, 1992). The idea of smooth transition between regimes is now a standard part of the modelling 'toolkit'; and the potential presence of multiple 'structural breaks' is a topic that has great relevance in a world of seemingly endless economic and political upheaval.

Moreover, although the technical means of solving the problems he had posed were many years away from Carmichael, his skillful reasoning and understanding of the issues involved in estimating the break date(s) and strength of adjustment of his arctan adjustment mechanism(s) were quite exceptional. A combination of graphical methods (which, as Biddle [1999] notes, were an important part of the methodology of the era), and analysis of the changes in the data and economic circumstances, led him close to the solutions he would have obtained had he the benefit of modern methods.

Table 2. Central Leather common stock prices: double break, arctan estimation

Dep. var: y	constant	x	arctan(x)	arctan($x-\beta$)	RSS	Δx	origin, $x=0$
				$\beta = -4.5$			
(1)	24.47	–1.7616	0.5191	–0.4193	1105.26	1.0	Break 1: 25th Dec, 1915
't'	(13.26)	(–3.958)	(11.156)	(–10.684)			Break 2: 25th June, 1920
				$\hat{\beta} = -4.5$			
(2)	21.84	–1.615	0.5230	–0.4910	1084.22	0.86	Break 1: 25th June, 1916
''t'	(11.26)	(–3.21)	(10.73)	(–11.84)			Break 2: 25th Dec, 1921
	Global minimum: estimation over $\boldsymbol{B}_1 \times \boldsymbol{B}_2 \times \boldsymbol{T}_1 \times \boldsymbol{T}_2$						
				$\hat{\beta} = -4.5$			
(3)	22.53	–1.845	0.5248	–0.4648	1078.08	$\Delta x_1 = 0.81$	Break 1: 25th June, 1916
't'	(11.88)	(–3.44)	(10.62)	(–11.84)		$\Delta x_2 = 1.03$	Break 2: 25th Dec, 1921

Notes: Carmichael's original estimates, $y = 24.47 - 1.7613\,\text{arctan}(x) - 0.4193\,\text{arctan}\left(x - 4.5\right)$

However, the inherently nonlinear nature of Carmichael's models was probably a step too far at the time and for some time thereafter. The linear regression model was then, and for much later, the dominant paradigm for analysis. It was not until the 1970s, and even then somewhat hesitantly, that the seeds of nonlinear analysis began to come to fruition; see Tong's (2010) interesting retrospective. The parallel of such developments with those in computing, especially 'micro' or personal computing, seems relevant and related, but is left for others to analyze.

REFERENCES

Aldrich, John. 1995. "Correlations Genuine and Spurious in Pearson and Yule." *Statistical Science* 10 (4): 364–376.

Andrews, Donald. 1993. "Tests for Parameter Instability and Structural Change with Unknown Change Point." *Econometrica* 61 (4): 821–856.

Bacon, David, and Donald Watts. 1971. "Estimating the Transition between Two Intersecting Lines." *Biometrika* 58 (3): 525–534.

Bai, Jushan. 1997. "Estimating Multiple Breaks One at a Time." *Econometric Theory* 13 (3): 315–352.

Bai, Jushan, and Pierre Perron. 1998. "Estimating and Testing Linear Models with Multiple Structural Changes." *Econometrica* 66 (1): 47–78.

———. 2003. "Computation and Analysis of Multiple Structural Change Models." *Journal of Applied Econometrics* 18 (1): 1–22.

Beveridge, Stephen, and Charles R. Nelson. 1981. "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle." *Journal of Monetary Economics* 7 (2): 151–174.

https://doi.org/10.1017/S1053837215000462 Published online by Cambridge University Press

Biddle, Jeff. 1999. "Statistical Economics, 1900–1950." *History of Political Economy* 31 (4): 607–651.

Carmichael, Fitzhugh. 1928. "The Arc Tangent in Trend Determination." *Journal of the American Statistical Association* 23 (163): 253–262.

Chan, Kung-Sik, and Howell Tong. 1986. "On Estimating Thresholds in Autoregressive Models." *Journal of Time Series Analysis* 7 (3): 179–190.

Cooper, Suzanne. 1998. "Multiple Regimes in U.S. Output Fluctuations." *Journal of Business and Economic Statistics* 16 (1): 92–100.

Cramer, Jan. 2002. "The Origins of the Logistic Regression." Tinbergen Institute Discussion Paper, TI 2002 119/4.

Frisch, Ragnar, and Frederick V. Waugh. 1933. "Partial Time Regressions as Compared with Individual Trends." Econometrica 1 (4): 387–401.

Giovanis, Eleftherios. 2008. "Additional Smoothing Transition Autoregressive Models." MPRA Paper No. 24657.

Granger, Clive, and Paul Newbold. 1974. "Spurious Regressions in Econometrics." *Journal of Econometrics* 2 (2): 111–120.

Gray, Stephen. 1996. "Modeling the Conditional Distribution of Interest Rates as Regime-Switching Process." *Journal of Financial Economics* 42 (1): 27–62.

Hall, Lincoln. 1925. "A Moving Secular Trend and Moving Integration." *Journal of the American Statistical Association* 20 (149): 13–24.

———. 1926. "The Determination of Past and Present Secular Trends." *Journal of the American Statistical Association* 21 (154): 206–212.

Hansen, Bruce. 2000. "Testing for Structural Change in Conditional Models." *Journal of Econometrics* 97 (1): 93–115.

———. 2011. "Threshold Autoregression in Economics." *Statistics and Its Interface* 4 (2): 123–127.

Harvey, Andrew C. 1985. "Trends and Cycles in Macroeconomic Time Series." *Journal of Business and Economic Statistics* 3 (3): 216–227.

Hooker, Reginald. 1901. "Correlation of the Marriage-rate with Trade." *Journal of the Royal Statistical Society* 64 (3): 485–492.

Hudson, Derek. 1966. "Fitting Segmented Curves whose Join Points have to be Estimated." *Journal of the American Statistical Association* 61 (316): 1097–1129.

Karsten, Karl. 1926. "The Harvard Business Indexes—A New Interpretation." *Journal of the American Statistical Association* 21 (156): 399–418.

Kuznets, Simon. 1928. "On the Analysis of Time Series." *Journal of the American Statistical Association* 23 (165): 398–410.

Lukkonen, Ritva, Pentti Saikkonen, and Timo Teräsvirta. 1988. "Testing Linearity Against Smooth Transition Autoregressive Models." *Biometrika* 75 (3): 491–499.

Mills, Terence. 2011. *The Foundations of Modern Time Series Analysis*. Basingstoke: Palgrave-Macmillan.

Mitchell, Wesley. 1927. *Business Cycles: The Problem and its Setting*. New York: NBER.

Morgan, Mary. 1990. *The History of Econometric Ideas*. Cambridge: Cambridge University Press.

Norberg, Arthur. 1990. "High-Technology Calculation in the Early 20th Century: Punched Card Machinery in Business and Government." *Technology and Culture* 31 (4): 753–779.

Pearl, Raymond, and Lowell Reed. 1923. "On the Mathematical Theory of Population Growth." *Metron* 3 (1): 6–19.

Pearson, Karl. 1896. "Mathematical Contributions to the Theory of Evolution, III: Regression, Heredity and Panmixia." *Philosophical Transactions of the Royal Society of London*, Series A (187): 253–318.

Perron, Pierre. 1989. "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis." *Econometrica* 57 (6): 1361–1401.

———. 1992. "Testing for a Unit Root in a Time Series with a Changing Mean." *Journal of Business and Economic Statistics* 8 (2): 153–162.

Persons, Warren. 1916. "Construction of a Business Barometer Based upon Annual Data." *American Economic Review* 6 (4): 739–769.

———. 1919. "General Considerations Relating to Secular Trend, 1903–18." *The Review of Economics and Statistics* 1 (1): 38–39.

Phillips, Peter. 2005. "Challenges of Trending Time Series Econometrics." *Mathematics and Computer Simulation* 68 (5–6): 401–416.

Quandt, Richard. 1958. "The Estimation of the Parameters of a Linear Regression Obeying Two Separate Regimes." *Journal of the American Statistical Association* 53 (284): 873–880.

———. 1960. "Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes." *Journal of the American Statistical Association* 55 (290): 324–330.

Robison, D. E. 1964. "Estimates for the Points of Intersection of Two Polynomial Regressions." *Journal of the American Statistical Association* 59 (305): 214–224.

Sanquer, Matthieu, Florent Chatelain, Mabrouka El-Guedri, and Nadine Matin. 2013. "A Smooth Transition Model for Multiple-regime Time Series." *IEEE Transactions on Signal Processing* 61 (7): 1835–1847.

Smith, Bradford B. 1926. "Combining the Advanatges of First-Difference and Deviation-from-Trend Methods of Correlating Time Series." *Journal of the American Statistical Association* 21 (153): 55–59.

Teräsvirta, Timo. 1994. "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models." *Journal of the American Statistical Association* 89 (425): 208–218.

Tjostheim, Dag. 1986. "Estimation in Nonlinear Time Series Models." *Stochastic Processes and their Applications* 21 (2): 251–273.

Tong, Howell. 1978. "On a Threshold Model." In Chi Hua Chen, ed., *Pattern Recognition and Signal Processing*. Amsterdam: Sijthoff and Noordhoff, pp. 575–586.

———. 1982. "Discontinuous Decision Processes and Threshold Autoregressive Time Series Modelling." *Biometrika* 69 (1): 274–276.

———. 1983. *Threshold Models in Non-linear Time Series Analysis.* New York: Springer.

———. 2010. "Threshold Models in Time Series Analysis—30 Years On." Research Report No 471. Pokfulam: University of Hong Kong.

Van Sickle, Jenna. 2011. "A History of Trigonometric Education in the United States: 1776–1900." PhD thesis, Columbia University.

Verhulst, Pierre-François. 1845. "Recherches mathématiques sur la loi d'accroissement de la population." *Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles* 18: 1–42.

Yule, G. Udny. 1897. "On the Theory of Correlation." *Journal of the Royal Statistical Society* 60 (4): 477–489.

———. 1926. "Why Do We Sometimes Get Nonsense-correlations between Time-series? A Study in Sampling and the Nature of Time Series." *Journal of the Royal Statistical Society* 89 (1): 1–63.

———. 1927. "On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers." *Philosophical Transactions of the Royal Society of London.* Series A (226): 267–298.