Using turning point information to study economic dynamics*

Don Harding†
The University of Melbourne

November 27, 2003

Abstract

In this paper I use the proportion of turning points located in the sample path to of time series data to describe that data. I show that the proportion of turning points can be directly related to the data generating process and therefore provides a methodology for estimation and testing of models. This methodology has two main advantages. First, it has direct links back to the visual intuition that practitioners obtain from the inspection of graphs. Second, it turns out to be very robust as the methodology developed here does not require that any of the moments of the series Y_t exist.

---

*This paper is revised version of Chapter 3 of my 2003 Yale Phd dissertation. JEL Classification: C12, C14, C22, E32. Keywords: Business cycle, turning points, elliptically symmetric distribution.

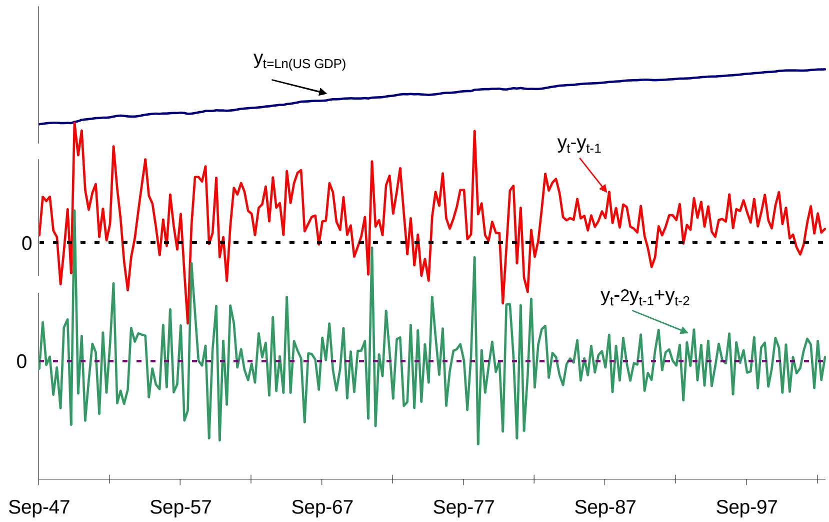
†I would like to thank Adrian Pagan for comments on an earlier draft of this paper. Of course, I am responsible for any errors.

When we behold one of the familiar graphs of economic time series .... we undoubtedly have, ... , the impression of an irregular regularity of fluctuations. Our first and foremost task is to measure them and to describe their mechanism. Schumpeter (1939)

1 Introduction

Many of the stylized facts and much of the intuition that economists develop about macroeconomic phenomena such as business cycles comes from the inspection of graphs of time series; Jugular's (1862) table-graph of credit cycles and Jevons (1862) graphs in his *Study of Periodic Commercial Fluctuations* provide early examples of such visual intuition. Today, it remains natural to start an empirical economic inquiry by viewing the graph of some economic time series such as that of ln(US GDP) shown in Figure 1.

Figure 1: Time series of Ln(US GDP), its first and second differences



When one looks at Figure 1 the human eye is drawn to certain features of the sample path. The first of these is the pattern of rises and falls in the level of the series — the absence of any such patterns would mark the series as unusual in economics and in the applied sciences. If a time series is generally moving upwards or downwards then, the eye notices periods in which the trend is temporarily reversed after that it might notice whether the growth rate is above or below trend and then perhaps the eye might locate slowdowns and pick ups in the growth rate and changes in volatility such as that which is evident after the mid 1980s in US GDP.

Typically, the eye cannot discern all of this information from a graph of the original series, particularly if it has a pronounced trend. But, as is illustrated in Figure 1, a viewing of graphs of higher differences of the series reveals some further information — for example, periods when $\Delta y_t < 0$ are periods in which the trend is temporarily reversed while those where $\Delta^2 y_t < 0$ are periods where the growth rate is declining. Such visual features of the data are often mentioned in discussion of time series but it is rare for practitioners to connect the visual features seen in the data with other summaries of the data such as the sample moments or estimates of the parameters of the data generating process (DGP). The result is often a feeling of some dissatisfaction regarding the connection made between the data description and the more formal modelling and testing parts of applied papers.

One source of this feeling of dissatisfaction is the fact that the human brain is highly adapted for the processing of visual intuition. Oomes (1980, Chapter 2 p. 12), for example, estimates that more than 50 per cent of the human brain is occupied with optically related processing. In recognition of the prodigious visual processing power of the human mind there has been an increase in interest in the visual analysis of time series data in computer science and statistics; see for example Berndt and Clifford (1996), Card, Mackinlay and Shneiderman (1999), Carlis and Konstan (1998), Silva and Catarci (2000). Motivated by that literature, this paper is based on the view that practitioners might find use for an econometric technology that develops the connection between the patterns that the eye sees and the DGP. My objective in this paper is to develop that econometric technology and to apply it to some interesting questions related to the investigation of economic time series.

Perhaps, the most well known of these visual methods relates to Burns and Mitchell's procedures for locating turning points in the sample path of $ln(Y_t)$. Section 2 explores how to generalize this method to investigate turning points in first and higher differences of the logarithm of the series. Harding (2003) shows that these turning points in higher order differences in $log(Y_t)$ are based on successively weaker notions of what constitutes decisive change.

The discussion above of the patterns that the eye discerns focused on the peaks and troughs in either the levels, deviation from trend or difference of the series. One could, with equal justification, focus on the patterns formed when the series crossed the trend, or when the differences of the series crossed zero. It turns out that there is a close relationship between such crossing points and turning points and there is a mathematics literature on the former. These relationships are explored in section 3.

The focus of interest in this paper is in using counts and proportions of turning points to construct statistics that can be used to study economic time series. Section 4 explores the main properties of counts and proportions of turning points. The latter always exists and is bounded above by one and below by zero. This means that we can always compare two series by comparing the proportion of turning points, something that one cannot always do with moments since some or all of the population moments of the series may not be finite. Section 4 shows that the proportion of turning points has a mononicity property that can be used to formalize the visual intuition that a decreasing amount of information about a dynamic process is obtained as one successively differences the series. Asymptotic theory and procedures to obtain robust standard errors for the proportion of turning points are developed thereby providing the foundation for hypythesis testing.

Constructing visually based summaries of economic time series is useful, but in the econometric literature interest focuses on linking the statistics used to describe the features of the series to the DGP. Section 5 discusses the relationship between counts of the number of turning points and the data generating process for y_t. Analytic results are presented for the case where the DGP comes from the class of elliptically symmetric distributions — simulation methods can be used if the DGP falls outside of this class of distributions.[1]

Applications of the methodology are given in section 6. Conclusions are presented in section 7.

2 A formal definition of calculus rule turning points

Using the notation that $\wedge_t$ and $\vee_t$ represent a peak at t and trough at t respectively the procedure that Harding and Pagan (2001) labeled as the

[1] Appendix A sets out the relevant properties of the class of elliptically symmetric distributions.

calculus rule for identifying turning points in $\Delta^r y_t$, is written formally as follows,[2]

$$\wedge_t^r = \mathbf{1}\left(\Delta^r y_t > Max\{\Delta^r y_{t-1}, \Delta^r y_{t+1}\}\right) \tag{1}$$

$$\vee_t^r = \mathbf{1}\left(\Delta^r y_t < Min\{\Delta^r y_{t-1}, \Delta^r y_{t+1}\}\right) \tag{2}$$

By varying r one can implement various definitions of decisive change. When $r = 0$ in equations (1) and (2) the notion of decisive change being used is that the series changes direction so that turning points are local maxima or minima in the level of the series. Such turning points are said to be from the **classical cycle** that was studied by Burns and Mitchell.[3] The last date before y_t shows a sustained decline (rise) is referred to as the **classical cycle peak (trough).**

It is useful to note that equations (1) and (2) can also be expressed as,

$$\wedge_t^r = \mathbf{1}\left(0 > Max\{-\Delta^{r+1} y_t, \Delta^{r+1} y_{t+1}\}\right) \tag{3}$$

$$\vee_t^r = \mathbf{1}\left(0 < Min\{-\Delta^{r+1} y_t, \Delta^{r+1} y_{t+1}\}\right) \tag{4}$$

Thus, all of the information that is needed to study turning points in $\Delta^r y_t$ resides in the DGP of $\Delta^{r+1} y_t$.

2.1 Higher order turning points

Higher order turning points involve successively weaker notions of decisive change and are associated with values of $r \geq 1$ in equations (1) and (2). The case where $r = 1$ receives some popular attention. It is worth giving the cycle defined in this way the special name of the acceleration cycle because

---

[2] I use the following notation.

Δ notation. For a positive integer r, $\Delta^r y_t = \Delta^{r-1} y_t - \Delta^{r-1} y_{t-1}$ and $\Delta^0 y_t = y_t$. For a negative integer the cumulation of y_t is calculated, viz $\Delta^{-1} y_t = \sum_{i=0}^{t} y_i$ and $\Delta^{-2} y_t = \sum_{j=0}^{t} \sum_{i=0}^{j} \Delta^{1-r} y_i$.

Indicator function notation. For a statement S, $\mathbf{1}(S)$ is an indicator function that takes the value 1 if the statement S is true and zero otherwise.

[3] Here I am using the term classical cycle rather loosely as Burns and Mitchell also required minimum phase and cycle lengths. Procedures for locating turningpoints that meet these criteria are discussed in chapter 5 of Harding (2003).

the peak measures the date at which the growth rate stops increasing and begins to decrease. Similarly, the trough measures the date at which the growth rate stops decreasing and begins to increase.[4]

2.2 Lower order turning points

Lower order turning points refer to turning points in $\Delta^r y_t$ for $r < 0$. That is they refer to turning points in the r^{th} cumulative sum of y_t and thus involve successively stronger notions of what constitutes decisive change. For economic time series the case where $r = -1$ is of some importance. The reason for this is that most observed economic time series are generally thought to be I(d) (integrated of order d) with $0 \leq d < 2$. If a series is integrated of order less than $\frac{1}{2}$ then it is covariance stationary and thus the cumulative sum of the series $\Delta^{-1} y_t$ is integrated with order greater than one. Later (in section 5) I show that for any series which is integrated of order higher than one the expected number of turningpoints equal to zero. Thus, the highest order of differencing at which the proportion of turning points is zero provides a measure of the extent of differencing required to induce stationarity in the series.

3 The relationship between turning points and crossing points

Turning points are not the only features of the sample path of time series that are discussed. For example, it is commonplace to also discuss features such as whether a series is currently above or below trend; whether the output gap is positive or negative; and for countries that have inflation targets, whether inflation is above or below the target range. This leads to the concept of crossing points viz, points at which the series y_t crosses some critical mark.[5]

To show that crossing points can be represented as turning points in a related series let o_t be a mean zero series and assume that up crossing points

[4] The Economic Cycle Research Institute, New York, refers to the acceleration cycle as the cycle in the growth rate. However, this terminology runs the risk of being confused with the growth cycle.

[5] There is a mathematics and statistical literature on crossingpoints. Some references to that literature are in Kedem (1980) but the most comprehensive exposition is in Kedem (1994).

and down crossing points are represented by binary series $\measuredangle_t$ and $\leftthreetimes_t$ respectively that take the value one when a crossing occurs and zero otherwise. These crossing points are defined by equations (5) and (6) respectively.

$$\measuredangle_t = \mathbf{1}\left(o_t < 0, o_{t+1} > 0\right) \tag{5}$$

$$\leftthreetimes_t = \mathbf{1}\left(o_t > 0, o_{t+1} < 0\right) \tag{6}$$

Now consider the series O_t that is the cumulation of o_t, that is $O_t = \sum_{j=1}^{t} o_j$. Then, equations (5) and (6) can be written in terms of O_t as follows

$$\measuredangle_t = \mathbf{1}\left(\Delta O_t < 0, \Delta O_{t+1} > 0\right) \tag{7}$$

$$\leftthreetimes_t = \mathbf{1}\left(\Delta O_t > 0, \Delta O_{t+1} < 0\right) \tag{8}$$

But these are equivalent to the definitions of classical cycle calculus rule turning points in O_t. Thus the turning point and crossing point approach are equivalent. I will focus on turning points because of the long history of using this concept when studying the business cycle. This equality between crossing points in o_t and turning points in O_t is of great convenience as it means I can use the mathematics literature on crossing points.

4 Counts of the number of turning points

Let $\vee_t^r$and $\wedge_t^r, t = 1, ..N$ be r^{th} order turning points as defined in equations (1) and (2) and let NTP_r be the number of r^{th} order turning points. That is,

$$NTP_r = \sum_{t=2}^{N-1} \left(\vee_t^r + \wedge_t^r\right) \tag{9}$$

This quantity cannot be negative. The sum starts from 2 and runs to T-1 because calculus rule turning points have minimum pase duration of one period and thus one cannot determine whether the end points are turning points. Thus,

$$N - 2 \geq NTP_r \geq 0 \tag{10}$$

As established in the proposition below there exists a weak form of monotonicity in the number of classical and higher order turning points.

Proposition 1 ***Monotonicity of higher order turning points.*** *Let NTP_r be the number of r^{th} order turning points as defined in equation (9). Then, the following inequality applies: $NTP_{r+1} \geq NTP_r - 1$.*

Proof. Let there be a r^{th} order peak at t and the next r^{th} order trough be at date τ. The r^{th} order peak at t requires that $\Delta^r y_t > 0$ and $\Delta^r y_{t+1} < 0$. From this it follows that $\Delta^{r+1} y_{t+1} < 0$. The trough at τ requires that $\Delta^r y_\tau < 0$ and $\Delta^r y_{\tau+1} > 0$. From this it follows that $\Delta^{r+1} y_{\tau+1} > 0$. Thus there must be at least one $(r+1)^{st}$ order trough located between the dates $t+1$ and $\tau+1$. Thus, leaving aside endpoint problems there must be at least one $(r+1)^{st}$ trough for every r^{th} order trough. For a similar reason, leaving aside endpoint problems, there must be at least one $(r+1)^{st}$ order peak for every r^{th} order peak. The endpoint issue mentioned above arises if there is a r^{th} order turning point at the endpoint. Letting $t^{LTP,r}$ be the date of the previous turning point, the argument used above told us that there was a like $(r+1)^{st}$ turning point in the interval between $t^{LTP,r}+1$ and $T+1$. But since the series is only of length N it is possible that there is no $(r+1)^{st}$ order turning point to correspond to the last r^{th} order turning point. Thus, $NTP_{r+1} \geq NTP_r - 1$. ■

This result is of particular interest as it allows the following corollary which will prove to be useful later in the paper.

Corollary 2 ***Monotonicity and boundedness of the proportion of higher order turning points.*** *Let $p_{r,N}$ be the proportion of time occupied by r^{th} order calculus rule turning points, that is*

$$p_{r,N} = \frac{NTP_r}{N-2}$$

And let $p_r = \lim_{N\to\infty} p_{r,N}$ be the proportion of turning points as the length of the series becomes infinite. Then, p_r is bounded below by zero and above by one and is non decreasing in r. That is, $1 \geq p_r \geq 0$ and $p_{r+1} \geq p_r$ for all r.

Proof. Proposition 2 above implies that the proportion of time occupied by higher order turning points in a series of length N is weakly non-decreasing in the sense that $p_{r+1,N} \geq p_{r,N} - \frac{1}{N}$. Thus, taking limits yields

$$p_{r+1} = \underset{N\to\infty}{lim} p_{r+1,N} \geq \underset{N\to\infty}{lim} \left(p_{r,N} - \frac{1}{N-2}\right) = p_r$$

from which it follows that $p_{r+1} \geq p_r$ for all r..The boundedness of p_r follows directly from inequality (10). ■

The intuition behind the result just given is that for the calculus rule the pattern of turning points converges to an infinite dimension vector of ones as $r \to \infty$. That is, successive differencing of the time series ultimately extracts all of the information about the DGP of that series.

4.1 Certain properties of $p_{r,N}$

The quantity $p_{r,N}$ is an estimator of the proportion of time occupied by rth order turning points. To obtain the properties of this estimator we need to make some assumptions about y_t. I will assume that there exists some number k for which $\Delta^k y_t$ it is strictly stationary. This implies that $E\vee_t^k$ and $E\wedge_t^k$ are both constants that I represent by $\underline{\vee}^k$ and $\overline{\wedge}^k$ respectively.[6]

With this notation in place it is straight forward to see that $Ep_{r,N}$ is an unbiased estimate of $\underline{\vee}^r + \overline{\wedge}^r$. This is because

$$\begin{aligned} Ep_{r,N} &= E\frac{NTP_r}{N-2} \\ &= \frac{1}{N-2}\sum_{t=2}^{N-1} E\left(\vee_t^r + \wedge_t^r\right) \\ &= \underline{\vee}^r + \overline{\wedge}^r \end{aligned}$$

Proposition 3 ***The variance of*** $p_{r,N}$ ***is bounded above by*** $\frac{2(\underline{\vee}^r + \overline{\wedge}^r)}{(N-2)}$.

[6] Strict stationarity of $\Delta^k y_t$ is a stronger assumption than I require but it simplifies the analysis.

Proof. The variance of $p_{r,N}$ is defined as

$$\begin{aligned} Var\left(p_{r,N}\right) &= E\left[\frac{1}{N-2}\sum_{t=2}^{N-1}\left(\vee_t^r+\wedge_t^r\right)-\underline{\vee}^r+\overline{\wedge}^r\right]^2 \\ &= E\left[\frac{1}{N-2}\sum_{t=2}^{N-1}\left(\vee_t^r+\wedge_t^r\right)\right]^2-\left(\underline{\vee}^r+\overline{\wedge}^r\right)^2 \qquad (11) \end{aligned}$$

The binary nature of turning points implies that the first term in equation (11) achieves its maximum where $\vee_t^r + \wedge_t^r$ is perfectly synchronized at lag k that is where $E\left(\vee_t^r+\wedge_t^r\right)\left(\vee_{t-k}^r+\wedge_{t-k}^r\right) = \left(\underline{\vee}^r+\overline{\wedge}^r\right)$. Moreover, it is clear again because of the binary nature of turning points that $k = \frac{1}{\underline{\vee}^r+\overline{\wedge}^r}$. An implication of the assumption that turning points are perfectly synchronized at lag k is that they are perfectly non synchronized at lags that are not integer multiples of k. That is $E\left(\vee_t^r+\wedge_t^r\right)\left(\vee_{t-j}^r+\wedge_{t-j}^r\right) = 0$ if $j \neq ik$ for i an integer. This now allows a bound to be placed on the variance of $p_{r,N}$. Specifically,

$$\begin{aligned} Var\left(p_{r,N}\right) \leq\; & \frac{1}{(N-2)^2}\sum_{t=2}^{N-1}E\left(\vee_t^r+\wedge_t^r\right)^2 \\ & +\frac{2}{(N-2)^2}\sum_{i=1}^{\frac{N-2}{k}}\sum_{t=2+ik}^{N-1}E\left(\vee_t^r+\wedge_t^r\right)\left(\vee_{t-ik}^r+\wedge_{t-ik}^r\right) \\ & -\left(\underline{\vee}^r+\overline{\wedge}^r\right)^2 \end{aligned}$$

where it is understood that the expectations are taken under the assumption that the turning points are perfectly synchronized at lag k. Taking the expectations and making use of the fact that $k = \frac{1}{\underline{\vee}^r+\overline{\wedge}^r}$ and simplifying yields

$$\begin{aligned} Var\left(p_{r,N}\right) &\leq \frac{\underline{\vee}^r+\overline{\wedge}^r}{(N-2)}+2\frac{\underline{\vee}^r+\overline{\wedge}^r}{(N-2)^2}\sum_{i=1}^{\frac{N-2}{k}}(N-2-ik)-\left(\underline{\vee}^r+\overline{\wedge}^r\right)^2 \\ &= \frac{2\left(\underline{\vee}^r+\overline{\wedge}^r\right)}{(N-2)} \end{aligned}$$

which is as stated in the proposition. ■

This result is important because it underpins proofs of the following two propositions about the conistency of $p_{r,N}$ for $\underline{\vee}^r+\overline{\wedge}^r$ and the asymmptotic distribution of $\sqrt{N}\left(p_{r,N}-\underline{\vee}^r+\overline{\wedge}^r\right)$

Proposition 4 *$p_{r,N}$ is a consistent estimator of $\underline{\vee}^r + \overline{\wedge}^r$.*

Proof. This follows immediately from the results that $p_{r,N}$ is unbiased for $\underline{\vee}^r + \overline{\wedge}$ and its variance is bounded above by $\frac{2(\underline{\vee}^r + \overline{\wedge}^r)}{(N-2)}$ which converges to zero as N goes to infinity. ■

Proposition 5 *$\sqrt{N}\left(p_{r,N} - \underline{\vee}^r + \overline{\wedge}^r\right)$ converges in distribution to $N\left(0, \sigma^2\right)$ where $\sigma^2 \leq 2\left(\underline{\vee}^r + \overline{\wedge}^r\right)$*

Proof. $\vee_t^r + \wedge_t^r$ is a stationary stochastic process with mean $\underline{\vee}^r + \overline{\wedge}^r$ and variance this bounded above by $\frac{2(\underline{\vee}^r + \overline{\wedge}^r)}{(N-2)}$ thus, the result follows from from proposition 7.11 of Hamilton (1994). ■

A consistent robust estimate of the variance $\widehat{\sigma}^2$ can be obtained by application of the heteroscedasticity and autocorrelation consistent (HACC) robust estimator suggested by Kiefer and Vogelsang (2002). That is let u_t^r, v_t^r and $\widehat{\sigma}^2$ be defined as follows

$$u_t^r = \vee_t^r + \wedge_t^r - p_{r,N}$$

$$v_t^r = \sum_{i=1}^{t} u_i^r$$

$$\widehat{\sigma}^2 = N^{-2} \sum_{t=1}^{N} \left(v_t^r\right)^2$$

In some instances interest will centre on the vector of i^{th} to r^{th} order turning point proportions denoted $\underline{p}_{i,r,N} = \begin{bmatrix} p_{i,N} & p_{i+1,N} & \cdots & p_{i-1,N} & p_{r,N} \end{bmatrix}'$. The propositions regarding asymptotics extend to this case and the method described above also extends to obtaining a HACC robust estimator for the variance covariance matrix of $\underline{p}_{i,r,N}$.

5 The relationship between expected number of turning points and the parameters of the DGP for y_t

Interest centres on the relationship between the expected proportion of turning points and the parameters of the DGP for y_t. Where the DGP involves the elliptically symmetric distribution analytic results can be obtained as is shown in section 5.1 below. In other cases one must proceed via simulation.

5.1 The expected proportion of turning points when the time series has an elliptically symmetric distribution

The class of elliptically symmetric distributions introduced by Kelker (1970) is of considerable interest as it contains the multivariate normal, multivariate Student's t, multivariate logistic, and multivariate Cauchy distribution along with a large range of other less familiar distributions. A particular realization from this class of distributions is described by a vector of parameters μ, a symmetric matrix of parameters Σ and a function g. If v has elliptically symmetric distribution then we will write that in shorthand as $v\tilde{}ES(\mu,\Sigma,g)$. Appendix A provides a more detail discussion of the class of elliptically symmetric distributions and the properties relevant to this paper.

Assume that the vector $\begin{bmatrix} \Delta y_1 & \cdots & \Delta y_T \end{bmatrix}'\tilde{}ES_T(\mu_t,\Sigma,g_T)$ where μ_t is a polynomial time trend of degree q viz, $\mu_t = \mu_0 + \mu_1 t + ..\mu_b t^q$, Σ is a $T \times T$ positive definite symmetric banded matrix with ii^{th} element σ_0^2 and ij^{th} element $\sigma_{|i-j|}$ for $i \neq j$ and g_T is a density function generator.[7] Construct the $2 \times T$ matrix $B_{t+1,r}$ such that

$$B_{t+1,r} = \begin{bmatrix} 0 & 0 & {}_r a_{t+1}^r & \cdots & {}_0 a_{t+1}^r & 0 \\ 0 & {}_r a_t^r & \cdots & {}_{0r} a_t^r & 0 & 0 \end{bmatrix} \tag{12}$$

where ${}_j a_t^r$ places the signed integer $\binom{r}{j}(-1)^j$ in the t^{th} column. Thus,when applied to $\begin{bmatrix} \Delta y_1 & \cdots & \Delta y_T \end{bmatrix}'$ the matrix $B_{t+1,r}$ produces a 2×1 vector of the $(r+1)^{st}$ differences of y_{t+1} and y_t. That is

[7] The assumption that all elements in a band of the matxix Σ are equal is necessary to ensure that the distribution is strictly stationary. One could proceed in cases where the distribution is non stationary but it would introduce difficulties that are unnecessary at this stage.

$$\begin{bmatrix} \Delta^{r+1}y_{t+1} \\ \Delta^{r+1}y_t \end{bmatrix} = B_{t+1,r} \begin{bmatrix} \Delta y_1 \\ \vdots \\ \Delta y_T \end{bmatrix} \tag{13}$$

Since pre multiplication by $B_{t+1,r}$ is a linear transformation, it follows from property 1 of elliptically symmetric distributions stated in Appendix A that $\begin{bmatrix} \Delta^{r+1}y_{t+1} & \Delta^{r+1}y_t \end{bmatrix}' \,\tilde{}\, ES\left(\begin{bmatrix} \Delta^{r+1}\mu_{t+1} & \Delta^{r+1}\mu_t \end{bmatrix}', B_{t+1,r}\Sigma B'_{t+1,r}, g_2\right)$. The assumption that Σ is symmetric and banded with ij^{th} element $\sigma_{|i-j|}$ allows us to write $B_{t+1,r}\Sigma B'_{t+1,r}$ as follows

$$B_{t+1,r}\Sigma B'_{t+1,r} = \omega_r \begin{bmatrix} 1 & \rho_r \\ \rho_r & 1 \end{bmatrix}, \qquad \omega_r > 0\,,\ and\ -1 < \rho_r < 1. \tag{14}$$

Thus,

$$\begin{bmatrix} \Delta^{r+1}\left(y_{t+1} - \mu_{t+1}\right)/\omega_r \\ \Delta^{r+1}\left(y_t - \mu_t\right)/\omega_r \end{bmatrix} \tilde{}\, ES\left(0, \begin{bmatrix} 1 & \rho_r \\ \rho_r & 1 \end{bmatrix}, g_2\right) \tag{15}$$

Appendix B provides formulas that relate ρ_r to the elements of Σ.

Now interest is focused on the probability of a r^{th} order calculus peak at date t viz $\Pr\left(\wedge_t^r = 1\right)$. Letting $z_t = \left(y_t - \mu_t\right)/\omega_r$ This can be calculated as follows,

$$\begin{aligned} \Pr\left(\wedge_t^r = 1\right) &= \Pr\left(\Delta^{r+1}y_{t+1} < 0, \Delta^{r+1}y_t > 0\right) \qquad (16) \\ &= \Pr\left(\Delta^{r+1}z_{t+1} < -\Delta^{r+1}\mu_{t+1}/\omega_r, \Delta^{r+1}z_t > -\Delta^{r+1}\mu_t/\omega_r\right) \end{aligned}$$

Appealing to the arcsine result obtained in Appendix A we have that

$$\Pr\left(\wedge_t^r = 1\right) = \frac{1}{4} - \frac{1}{2\pi}\arcsin\rho_r + h\left(\Delta^{r+1}\mu_{t+1}/\omega_r, \rho_r\right) \tag{17}$$

where

$$\begin{aligned} h\left(\Delta^{r+1}\mu_{t+1}/\omega_r, \rho_r\right) &= \int_{-\Delta^{r+1}\mu_t/\omega_r}^{0}\int_{-\infty}^{-\Delta^{r+1}\mu_{t+1}/\omega_r} \frac{a_2}{\sqrt{(1-\rho_r^2)}}g_2\left(\frac{x^2 - 2\rho_r xy + y^2}{(1-\rho_r^2)}\right)dydx \\ & \qquad (18) \\ &\quad - \int_0^{\infty}\int_{\Delta^{r+1}\mu_{t+1}/\omega_r}^{0} \frac{a_2}{\sqrt{(1-\rho_r^2)}}g_2\left(\frac{x^2 - 2\rho_r xy + y^2}{(1-\rho_r^2)}\right)dydx \end{aligned}$$

Now the probability of a r^{th} order trough is

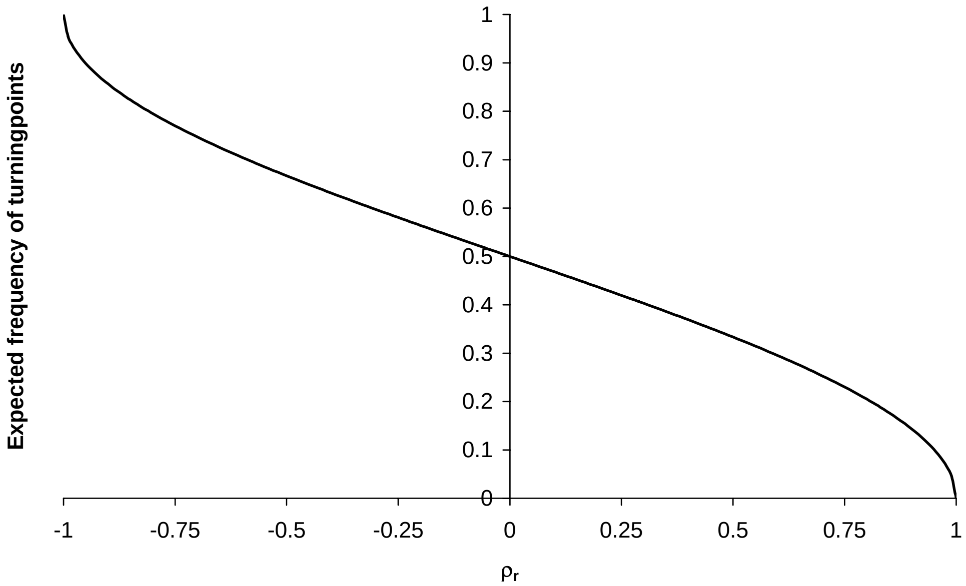
$$\Pr\left(\vee_t^r = 1\right) = \Pr\left(\Delta^{r+1} y_{t+1} > 0, \Delta^{r+1} y_t < 0\right) \tag{19}$$

appealing to the elliptical symmetry assumption yields the conclusion that $\Pr\left(\vee_t^r = 1\right) = \Pr\left(\wedge_t^r = 1\right)$. Since these are binary variables $\Pr\left(\wedge_t^r = 1\right) = E\left(\wedge_t^r\right)$ and thus $E\left(\vee_t^r + \wedge_t^r\right) = 2\Pr\left(\wedge_t^r = 1\right)$. Making use of equation (19) yields

$$\begin{aligned} E\left(\vee_t^r + \wedge_t^r\right) &= \Pr\left(\wedge_t^r = 1\right) + \Pr\left(\vee_t^r = 1\right) \\ &= \frac{1}{2} - \frac{1}{\pi}\arcsin\rho_r + 2h\left(\Delta^{r+1}\mu_{t+1}/\omega_r, \rho_r\right) \end{aligned} \tag{20}$$

which establishes that the expected proportion of r^{th} order turning points can be directly related to the parameters of the data generating process.

Figure 2: Relationship between $\frac{1}{2} - \frac{1}{\pi}\arcsin\rho_r$ and ρ_r



When seeking to gain intuition about equation (20) it is useful to restrict attention to the case where $\Delta^{r+1}\mu_t = 0$ so that $E\left(\vee_t^r + \wedge_t^r\right) = \frac{1}{2} - \frac{1}{\pi}\arcsin\rho_r$. Figure 2 show a plot of $\frac{1}{2} - \frac{1}{\pi}\arcsin\rho_r$ against ρ_r illustrating the important point that the expected proportion of turning points is monotonically decreasing in ρ_r and $E\left(\vee_t^r + \wedge_t^r\right)$ is bounded below by zero and above by one. Thus $\lim_{\rho_r \nearrow 1} E\left(\vee_t^r + \wedge_t^r\right) = 0$ and $\lim_{\rho_r \searrow -1} E\left(\vee_t^r + \wedge_t^r\right) = 1$. This leads to the conclusion that if y_t is integrated of order j the expected proportion of r^{th} order turning points is zero for $r < j - 1$ and is positive and is non decreasing in r for $r \geq j - 1$. This result looks like it will prove to be useful in testing for the order of differencing that results in a series being stationary. I have not included such tests in this draft of the paper since they involve issues with testing at the boundary of the parameter space.

A second application of the monotonicity of $\frac{1}{2} - \frac{1}{\pi} \arcsin \rho_r$ is that when $p_{r,N}$ is plotted for two time series the finding that the proportion of turning points for one series always lies below the other has the interpretation that the series with the lower $p_{r,N}$ is more persistent than than the other series in the sense that it has a lower value of ρ_r.

5.2 The expected proportion of turning points for an AR 1 process

As an illustration consider a gaussian AR 1 process in growth rates such as

$$\Delta y_t = (1 - \alpha)\mu + \alpha \Delta y_{t-1} + \varepsilon_t \qquad \varepsilon_t \tilde{} iid\, N(0, 1)$$

which has moving average representation

$$\Delta y_t = \mu + \sum_{j=0}^{\infty} \alpha^j \varepsilon_{t-j}$$

And thus, $\rho_0 = \alpha$, $\rho_1 = \frac{2\alpha - \alpha^2 - 1}{2}$.

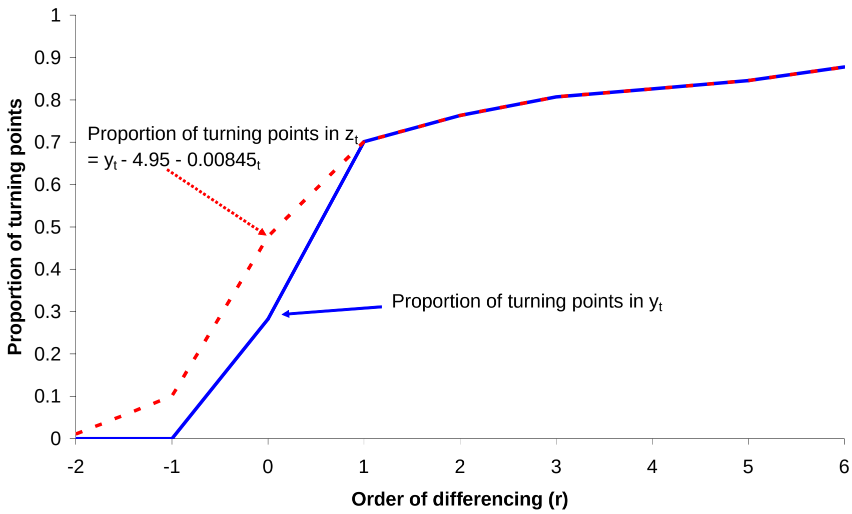
As an experiment I apply this model to US GDP for the period 1947.1 to 2001.4. On this data a typical estimate of α would be 0.34. Thus the AR 1 model yields an estimated proportion of 0 order turning points (in detrended ln(US GDP) of 0.44 whereas the data for 1947.1 to 2001.4 shows an estimated proportion of 0.48. The estimated proportion of first order turning points from the AR 1 model is 0.57 whereas the data yieds an estimate of 0.70. Thus the proprtion of turning points in the data conflicts with those predicted by an AR 1 model estimated on the data. The proportion of turning points predicted by the AR 1 model is less than that seen in the data for both orders of differencing and for the reasons given earlier this suggests that the

DGP is less persistent than the estimated AR 1 in growth rates. I have not put standard errors on these estimates for reasons of time but the discussion in section 4 shows how it can be done.

6 Applications

Figure 3 shows the proportion of turning points in the logarithm of US GDP (the solid line) and the proportion of turning points in the deviation of GDP from a linear trend (the dashed line). The following information can be extracted from Figure 3. First, the fact that the two proportions diverge for order of differencing below 1 and converge for orders of differencing above 1 indicates that ln(GDP) contains a deterministic component that is well approximated by a linear time trend. Second the fact that the proportion of turning points in z_t is substantially greater than zero is evidence that the series is not I(2), also the proportion of turning points in $\Delta^{-1}z_t$ is 0.1 which provides some evidence against the hypothesis that US GDP is I(1) and for the hypothesis that it is I(0). Finally, the proportion of turning points in $\Delta^{-2}z_t$ is 0.01 which provides evidence against the hypothesis that US GDP is I(-1).

Figure 3: Proportion of turning points in ln(US GDP) and its deviation from trend, 1947.1 to 2001.4



In the discussion above I have not referred to standard errors or developed formal statistical tests. In an earlier section I have developed a technology to do that but time prevents applying that technology in this paper. In the remainder of this section I present simulation based tests that were developed in Harding (2003). My objective in reporting those tests here is to assure the reader that the standard errors obtained for the proportion of turning points are quite small and that it is possible to use turning point based tests to distinguish between competing hypotheses.

As discussed in Blanchard and Fischer (1989), the maintained hypothesis that the major macroeconomic variables follow linear or log linear stochastic processes has influenced much of the direction of economic research. Given this hypothesis, empirical attention has focused on whether the linear stochastic process is trend stationary or difference stationary. For the post WWII United States this leads to the question of whether GDP is better described by equation (21) in the trend stationary case

$$y_t = 0.297 + .000318t + 1.316y_{t-1} - 0.356y_{t-2} + 0.00943e_t \tag{21}$$

or (22) in the difference stationary case.[8]

$$\Delta y_t = 0.00548 + 0.343\Delta y_{t-1} + 0.00943e_t \tag{22}$$

I assume that e_t ~iid N$(0, \sigma^2)$ — earlier I have shown that many of the results would hold provided $(e_1, .., e_T)$ comes from an elliptically symmetric distribution. The objective of the simulations below is to explore whether turning point based tests can shed light on the adequacy of these two models in terms of their capacity to match the dynamics of GDP.

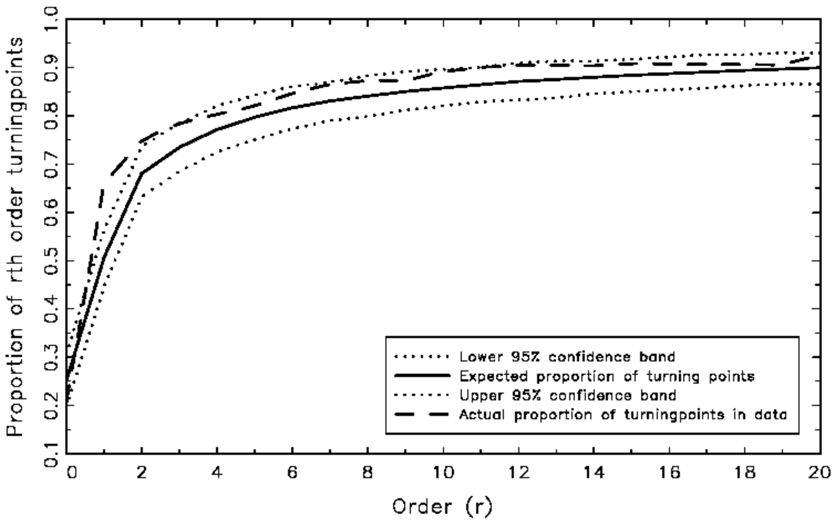
6.1 Simulation of a stationary AR(2) with Gaussian shocks

Figure (4) compares actual and simulated proportions of turning points; the simulations were for equation (21) in with iid Gaussian shocks and $N_{Sim} = 1000$. The property established earlier that $p_{r,N}$ is non decreasing in r is apparent in the plot of the actual proportion of turning points. The monotonicity in r of the population proportion of turning points is illustrated by the monotonicity of the average proportion of turning points from the simulation. The latter quantity is a good approximation to the comparable population quantity for large N_{Sim}.

[8] These equations were estimated on the log of US GDP in billions of chained 1996 dollars for the period 1947.1 to 2002.4.

In Figure 4 the actual proportion of turning points in the data is shown by the dashed line, the mean proportion of turning points from the simulation is shown by the solid line and the upper and lower 95 per cent confidence bands are shown by the dotted lines.[9] It is clear that the first and second order turning points lie above the upper confidence band suggesting that the model does not adequately capture the dynamics of the data.

Figure 4: Classical and higher order turningpoint frequencies actual and simulated for AR(2) in (log) levels of US GDP, 1947.1 to 2002.4



6.2 A turning point-based test of model fit

To supplement Figure 4 it would be useful to have a single summary statistic to quantify the fit of the simulated turning point frequencies to the data. To achieve this let $\widetilde{p}_r^i$ be the proportion of r^{th} order turning points in the i^{th} simulation where turning points are located by the calculus rule (k=1) and P be the matrix that contains all the simulated proportions for $i = 1, .., N_{Sim}$ and $r = 0, .., \overline{r}$, viz

[9] By 95 per cent confidence bands I mean that 97.5 per cent of simulations were above the lower band and 97.5 per cent of simulations were below the upper band. Thus 95 per cent of simmulations fell between the upper and lower bands.

$$P = \begin{bmatrix} \widetilde{p}_0^1 & \cdots & \widetilde{p}_{\overline{r}}^1 \\ \vdots & \ddots & \vdots \\ \widetilde{p}_0^{N_{Sim}} & \cdots & \widetilde{p}_{\overline{r}}^{N_{Sim}} \end{bmatrix} \tag{23}$$

The vector containing the mean proportion of turning points in the simulation $(\overline{P})$ is obtained as,

$$\overline{P} = \iota'_{N_{Sim}} P / N_{Sim} \tag{24}$$

Ω the covariance matrix from the simulated data is obtained as follows

$$\Omega = \left(P - \iota_{N_{Sim}} \overline{P}\right)' \left(P - \iota_{N_{Sim}} \overline{P}\right) / N_{Sim} \tag{25}$$

where $\iota_{N_{Sim}}$ is a $(N_{Sim} \times 1)$ vector of ones. Let P_i be the i^{th} row of P and form the scalars $Q_i = \left(P_i - \overline{P}\right) \Omega^{-1} \left(P_i - \overline{P}\right)'$ and $Q = \left(P_{data} - \overline{P}\right) \Omega^{-1} \left(P_{data} - \overline{P}\right)'$ then one can test the overall ability of the model to match the turning point information in the data by comparing Q with the distribution of Q_i $i = 1, .., N_{sim}$ from the simulation. The simplest method of effecting such a comparison is by computing the p-value viz,

$$\text{p-value} = \frac{1}{N_{Sim}} \sum_{i=1}^{N_{Sim}} 1\left(Q_i > Q\right) \tag{26}$$

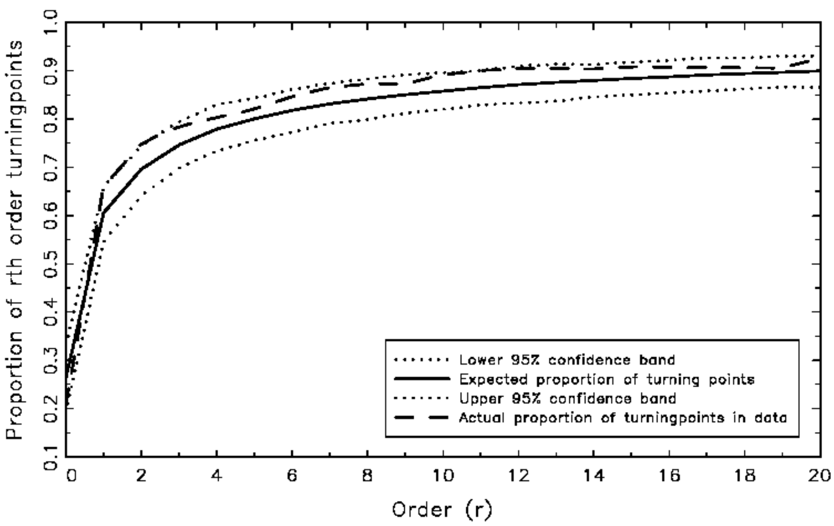
For the trend stationary AR(2) in equation (21) with a choice of $\overline{r} = 4$, $Q = 13.00$ with p-value=0.028.[10] Thus, one can conclude that the estimated trend stationary model does not adequately capture the dynamics of the data. It is important to note that this conclusion relates to the estimated model since I conditioned on the parameters of that model. If one sought to make a statement as to whether a model that is "close" to the estimated model could have generated the pattern of dynamics in the data then one would need to define "close" and base the simulations on that definition. One way of doing this would be to draw the parameters from their distribution N_{draw} times and for each parameter draw simulate the model N_{Sim} times yielding a total of $N_{draw} + N_{Sim}$ vectors of turning point frequencies. I have not made this extension here since I am primarily interested in the question of whether the estimated model could have generated the dynamics of the data.

[10] For a choice of $\overline{r} = 6$, $z_{data} = 14.6$ with a p-value=0.045.

6.3 Simulation of the difference stationary model with Gaussian shocks

Now attention turns to the difference stationary case of equation (22). Figure (5) compares actual and simulated turning point frequencies; the simulations were for equation (22) with iid Gaussian shocks and $N_{Sim} = 1000$. As was the case with Figure 4 the dashed line represent the observed frequencies, the solid line the mean of the simulated frequencies, and the two dotted lines are the upper and lower 95 per cent confidence bands. The difference stationary model does not appear to fit the data well. Most notably for $r > 1$ the data yields more turning points than expected on the basis of the model. For the reasons discussed earlier this suggests that there is less persistence in the data than in the model.

Figure 5: Classical and higher order turningpoint frequencies, actual and simulated for AR(1) in growth rates of US GDP, 1947.1 to 2002.4



Again with $\overline{r} = 4$ the test statistic $Q = 12.7$ with p-value$= 0.032$ for the difference stationary case. Thus, I conclude that the estimated difference stationary model in equation (22) does not adequately capture the dynamics of the data.[11]

[11] For a choice of $\overline{r} = 6$, $z_{data} = 14.3$ with a p-value=0.052.

7 Summary and Conclusions

In this paper it is shown that calculus rule turning points can be given a precise mathematical formulation. I have argued that turning points are very useful when studying dynamics because the expected proportion of turning points in a sample always exists and is bounded below by zero and above by one. This is not true of, for example, population moments since some or all of them may not exist. It was shown that one can measure what constitutes decisive change by locating turning points in higher differences of Δy_t. It was also shown that if the deterministic trend is removed from y_t then turning points located in $\Delta^r y_t$ are identical to crossing points in $\Delta^{r-1} y_t$ so there is a strong connection with the mathematics and statistics literatures on crossing points.

It was shown that expected proportions of turning points are related to the parameters of the DGP for the series of interest and thus one can recover estimates of the parameters from turning point frequencies. Specifically, if the DGP comes from the elliptically symmetric class of distributions then turning point frequencies and the parameters of the DGP are related by the arcsine rule. If the DGP does not come from the elliptically symmetric class of distributions then one can proceed via simulation to find the model parameters that yield turning point frequencies that match those found in the data.

It was also shown that turning points can be used to develop tests of how well particular models fit the data. Moreover, it was shown that given the assumption that the DGP comes from elliptically symmetric class of distributions it was possible to interpret the rejection of the model in a particular way.

I have shown that one can assess the fit of a dynamic model to the data by comparing the actual frequency of turning points observed in the data with the simulated frequencies of turning points from the model. The Gauss procedures written to do this can handle any model that is capable of being simulated. In the examples considered here I have analyzed two low order autoregressions where the question of interest is how much persistence there is in the data. I showed that simulation of the two models with Gaussian shocks led to their rejection.

I used the fact that when the DGP comes from the elliptically symmetric class of distributions, the arcsine rule could be used to interpret the simulations. This rule suggests that the data is generated by a process that is less persistent than either of the processes embedded in the models considered. In this sense I read the results as supporting the trend stationary case.

Thus, in addition to being a useful tool for studying dynamics, turning points frequencies seem likely to be useful as an adjunct to unit root testing. Of course one will need to undertake monte carlo studies to assess how well these tests perform.

Appendices

A Some properties of elliptically symmetric distributions

My objective in this appendix is to provide a brief discussion of this class of distributions and outline its relevance to the analysis of dynamics via turnng points. There are several ways of defining the class of elliptically symmetric distributions. I found the approach below useful.

Let X be a $J \times 1$ random vector with distribution function F(X) and $\mathbf{v}$ be a $J \times 1$ vector then the characteristic function $\varphi_X(v)$ is defined as follows

$$\varphi_X(v) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{iv'X} dF$$

Let Φ represent the set of continuous scalar functions that map the non negative real line into itself viz,

$$\Phi = \{\phi(.) : [0, \infty) \rightarrow [0, \infty)\}$$

If the characteristic function of X is of the form $\varphi_X(\mathbf{v}) = e^{i\mathbf{v}'\mu}\phi(\mathbf{v}'\Sigma\mathbf{v})$ for $\phi \epsilon \Phi_J$ and Σ a symmetric positive definite matrix we say that X belongs to the class of elliptically symmetric distribution which we write in shorthand as $X\tilde{\ } ES_J(\mu, \Sigma, \phi)$.

Of course if $\varphi_X(\mathbf{v})$ is a valid characteristic function it must satisfy the fundamental theorem of characteristic functions which relates to inversion. This places further restrictions on ϕ. The multivariate inversion formula is tedious to write out for the general case and I will not do so here, the reader is referred to Stuart and Ord (1994, Sec 4.17 p. 140) for details. However, it is instructive to consider the case where the distribution function is continuous and thus a density exists. In this case the inversion theorem links the density and the characteristic function as follows

$$f(x) = \frac{1}{(2\pi)^J} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-i\mathbf{v}'x} \varphi_X(\mathbf{v}) \, d\mathbf{v} \qquad (27)$$

For the elliptically symmetric class of distributions this integral becomes

$$
\begin{aligned}
f(x) &= \frac{1}{(2\pi)^J}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{-i\mathbf{v}'x}e^{i\mathbf{v}'\mu}\phi(\mathbf{v}'\Sigma\mathbf{v})\,d\mathbf{v} \\
&= \frac{1}{(2\pi)^J}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{-i\mathbf{v}'(x-\mu)}\phi(\mathbf{v}'\Sigma\mathbf{v})\,d\mathbf{v} \qquad (28)
\end{aligned}
$$

Fang, Kotz and Ng (1990) shows that if X has an elliptically symmetric distribution then the density (if it exists) is given by equation (29)

$$
f_X(x) = \frac{a_J}{\sqrt{|\Sigma|}} g_J\left((x-\mu)'\Sigma^{-1}(x-\mu)\right) \qquad (29)
$$

where $g_J(.)$ is a function that maps the non-negative real line into itself. The constant a_J is given by equation (30).

$$
a_J = \frac{\Gamma(J/2)}{(2\pi)^{\frac{J}{2}}}\left[\int_0^{\infty} v^{\frac{J}{2}-1} g_J(v)\,dv\right]^{-1} \qquad (30)
$$

Thus, the density exists if inequality (31) is satisfied

$$
\int_0^{\infty} v^{\frac{J}{2}-1} g_J(v)\,dv < \infty \qquad (31)
$$

Where the density exists it is useful to also use $X\tilde{\ }ES_J(\mu, \Sigma, g_J)$ as shorthand for the statement that 'X has an elliptically symmetric distribution with density function proportional to g_J the centrality parameter vector μ and the dispersion matrix Σ'.

Elliptically symmetric distributions have a number of useful properties some of which are stated below.

Property 1: Linear combinations of X have an elliptically symmetric distribution If $X\tilde{\ }ES(\mu, \Sigma, \phi)$ with density $f_X(x; g_J, \mu, \Sigma)$, B is a $(q \times J)$ matrix of rank q and b is a $(q \times 1)$ vector, then $(BX+b)\tilde{\ }ES_q(B\mu + b, B\Sigma B', \phi)$. And, if the density function $f_X(x; g_J, \mu, \Sigma)$ exists, then $f_{BX+b}(.;.) = f_X(x; g_q, B\mu + b, B\Sigma B')$. See Fang et al. (1990, p. 43).

Property 2: Marginal density. Partition X, μ and Σ as follows. $X = [X_1, X_2]'$, $\mu = [\mu_1, \mu_2]'$ and $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$ then if $X\tilde{\ }ES(\mu, \Sigma, \phi)$ it is the case that $X_1\tilde{\ }ES(\mu_1, \Sigma_{11}, \phi)$ and $X_2\tilde{\ }ES(\mu_2, \Sigma_{22}, \phi)$. That is the sub vectors have elliptically symmetric distribution with the same characteristic generating function ϕ. See Fang et al. (1990, p. 43).

Property 3. Conditional density Using the partition of X above, and letting $X_1|X_2 = x$ denote the random vector X_1 conditional on $X_2 = x$. Then

$X_1|X_2 = x$ has elliptically symmetric distribution with conditional mean $\mu_{1|2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x - \mu_2)$ and conditional variance $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$; see Fang et al. (1990, p. 45-46) for further details.

Property 4. Arcsine rule Let $\begin{bmatrix} X & Y \end{bmatrix}' \tilde{} ES\left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, g_J\right)$ then, $\Pr(X > c_1, Y < c_2)$ for $c_1 \leq 0, c_2 \leq 0$ is given by equations (32) and (33).

$$\Pr(X > c_1, Y < c_2) = \frac{1}{4} - \frac{1}{2\pi}\arcsin\rho + h(c_1, c_2, \rho) \tag{32}$$

where

$$\begin{aligned} h(c_1, c_2, \rho) &= \int_{c_1}^{0}\int_{-\infty}^{c_2} \frac{a_2}{\sqrt{(1-\rho^2)}} g_2\left(\frac{x^2 - 2\rho xy + y^2}{(1-\rho^2)}\right) dydx \\ &\quad - \int_{0}^{\infty}\int_{c_2}^{0} \frac{a_2}{\sqrt{(1-\rho^2)}} g_2\left(\frac{x^2 - 2\rho xy + y^2}{(1-\rho^2)}\right) dydx \end{aligned} \tag{33}$$

Property 4 is not in the literature, which focuses on the case $c_1 = c_2 = 0$. But property 4 follows in a straightforward way from the following argument. Since the probability of interest is $\Pr(X > c_1, Y < c_2)$ for $c_1 \leq 0, c_2 \leq 0$, it can be obtained as

$$\begin{aligned} \Pr(X > c_1, Y < c_2) &= \int_{c_1}^{\infty}\int_{-\infty}^{c_2} \frac{a_2}{\sqrt{(1-\rho^2)}} g_2\left(\frac{x^2 - 2\rho xy + y^2}{(1-\rho^2)}\right) dydx \\ &= \int_{0}^{\infty}\int_{-\infty}^{0} \frac{a_2}{\sqrt{(1-\rho^2)}} g_2\left(\frac{x^2 - 2\rho xy + y^2}{(1-\rho^2)}\right) dydx \\ &\quad + \int_{c_1}^{0}\int_{-\infty}^{c_2} \frac{a_2}{\sqrt{(1-\rho^2)}} g_2\left(\frac{x^2 - 2\rho xy + y^2}{(1-\rho^2)}\right) dydx \\ &\quad - \int_{0}^{\infty}\int_{c_2}^{0} \frac{a_2}{\sqrt{(1-\rho^2)}} g_2\left(\frac{x^2 - 2\rho xy + y^2}{(1-\rho^2)}\right) dydx \end{aligned} \tag{34}$$

Making use of the result that $\Pr(X > 0, Y < 0) = \frac{1}{4} - \frac{1}{2\pi}\arcsin\rho$ see Fang et al. (1990), and naming the second part of the integral $h(c_1, c_2, \rho)$ the result stated in proposition 4 holds.[12]

Property 5. Moments of elliptically symmetric distributions. If $X \tilde{} ES_J(\mu, \Sigma, g_J)$ and the first moment of X is finite then $E(X) = \mu$. If

[12] It is interesting to note that the mathematics behind the arcsin rule for the gaussian case originates in the late 19th century with the work of Stieltjes (1889) and Sheppard (1899). One wonders why Mitchell did not use this mathematics to provide a rigorous foundation for turningpoint analysis.

the second moment of X is finite then $Cov(X) = c\Sigma$ for some constant c— in the Gaussian case where Σ is nonsingular $c = 1$ and thus $Cov(X) = \Sigma$. Higher moments are obtained by further differentiation of the characteristic function $\varphi_X(\mathbf{v}) = e^{i\mathbf{v}'\mu}\phi(\mathbf{v}'\Sigma\mathbf{v})$ wrt v and evaluating the derivative at $v = 0$.

A.1 Generating linear time series with elliptically symmetric distributions

This section turns to the question of under what circumstances will a time series $\{y_t\}_{t=1}^T$ have an elliptically symmetric distribution. Let y_t be generated by the linear autoregression described in matrix notation by equation (35)

$$A\begin{bmatrix} y_T \\ \vdots \\ y_1 \end{bmatrix} = \begin{bmatrix} e_T \\ \vdots \\ e_1 \end{bmatrix} \tag{35}$$

where A is a $T\times T$ invertible matrix. Assume that $\begin{bmatrix} e_T & \cdots & e_1 \end{bmatrix}' \,\tilde{}\, ES_T(0, \Sigma, g_T)$ then, by property 1 above $\begin{bmatrix} y_T & \cdots & y_1 \end{bmatrix}' \,\tilde{}\, ES_T(0, A^{-1}\Sigma A'^{-1}, g_T)$. Now this is a very general class of linear model and some considerable attention needs to be paid to the specification of the A and Σ matrices.

Two sets of restrictions are common in the time series literature, one on A, the other on Σ. The former restriction requires that the matrix A have zeros below the diagonal, ones on the diagonal, and typical element $A_{i,j+i} = \alpha_j$ for $j = 1, ..p$ where where the roots of $1 - \alpha_1 z - .. - \alpha_p z^p = 0$ lie outside the unit circle. That is

$$A = \begin{bmatrix} 1 & -\alpha_1 & \cdots & -\alpha_p & 0 \\ 0 & 1 & & & -\alpha_p \\ 0 & & & & \cdots \\ \vdots & & & 1 & -\alpha_1 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} \tag{36}$$

The second restriction is to let $\Sigma = B'B$ where B is an upper diagonal matrix with zeros below the diagonal, ones on the diagonal and typical element $B_{i,j+i} = \beta_j$ for $j = 1, ..q$. That is

$$B = \begin{bmatrix} 1 & \beta_1 & & \beta_q & 0 \\ 0 & 1 & & & \beta_q \\ 0 & & & & \vdots \\ \vdots & & & 1 & \beta_1 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} \tag{37}$$

Thus by construction with these restrictions y_t is an ARMA(p, q) with elliptically symmetric distribution.

B Relating ρ_r and ω_r to the DGP of y_t

The quantity ρ_r used in equation (14) is given by equation (38) where it is related to r and the dispersion parameters of the DGP of y.

$$\rho_r = \frac{\frac{\sigma_1}{\sigma_0} + \sum_{j=2}^{r+1} (-1)^{1-j} \frac{r!r!}{(r-j+1)!(r+j-1)!} \frac{\sigma_j + \sigma_{j-2}}{\sigma_0}}{1 + 2\sum_{j=1}^{r} (-1)^j \frac{r!r!}{(r-j)!(r+j)!} \frac{\sigma_j}{\sigma_0}} \tag{38}$$

The quantity ρ_r is obtained as the following ratio $\rho_r = \frac{\rho_r \omega_r}{\omega_r}$ where the quantities ω_r and $\rho_r \omega_r$ in equation (14) are linked to the parameters of the Σ matrix $\sigma_{|i-j|}$ via equations (39) and (40).

$$\omega_r = \sigma_0 \binom{2r}{r} + 2\sum_{j=1}^{r} (-1)^j \sigma_j \binom{2r}{r-j} \tag{39}$$

$$\rho_r \omega_r = \sigma_1 \binom{2r}{r} + \sum_{j=2}^{r+1} (-1)^{1-j} \binom{2r}{r-j+1} [\sigma_j + \sigma_{j-2}] \tag{40}$$

This following sections established these two results for ω_r and $\rho_r \omega_r$

B.1 Obtaining ω_r

Equation (39) can be obtained by expanding the (1,1) term in $B_{t+1,r} \Sigma B'_{t+1,r}$ into its component terms as in equation (41).

$$\omega_r = \sigma_0 \sum_{i=0}^{r} \left[(-1)^i \binom{r}{i} \right]^2 + 2\sum_{j=1}^{r} \left[\sigma_j \sum_{i=0}^{r-j} (-1)^i \binom{r}{i} (-1)^{i+j} \binom{r}{j+i} \right] \tag{41}$$

which can be simplified to yield equation (39) by making use of two results from combinatorics. The first result is that

$$\binom{x}{v} = \binom{x}{x-v} \tag{42}$$

Applying this result to equation (41) yields

$$\omega_r = \sigma_0 \sum_{i=0}^{r} \binom{r}{i}\binom{r}{r-i} + 2\sum_{j=1}^{r} (-1)^j \sigma_j \sum_{i=0}^{r-j} \binom{r}{i}\binom{r}{r-j-i} \tag{43}$$

The second combinatoric result is that

$$\sum_{i=0}^{w} \binom{x}{i}\binom{v}{w-i} = \binom{x+v}{w} \tag{44}$$

Applying this result to equation 43 yields

$$\omega_r = \sigma_0 \binom{2r}{r} + 2\sum_{j=1}^{r} (-1)^j \sigma_j \binom{2r}{r-j}$$

which is the result stated in equation (39).

B.2 Obtaining $\omega_r \rho_r$

Now equation (40) can be obtained in a similar fashion viz, expanding the (1,2) term in $B_{t+1,r}\Sigma B'_{t+1,r}$ into its component terms as in equation (45)

$$\begin{aligned} \rho_r\omega_r &= \sigma_1 \sum_{i=0}^{r} (-1)^i \binom{r}{i} (-1)^i \binom{r}{i} \\ &\quad + \sum_{j=2}^{r+1} [\sigma_j + \sigma_{j-2}] \sum_{i=j-1}^{r} (-1)^i \binom{r}{i} (-1)^{i-j+1} \binom{r}{i-j+1} \end{aligned} \tag{45}$$

Making use of the combinatoric result (42) equation (45) becomes

$$\rho_r\omega_r = \sigma_1 \sum_{i=0}^{r} \binom{r}{i}\binom{r}{r-i} + \sum_{j=2}^{r+1} (-1)^{1-j} [\sigma_j + \sigma_{j-2}] \sum_{i=j-1}^{r} \binom{r}{i}\binom{r}{r+j-1-i} \tag{46}$$

Simplifying this expression using the combinatoric results used earlier yields equation (47)

$$\rho_r\omega_r = \sigma_1 \binom{2r}{r} + \sum_{j=2}^{r+1} (-1)^{1-j} \binom{2r}{r-j+1} [\sigma_j + \sigma_{j-2}] \tag{47}$$

References

Apostol, T. M.: 1974, *Mathematical Analysis*, Addison-Wesley, Menlo Park, California.

Berndt, D. J. and Clifford, J.: 1996, Finding patterns in time series:a dynamic programming approach, , pp. 229–248.

Blanchard, O. J. and Fischer, S.: 1989, *Lectures on Macroeconomics*, MIT Press.

Campbell, J. Y., Lo, A. W. and MacKinlay, A. C.: 1997, *The Econometrics of Financial Markets*, Princeton University Press.

Card, S. K., Mackinlay, J. D. and Shneiderman, B.: 1999, *Readings in Information Visualization: Using Vision to Think*, Morgan-Kaufmann Publishers.

Carlis, J. V. and Konstan, J. A.: 1998, Interactive visualizations of serial periodic data, *Proceedings of UIST 98 (San Francisco CA, November)*, ACM Press, pp. 29–38.

Edwards, R. and Magee, J.: 1966, *Technical Analysis of Stock Trends*, John Magee.

Fang, K. T., Kotz, S. and Ng, K. W.: 1990, *Symmetric Multivariate and Related Distributions*, Chapman and Hall.

Fayyad, U. M., Piatesky-Shapiro, G., Smyth, P. and Uthursamy, R.: 1996, *Advances in Knolwedge Discovery and Data Mining*, AAAI Press.

Harding, D.: 2003, *Essays on the Business Cycle*, Phd thesis submitted, Yale.

Harding, D. and Pagan, A. R.: 2001, Some econometric problems arising from regressions with constructed state variables. The University of Melbourne.

Harding, D. and Pagan, A. R.: 2002, Synchronisation of cycles. The University of Melbourne.

Jevons, W. S.: 1862, On the study of periodic commercial fluctuations, *Report of BAAS* .

Jugular, C.: 1862, *Des Crises commerciales et leur retour periodique en France, en Angleterre, et aux Etats-Unis*, Guillaumin, Paris.

Kedem, B.: 1980, *Binary Time Series*, M. Dekker.

Kedem, B.: 1994, *Time Series Analysis by Higher Order Crossings*, IEEE Press.

Kelker, D.: 1970, Distribution theory of spherical distributions and location scale parameter generalization, *Sankhya* **32**, 419–430.

Little, J. B. and Rhodes, L.: 1978, *Understanding Wall Street*, Liberty Publishing, Cockeysville, MD.

Lo, A. W., Mamaysky, H. and Wang, J.: 2000, Foundations of technical analysis: Computational algorithms, statistical inference and empirical implementation, *The Journal of Finance* **LV**(4), 1705–1765.

Murphy, J.: 1983, *Technical Analysis of the Futures Markets*, New York Institute of Finance, New York.

Oomes, S.: 1980, *Human visual perception of spatial structure: symmetry, orientation, and attitude*, PhD thesis, University of Nijmegen, The Netherlands. Available on web at http://vision-lab.psy.ohio-state.edu/oomes/dissertation.htm.

Pagan, A. R. and Ullah, A.: 1999, *Nonparametric Econometrics*, Cambridge University Press, Cambridge.

Schumpeter, J. A.: 1939, *Business Cycles: A theoretical, historical and statistical analysis of the Capitalist process*, McGraw-Hill, New York.

Sheppard, W. F.: 1899, On the application of the theory of error to the cases of normal distribution and normal correlation, *Philosophical Transations of the Royal Society London* **A 192**, 101–167.

Silva, S. F. and Catarci, T.: 2000, Visualization of linear time-oriented data: a survey, *Proceedings of the First International Conference on Web Information Systems Engineering*, IEEE Computer Society.

Stieltjes, T. J.: 1889, Extrait d'une lettre addressee a m. hermite, *TBulletin of Science and Mathematics, 2nd Series* **13**(2), 170–72.

Stuart, A. and Ord, J. K.: 1994, *Kendall's Advanced theory of Statistics: Volume 1, Distribution Theory*, sixth edn, Edward Arnold, London.