DEPARTMENT OF ECONOMICS

COLLEGE OF BUSINESS AND ECONOMICS

UNIVERSITY OF CANTERBURY

CHRISTCHURCH, NEW ZEALAND

Cliometrics and Time Series Econometrics: Some Theory and Applications

David Greasley and Les Oxley

WORKING PAPER

No: 56/2010

Department of Economics
College of Business and Economics
University of Canterbury
Private Bag 4800, Christchurch
New Zealand

WORKING PAPER No. 56/2010

Cliometrics and Time Series Econometrics: Some Theory and Applications

David Greasley[1] and Les Oxley[2†]

September 4, 2010

Abstract: The paper discusses a range of modern time series methods that have become popular in the past 20 years and considers their usefulness for cliometrics research both in theory and via a range of applications. Issues such as, spurious regression, unit roots, cointegration, persistence, causality, structural time series methods, including time varying parameter models, are introduced as are the estimation and testing implications that they involve. Applications include a discussion of the timing and potential causes of the British Industrial Revolution, income 'convergence' and the long run behaviour of English Real Wages 1264 – 1913. Finally some new and potentially useful developments are discussed including the mildly explosive processes; graphical modelling and long memory.

Keywords Time series; cointegration, unit roots, persistence, causality, cliometrics, convergence, long memory, graphical modelling, British Industrial Revolution.

JEL classifications N33, O47, O56, C22, C32

Acknowledgements: We wish to acknowledge the numerous, wide ranging and stimulating conversations on time series econometrics with the late Sir Clive Granger – his brilliance will be deeply missed. Section 4.1 benefited from the research assistance of Elizabeth Duston. The second author wishes to acknowledge financial support from the Marsden Fund.

[1] School of History, Classics and Archaeology, University of Edinburgh
[2] Department of Economics and Finance, University of Canterbury, Private Bag 4800, Christchurch, New Zealand
† Corresponding author: Email. les.oxley@canterbury.ac.nz

1. Introduction

The power of a popular test is irrelevant. A test that is never used has zero power. (McAleer, 1994, 2005)

The publications of Granger and Newbold (1974), Dickey and Fuller (1979, 1981), Nelson and Plosser (1982), Engle and Granger (1987) and Johansen (1988, 1995) have changed the way we think about and undertake time series econometrics. Although discussion of trends and their importance in economic time series can be traced to Yule (1926) and Kendall (1954), until the 1970s and '80s the field remained mostly a curiosity. Statistical research focussed on the isolation of trends from cycles in a world (implicitly) assumed to be generating stationary data. This was the world that quantitative economic historians occupied and where some remain. However, the message of Granger and Newbold (1974) were simple, yet powerful:

> 'In our opinion the econometrician can no longer ignore the time series properties of the variables with which he is concerned - except at his peril. The fact that many economic 'levels' are near random walks or integrated processes means that considerable care has to be taken in specifying one's equations.'

The seminal time series papers and the research agendas they created have particular relevance for quantitative economic history. Perhaps in more than any other area of applied economics, time series cliometrics utilizes long time spans of data which will often have the characteristics of non-stationarity in levels and/or which might experience structural change, persistence, large 'shocks', large outliers, conditional heteroskedasticity and potentially switching time series properties. The traditional cliometrics topics relating for example to 'trends', 'cycles' and 'path dependency' have

been both challenged and given new and different meanings as a consequence of the ways we now typically analyse, estimate and test time series data.

Whether or not macroeconomic time series exhibit 'unit roots' still remains a hotly debated issue see for example Darne (2009), and research continues to seek out, if not the 'truth', at least the 'data generating process' (DGP). However, whatever the outcome, current cliometrics research involving time series data has been fundamentally changed by the recent developments in time-series econometrics.

The purpose of this paper is to both inform those who may be unfamiliar with the changes in econometric methods that have ensued as a consequence of this econometric revolution of the nature and effects of these changes and also to provide some examples of how these new approaches have been (and might be) used to address some traditional and new areas of cliometrics. The intention is not to provide a full presentation of all the technical properties of each and every test and estimation method, but rather to motivate the need for and use of such tests (and methods) and refer the reader to software where such tests and methods can be applied. Most of the tests and methods discussed below can be easily and robustly implemented via packages such as *EViews* 6 and 7; *STATA* 9, 10 or 11; *STAMP*, and *RATS*. The excellent user manuals that accompany these packages are also a comprehensive source of technical detail which, on occasion, we will also refer to in this paper. Exceptions to these intentions relate to sections where we highlight new and emerging areas where, as one might expect, implementation in the packaged software tend to lag the theoretical developments.

The paper will start by outlining the nature of what we believe are the more important changes in method and interpretation that have occurred as a consequence of

the new time series developments (non-stationarity and unit root testing; measures of persistence; cointegration; Granger causality). Examples drawn from the cliometrics literature will complement the more technical aspects of the discussion. Finally some new and emerging areas of time series econometrics will be discussed in relation to their potential applications to cliometrics research including long memory models and applications of graphical methods.

In a single paper it is impossible to cover all areas of relevance to cliometrics research with time series data. As a consequence, the following areas are excluded or given little emphasis, or simply enter by way of a specific example; spectral-based methods; Bayesian-based methods; non-parametric and semi-parametric approaches.

Prior to the new developments in time series methods, the 'meat and drink' of time series quantitative economic history was the detection and measurement of (deterministic) 'trends' (if only to then subsequently remove them), their potential shifting location based upon tests for 'structural breaks' and the decomposition of data into trends, cycles and possibly deviations from 'long run trend'. The new econometrics considers stochastic trends versus deterministic trends; nonlinear trends; common trends (and common cycles, see Vahid and Engle, 1993); detrending and trend extraction (Hodrick and Prescott, 1997, Harvey, 1989, Harvey and Jaeger, 1993). Phillips (2005) provides an excellent overview of the challenges faced by the notion of trends, arguing that we have 'only scratched the surface' when it comes to understanding what trends are, how to model them and the consequences of getting that modelling wrong.

The paper comprises the following sections. In Section 2 we consider the notion and the importance of spurious regression as a precursor to Section 3 which introduces a

range of statistical issues relating to ideas of and tests for non-stationarity, cointegration, Granger causality, persistence and structural time series modelling. Section 4 comprises a range of empirical applications which utilise the estimation methods and tests discussed in Section 3. Section 4.1 contains a brief overview of the time series papers and topics published in the two main cliometrics journals, *Explorations in Economic History* (2000-2009) and *Cliometrica*, 2007-2009. Sections 4.2 - 4.5 present applications of such topics as; when did the British Industrial Revolution begin and what were its causes; development blocks and New Zealand economic development, testing for convergence in real GDP per capita; and new results interpreting English real wages data 1264 – 1913. Section 5 introduces new developments/applications with potential for cliometrics including the mildly explosive process of Phillips and Yu (2009); graphical modelling and implications for causality testing; and long memory estimation. Section 6 offers some final thoughts and Section 7 concludes.

2. Prologue: Spurious Regression

Granger and Newbold (1974) state that:

> 'It is common to see reported .. time series regression equations with apparently high degree of fit as measured by R^2 , but with very low reported Durbin Watson statistics. We find this strange given that almost every econometrics textbook warns of the dangers of autocorrelated errors. ...The most extreme example we encountered was an R^2=0.99 and d=0.093....There are three main consequences of autocorrelated errors:
>
> 1. Estimates are inefficient
> 2. Forecasts based upon the regression are sub-optimal
> 3. The usual significance tests on the coefficients are invalid"

They then concentrate on point 3 and the 'discovery' of spurious relationships – the 'nonsense correlations' between a pair of independent I(1) processes reported in Yule (1926). In what was an empirically-based paper, they show that with non-stationary series, a high R^2 should not, on the basis of traditional tests, be regarded as evidence of a significant relationship between autocorrelated series. In a practical sense, whenever $DW<R^2$ alarm bells should sound as the classical properties of the error term have been violated with empirically, a highly persistent error equated with highly persistent dependent and independent variables. These highly persistent variables will typically have the property, in levels, of non-stationary series.

Phillips (1986, 1998) provides a theoretical foundation to the empirical examples of Granger and Newbold (1974). Consider a regression of:

$$X_t = \alpha + \beta Y_t + \varepsilon_t \qquad (1)$$

where X and Y are independent random walks and ε_t is a zero mean Gaussian white noise process. Under these conditions Phillips shows that the OLS estimates of the model of equation (1) have no interpretable *t* statistics for α and β, as the distributions of these statistics diverge as the sample size increases. The estimate of β converges to some random variable whose value changes from sample to sample and the Durbin-Watson statistics for the equation tends to zero.

Rather worryingly, it's not just random walks that can cause spurious regression-type results. Granger, Hyung and Jeon (2001) consider a variety of independent stationary (but highly persistent) processes and produced some 'worrying' results. For example, if the two processes were AR1, one with a

coefficient of 0.5 and the other a coefficient >0.75, then 20-26% of the regressions would spuriously suggest a relationship.

3. Some statistical issues

3.1 Overview

It has now become standard practice in time series econometrics to use univariate tests to consider the existence of a unit root as a pre-test prior to subsequent estimation or inference. Such a practice is seen as consistent with the guidelines of Granger and Newbold (1974) on the dangers of 'spurious regression' and the need to consider only 'balanced' relationships. Here 'balance' relates to a situation where a relationship to be estimated/tested includes only variables with the same (or lower) orders of integration. Order of integration relates to the numbers of times a series needs to be differenced to produce the property of stationarity. For example, a stationary series is said to be 'integrated of order 0', denoted I(0) as it needs to be differenced 'zero times' to become stationary (as it already is stationary). 'Integrated of order 1' means the series is rendered stationary by differenced 1 times (first differenced); I(2) differenced twice (difference of the difference), etc. The idea of 'balance' is that any attempt to explain (say) an I(1) variable by a series of I(0) would be theoretically impossible (although due to the low power of some of the empirical tests we might produce evidence of such an occurrence – perhaps the source of some assumed 'puzzling results' in econometrics see the *Epilogue* section 6.0 below). To be a balanced relationship, the 'order of integration' of the variables in a relationship to be estimated must be of equal orders of integration. In an 'all variables are stationary world' this is ensured by definition, but in a world where some variables are non-stationary in levels, some in differences, some stationary as linear combinations (cointegrated), etc., careful thought and modelling is required.

Below we will provide a very brief overview of unit root tests and developments, including the effects of 'structural breaks' and in section 4 we consider how the unit root testing approach can be applied to cliometrics research.

3.2 Plain vanilla Unit Root tests

At the centre of some of the practical implications of the difference between stationary and non-stationary processes is the *persistence* of shocks, i.e., transitory or permanent. One of the simplest ways to model and subsequently *infer* persistence is to investigate the properties of uni-variate series, in particular, whether they are *trend stationary* (TS) or *Difference Stationary* (DS). The class of model most commonly used to describe temporary, i.e. non-persistent, deviations about a trend is:

$$y_t = a + bt + u_t \qquad (2)$$

where y_t is typically the natural logarithm of the variable of interest, t describes the trend and u_t is a stationary invertible auto-regressive moving average (ARMA) process. This process is stationary in levels or *trend stationary*, (TS).

The simplest class of model which captures permanent, i.e. persistent fluctuations is the random walk :

$$y_t = \rho y_{t-1} + e_t \qquad (3)$$

The random walk is non stationary in levels, but stationary when differenced, i.e. *difference stationary*, (DS) where in this case ρ=1.

It has become common practice to discriminate between DS and TS processes and hence *infer* persistence or otherwise by using the Dickey-Fuller unit root tests see Dickey and Fuller (1981). The usual form of the test treats DS as the null hypothesis and involves estimation of a regression like (4) below which presents the test in its 'augmented' form. Transforming (3) above, now the dependent variable is expressed as a first difference and α=(ρ-1), giving a null hypothesis (unit root) of Ho: α=0; H_1: α<0 ,

$$\Delta y_t = \alpha y_{t-1} + \sum_{i=1}^{p} \varphi_i \, \Delta y_{t-i} + \upsilon_t \qquad (4)$$

where υ_t is assumed to be serially uncorrelated. The original, (un)augmented, Dickey-Fuller test simply sets all the lags of the dependent variable equal to zero. Assuming drift adds an intercept, μ, to (4) above:

$$\Delta y_t = \mu + \alpha y_{t-1} + \sum_{i=1}^{p} \varphi_i \, \Delta y_{t-i} + \upsilon_t \qquad (4')$$

and the possibility of a deterministic trend gives:

$$\Delta y_t = \mu + \alpha y_{t-1} + \gamma t + \sum_{i=1}^{p} \varphi_i \, \Delta y_{t-i} + \upsilon_t \qquad (4'')$$

In all cases, however, the hypothesis of interest remains the DS null, H_0: α=0

As is now well known, under the null hypothesis the usual *t* ratios are distributed as Dickey-Fuller τ (1981) rather than Student's *t*. This is the approach pioneered by Nelson and Plosser (1982) and followed on numerous occasions (see Gaffeo, Gallegati and Gallegati (2005) for an excellent, up to date review).

3.2.1 The Dickey-Fuller test with Generalised Least Squares detrending, (ADF-GLS)

There have been many developments in unit root testing since Dickey and Fuller, including a range of Bayesian-based methods we are not going to consider here. Instead we will present some recent developments in testing for a unit root where, in particular, these new tests can be found in eg., EViews, STATA and RATS.

Elliott, Rothemburg and Stock (1996) use a modification of the ADF tests where the data are detrended so that explanatory variables are removed from the data prior to running the test regression. This ADF-GLS test is based upon the following regression:

$$\Delta y^{\tau}_{t} = \alpha y_{t-1} + \sum_{i=1}^{k} \varphi_i \Delta y^{\tau}_{t-i} + u_t \qquad (5)$$

where u_t is assumed *iid N(0, σ^2)* and Δy^{τ}_{t} is the locally detrended process. Under the null hypothesis, H_0: α=0, the ADF-GLS $^{\tau}$ is the *t*-ratio on α with rejection inferred if it is significantly less than zero when compared with the response surface estimates in Cheung and Lai (1995).

3.2.2 The Phillips-Perron (PP) Test

The size distortions in the original DF test due to serial correlation in the error term was the reason for the adoption of lagged dependent variables in its 'augmented' form of the DF test, the ADF. This augmented form of the DF remains the most common unit root test used in practice.

Phillips and Perron (1988) take a different approach to the potential effects of serial correlation – they use semi-parametric estimation of the long run effects of the short run dynamics. In particular, consider the auxiliary regression:

$$\Delta y_t = \mu + \alpha y_{t-1} + \gamma t + \eta_t \qquad (6)$$

where $\eta_t = \phi(L)\varepsilon_t,\ \varepsilon_t \sim iid\ N(0,1).$ If we define the residuals from the OLS regression of y_t on a constant and *t*, the Phillips and Perron test statistic can be defined as:

$$Z_t = \hat{\sigma}_{\xi}\,\hat{\omega}^{-1} t_{\alpha} - \hat{\lambda}\left[\hat{\omega}\left(T^{-2}\sum_{t=2}^{T}\hat{D}_{t-1}^{2}\right)^{1/2}\right]^{-1}$$

Where ω^2 and λ are nuisance parameters consistently estimated by applying the Newey-West (1987) estimator. Under the null hypothesis of a unit root, Z_t converges in the limit to the Dickey-Fuller distribution, although they may differ in finite samples. Empirically, however, Schwert (1989) showed that Z_t is biased towards rejection of the null if the error term is an MA(p) process with negative first order serial correlation. This is something to be considered in empirical applications.

3.2.3 The Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) Test

In contrast to the tests considered above, these authors assume that the null hypothesis of the univariate process is stationarity. The test is based on modelling the series as the sum of one stationary and one non-stationary component and testing the null hypothesis that the variance of the non-stationary component is zero. The trend stationary null is rejected when the KPSS statistic is larger than the approximate critical values tabulated in Kwiatkowski, Phillips, Schmidt, and Shin (1992). Given the null of stationarity, this test is often uses as part of a 'battery' of tests to consider the robustness of other (non-stationary null) test results.

3.2.4 Ng and Perron Tests

Ng and Perron (2001) develop four test statistics based upon the Dickey-Fuller GLS approach with detrended data. In particular, they consider two variants of Phillips and

Perron (1988), the Bhargava (1986) R_1 statistic, and the Elliott, Rothemberg and Stock, Point Optimal statistic. For further details and implementation, see eg., *Eviews* 7.

3.3 Unit roots with exogenous or endogenous change

Perron (1989) demonstrated how structural breaks in a series can lead to biased unit root test results (in favour of DS). He uses the idea of exogenously determined breaks informed by prior knowledge. Such exogenous assumptions have effects on the timing and properties of the critical values that are used to compare with the test results. Zivot and Andrews (1992) allow for endogenously determined breaks chosen on the basis of particular statistical criteria in an economically atheoretical way. Their critical values are likewise affected by the testing methods and in the original form the number of breaks permitted is limited.

Ignoring the presence of breaks, when they actually occur can lead to spurious non-rejection of the null hypothesis of a unit root. The main differences in the testing procedures proposed by Perron (1989), and subsequently Zivot and Andrews (1992), over the original Dickey-Fuller approach, involves the addition of various dummy variables to (4) to capture changes in the intercept and/or time trend and the use of recursive estimation methods, i.e.,

$$y_t = \mu + \rho y_{t-1} + \beta t + \gamma DT + \theta DU + \sum_{i=1}^{p} \varphi_i \Delta y_{t-i} + u_t \qquad (7)$$

where DU=1 if t>TB, 0 otherwise and DT = t if t>TB and 0 otherwise and TB refers to the time of the break.

In the original Perron formulation, the variables DU and DT were included to capture the possibility of "crashes" (DU) trend changes (DT) and joint crashes and trend changes (DU and DT). In the empirical section below we consider the possibility of

"jumps" where the coefficient attached to DU could be positive rather than negative as in the case of a "crash". Likewise a joint "jump" and trend change (denoted j&t), would involve the inclusion of both DU and DT in the equation. Tests of the null hypothesis of DS still involve $H_0: \rho = 1, \beta = 0$ although critical values are now given in Perron (1989) for exogenous breaks or Zivot and Andrews (1992) for endogenously located breaks.

The original motivation for the Perron (1989) approach and modifications by Zivot and Andrews (1992) was in response to the lack of power to reject the null of a unit root in the presence of structural change. However, Perron and Zivot and Andrews are only one-break models. If the DGP involves more than one break as might be expected in the very long time series often used in cliometrics, we are left with the same problem the original approach was attempting to remove - biased unit root test results (in favour of DS). Vogelsang (1997), presents results showing the loss of power that ensures when using a one-break model in a world of two breaks. Empirically, Ben David and Papell (1998) present evidence of more than one break and Lumsdaine and Papell (1997), discussed below, consider a generalisation of the endogenous break-point procedure of Zivot and Andrews (1992).

3.3.1 Lumsdaine-Papell test

The Lumsdaine and Papell (1997) test extends the Zivot and Andrews test equation (7) above, by adding in additional dummy variables for intercept and slope changes as shown below:

$$y_t = \mu + \rho y_{t\text{-}1} + \beta t + \gamma DT_1 + \theta DU_2 + \gamma DT_2 + \theta DU_1 + \sum_{i=1}^{p} \varphi_i \Delta y_{t\text{-}i} + u_t \quad (8)$$

As in the single break tests of Zivot and Andrews, three types of models can be considered, but now there are more variations including two breaks in the intercept; two breaks in the slope, etc. Being a variation of a standard unit root test, the *t* statistic on ρ is compared to the relevant critical value found in Lumsdaine and Papell (1997).

3.3.2 Lee-Strazicich (2001, 2003) tests

Lee and Strazicich (2001, 2003) take a similar approach to Lumsdaine and Papell (1997), but their test statistic uses a minimum Lagrange Multiplier, *LM*, test criteria. This approach is based upon the results from Schmidt and Phillips (1992) on the potential for unit root tests to report spurious rejections when the null includes a genuine structural break. The Lee and Strazicich *LM*-based test therefore starts with an assumption that the null hypothesis is a unit root with up to two breaks and as should be clear, the *t* statistic to test the null arises via a *LM* principle based upon the score. The ability to permit (up to) two breaks in the null and two breaks in the level or slope of the alternative, makes the approach particularly flexible and attractive. This test procedure was recently utilised by Greasley, Madsen and Wohar (2010) to consider the empirics of long run growth.

*3.4 Panel Unit Root Test*s

It is now well know that conventional univariate ADF unit root tests often tend to suffer from low power when applied to series of only moderate length. The idea of panel-based unit root testing is to pool the data across individual members of a panel to address this

issue by making available considerably more information regarding the series under investigation. Panel unit root ADF techniques are intended to allow researchers to selectively pool information regarding common long-run relationships from across the panel while allowing the associated short-run dynamics and fixed effects to be heterogeneous across different members of the panel see Maddala and Wu, (1999). We consider below three common forms of panel unit root tests, Levine et al. (2002, hereafter LLC), Im Peseran and Shin (2003, hereafter IPS) and Hadri (2000).

LLC propose an ADF test with a panel setting that restricts parameters γ_i by assuming them identical across cross-sections as follows:

$$\Delta y_{it} = \alpha_i + \gamma_i y_{it-1} + \sum_{j=1}^{k} \alpha_j \Delta y_{it-j} + e_{it} \qquad (9)$$

where $t = 1,2,\cdots,T$ refers to the time periods and $i = 1,2,\cdots,N$ refers the numbers in the panel. The null hypothesis of LLC test is $\gamma_i = \gamma = 0$ for all *i* indicating that the panel data are non-stationary while the alternative hypothesis is $\gamma_1 = \gamma_2 = \cdots = \gamma < 0$. This test is based on the statistics, $t_y = \hat{\gamma} / s.e.(\hat{\gamma})$.

It is clear that the null hypothesis of the LLC test is very restrictive and the test of IPS (2003) relaxes this assumption by allowing γ to vary across *i* under the alternative hypothesis. The null hypothesis of the IPS test is therefore that $\gamma_i = 0$ for all *i*, while the alternative hypothesis is $\gamma_i < 0$ for all *i*. The IPS test then uses the mean-group approach and obtains the average of t_y to compute the following statistic:

$$\tilde{Z} = \sqrt{N}(\bar{t} - E(\bar{t})) / \sqrt{\text{var}(\bar{t})}$$

where $\bar{t} = (1/N)\sum_{i=1}^{N} t_{y_i} E(\bar{t})$ and $Var(\bar{t})$ represents the mean and variance of each t_y, respectively. The statistic $\tilde{Z}$ converges to a *Normal* distribution, and we can compute the significance level in a simple way.

By contrast, Hadri (2000) argues that the null hypothesis should be reversed to be stationarity in order to produce a test with more power. His Lagrange Multiplier (*LM*) statistics is given by the follow expression:

$$LM = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{T^{-2}\sum_{t=1}^{T}\sum_{j=1}^{t}\hat{\varepsilon}_{ij}}{\hat{\sigma}_{\varepsilon}^{2}}\right) \qquad (10)$$

Where $\hat{\sigma}_{\varepsilon}^{2}$ is the consistent Newey-West (1987) estimate of the long-run variance of disturbance terms (ε_{ij}).

Panel-based tests, which are now available via for example Eviews 7 include; Levin, Lin and Chu (2002), Breitung (2000), Im, Pesaran and Shin (2003), Fisher-type tests using ADF and PP tests (Maddala and Wu (1999) and Choi (2001)), and Hadri (2000). A summary of some aspects of these tests is provided below. All use lags as corrections for autocorrelation, as with the ADF, and permit the option of fixed effects; individual effects and individual trends where relevant to the tests:

Levin, Lin and Chu (2002); H_0: unit root; H_1: no unit root.

Breitung (2000) ; H_0: unit root; H_1: no unit root.

Im, Pesaran and Smith (2003) ; H_0: unit root; H_1: Some cross-sections without a unit root.

Maddala and Wu, Fisher-ADF, (1999); H_0: unit root; H_1: Some cross-sections without a unit root.

Maddala and Wu, Fisher-PP (1999); H_0: unit root; H_1: Some cross-sections without a unit root.

Hadri (2000); H_0: No unit root; H_1: unit root.

3.5 Direct measures of persistence

One of the possible outcomes of testing for whether a series is a DS v. TS processes is to infer that that shocks implied by the process will have either infinite or zero persistence respectively. Campbell and Mankiw (1987) and Cochrane (1988), amongst others consider this as extreme and provide methods to measure the actual *persistence* of shocks, i.e. how much does a one-unit shock to output (say), affect forecasts into the future? Furthermore, Cochrane (1988) demonstrates how any DS process can be represented as the sum of a stationary and random walk component where the issue of persistence revolves around the *size of the random walk element*. In particular assume y is a linear DS process i.e.,

$$\Delta y_t = (1-L)y_t = \mu + A(L)\ \varepsilon_t = \mu + \sum_{j=0}^{\infty} a_j\ \varepsilon_{t-j} \qquad (11)$$

Utilising the Beveridge and Nelson (1981) decomposition, let

$$y_t = z_t + c_t \qquad (12)$$

where

$$z_t = \mu + z_{t-1} + \left(\sum_{j=o}^{\infty} a_j\right) \varepsilon_j$$

$$-c_t = \left(\sum_{j=1}^{\infty} a_j\right) \varepsilon_t + \left(\sum_{j=2}^{\infty} a_j\right) \varepsilon_{t-1} + \left(\sum_{j=3}^{\infty} a_j\right) \varepsilon_{t-2} + \ldots..$$

Here z_t is to be considered the permanent and c_t, the temporary component of y_t. Long-term forecasts of y_t are unaffected by c_t, the temporary component.

Cochrane (1988) considers the innovation variance of the random walk component as a natural measure of the importance of the random walk element . He

gives two equivalent formulations of this measure. The variance of the random walk element $\sigma^2_{\Delta z}$ is given by:

$$\sigma^2_{\Delta z} = \left(\sum_{j=o}^{\infty} a_j \right)^2 \sigma^2_{\varepsilon} = \left| A(1) \right|^2 \sigma^2_{\varepsilon}$$

Equivalently, $\sigma^2_{\Delta z}$ is equal to the spectral density of Δy_t at frequency zero, i.e.,

$$\sigma^2_{\Delta z} = \left(\sum_{j=o}^{\infty} a_j \right)^2 \sigma^2_{\varepsilon} = S_{\Delta y}\left(e^{-i0}\right)\sigma^2_{\varepsilon}$$

which can be estimated by the Bartlett estimator. However, as demonstrated by Cochrane (1988), the Bartlett estimator will be biased in small samples where the bias can be corrected by multiplying the estimates by T/(T-k), where T is the effective sample size and k the window size. One of the main advantages of the $\sigma^2_{\Delta z}$ measure over parsimonious ARMA representations, is that it captures all the effects of a unit root on the behaviour of a series in a finite sample. Notice that tests of TS in this framework involve a test of $\sigma^2_{\Delta z} = 0$. However, further note that the size of the random walk element is a continuous choice where series can be more fruitfully categorised by the size of $\sigma^2_{\Delta z}$ or the *persistence* of y.

However, the persistence measure is derived assuming an underlying linear model, (without discontinuities). If y_t were a pure random walk, the variance of the kth difference would grow linearly with k. If it were TS, however, the variance of the kth difference approaches a constant. Non-linearities in the underlying series would invalidate this relationship and the relevance of the Cochrane-type persistence measure.

In circumstances when breaks in the series might be expected the following approach can be adopted; firstly check the order of integration of the data using standard

or extended versions of the Augmented Dickey-Fuller test. If the series is TS without breaks measure the degree of persistence using the methods discussed above, however, if it is TS with breaks use the timing of the breaks, as a first approximation, to distinguish between persistence measure periods. If it is DS after testing for breaks, present persistence measures for the full and, where necessary sub-samples. Check the sensitivity of the results to specific sample periods and investigate the existence of significant non-linear elements. (see below section 4.2 and Greasley and Oxley (1997d, and 1998a, for an example)

3.6 Cointegration

In their classic paper, Engle and Granger (1987) show that a linear combination of two or more I(1) series may be stationary, or I(0). In this case, the series are said to be c*ointegrated.* The linear combination, if it exists, defines a cointegrating equation with the resulting cointegrating vector of weights characterizing the long-run relationship between the variables. Stated more formally:

> The components of the vector $x_t = (x_{1t}, x_{2t}, x_{3t}, ...x_{nt})'$ are said to be cointegrated of order *d,b* and denoted $x_t \sim CI(d,b)$ if the following two conditions hold: i) All elements of x_t are integrated of order d and ii) there exists a vector (the cointegrating vector) $\beta = (\beta_1, \beta_2, \beta_3, ...\beta_n)$ such that the linear combination $\beta x_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + ...\beta_n x_{nt}$ is integrated of a lower order (*d-b*) where *b*>0.

3.6.1 Single equation Engle-Granger (1987) 2 step methods (residual based tests)

This original form of testing for cointegration is effectively a test of the time series properties of the residuals in an (OLS) regression of the levels of the variables where the null hypothesis is of no-cointegration (no significant linear combination).

Consider the following:

$$y_t = \beta_0 + \beta_1 x_t + e_t \qquad (13)$$

Where *x* and *y* are variables of interest; the β are coefficients to be estimated and e_t a random disturbance term. If *x* and *y* are non-stationary I(1), they have no tendency to revert to the mean, long run level. However, if *x* and *y* are cointegrated, that is a linear combination of *x* and *y* are stationary, or I(0), then we can think of the relationship above as exhibiting a long run equilibrium.

The nature of the residual-based cointegraion tests is that we can rewrite (13) above as:

$$e_t = y_t - \beta_0 - \beta_1 x_t \qquad (14)$$

and if this linear combination of integrated variables is cointegrated, then e_t must be I(0) or stationary (this is another example of the 'balance' property discussed above). Engle-Granger *two-step methods*, therefore involve OLS regressions of equations like (13) above (*step-one*) followed by a test of the order of integration of the error term from that equation (e_t) as *step-two*. Because the residual is derived from the step-one process it has the property of a generated regressor (see Pagan 1984 and Oxley and McAleer 1993) and the critical values of the ADF test need to account for this property. Software packages like *Eviews* take account of this automatically, otherwise incorrect inference could result.

Variations of the *residual-based* approach include Phillips and Ouliaris (1990) residual-based tests, Hansen (1992), and Park (1992).

If cointegration exists, and it should be stressed that not all integrated variables are co-integrated, see Granger (1986), not only is OLS an appropriate method to use it also has the property of *super-consistency* such that OLS estimates converge to the true value at the rate *n* compared to the usual rate √n where n is the sample size.

If we consider a special case of (14) above where e_t=0, then effectively the relationship estimated would be in long run equilibrium. However, typically $e_t \neq 0$ and the term represents the *equilibrium error* often referred to as the *error-correction term*. Note, that this is captured as the second stage of the cointegration testing procedure. The error-correction term plays an important role in the *Error (or Equilibrium) Correction Models (ECM)* which involves estimation of the short-run dynamics consistent with the long-run equilibrium captured by the cointegrating relationship. If cointegration exists between the I(1), integrated variables, the error-correction term (e_t) must be I(0). To be a valid ECM model of the short-run dynamics, therefore all the other variables in an ECM must be stationary. As the long run model involves I(1) in levels variables, the ECM will typically include variables in first-differences for example:

$$\Delta y_t = \alpha + \beta_1 \Delta x_t + \beta_2 \Delta x_{t-1} + \beta_3 \Delta z_{t-2} + \beta_4 e_{t-1} + \zeta_t \qquad (15)$$

where equation (15) may have been derived as part of a *General-to-Specific* search process yielding a relationship with an error term ζ_t with appropriate properties. Note e_t enters with a lag and an expectation that $\beta_4 < 0$. As e_{t-1} has been derived from the long run cointegrating model, it effectively links the long run and short run aspects of the

relationship. For discussions on alternative ways of constructing and interpreting the ECM see Muscatelli and Hurn (1992).

3.6.2 Johansen maximum likelihood estimation-based cointegration methods

As can be seen from the discussion above, Engle-Granger type methods for establishing the existence (or otherwise) of cointegration are simple, based upon an OLS regression of current valued variables, and powerful, having the property of superconsistency should cointegration be established. However, the EG methods have a number of shortcomings. The first is that the normalisation can matter. Normalisation here effectively means 'which variable is on the left hand side'. Consider our relationship described by equation (13) above.

As written we have chosen to assume y is the variable explained by x. If x and y were I(1), but cointegrated, it shouldn't matter whether we estimate $y=f(x)$ or $x=f(y)$ as the test for cointegration is based upon the e_t. This is the case asymptotically, but not necessarily in small samples. If we extend the relationship to include other variables the potential problem is expanded. Furthermore, the EG approach can only identify (up to) one cointegrating relationship. This isn't a problem in the bivariate case, but again once we go beyond two variables the potential to identify more than one significant cointegrating relationship is possible. Whether this means there are more than one 'economic' equilibrium is a mute point and requires careful consideration of the 'economic sense' of the linear combination(s) identified by the testing procedure.

In contract to Engle-Granger type methods, the Johansen (1988, 1995) approach utilizes a multivariate model where the sensitivity to normalisation problem outlined

above, disappears. At this stage it is also worth stressing that cointegration identifies linear combinations of variables integrated of a lower order. If we have *two* variables this implies at most, *one* (unique) cointegrating relationship; if we have *three* variables at most *two* cointegrating relationships; *four* variables at most *three* cointegrating relationships; *k* variables at most *k-1* cointegrating relationships, etc. If we establish via testing, *k significant* cointegrating relationships from *k* variables then it says <u>*any*</u> linear combination of the *k* variables is stationary which implies all the *k* variables are in fact stationary in levels (not non-stationary as thought or suggested by univariate pre-tests)! Of course, it should be stressed again that there may be no significant cointegrating relationships at all and in this case there is no tendency for the variables in the model to move together over time as the equilibrium interpretation of the mathematical notion of cointegration would imply.

The Johansen approach utilises this rank property as the basis for its tests of the existence of cointegration, in particular consider a multivariate version of the ADF equation:

$$x_t = A_1 x_{t-1} + A_2 x_{t-2} + A_3 x_{t-3} + ... A_p x_{t-p} + \upsilon_t \qquad (16)$$

Defining this in vector form so that $x_t = (x_{1t}, x_{2t}, x_{3t}, ... x_{nt})'$ we can rewrite it as

$$\Delta x_t = \pi x_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta x_{t-i} + \upsilon_t \qquad (17)$$

where $\pi = -(1 - \sum_{i=1}^{p} A_i)$ and $\pi_i = -\sum_{j=i+1}^{p} A_j$.

Here we can now see how the rank of π is crucial in determining the number of cointegating relationships and the issue of potential neglected stationarity . If the rank of π=o then no significant cointegrating relationship exists. If the rank of π=n, then any linear combination is stationary. If the rank=1, a single cointegrating relationship exists and πx_{t-1} represents the error correction term, whereas if the rank of π>0, <n, then multiple cointegrating relationships exist.

Testing for the number of significant cointegrating relationships using the Johansen approach involves checking the significance of the characteristic roots of π. Using the property that the rank of a matrix is equal to the number of characteristic roots that differ from zero, Johansen proposes two tests for cointegration; one based upon the trace and one on the maximum eigenvalue. In particular, if we denote the ordered *n* characteristic roots of the matrix π are denoted $\lambda_1 > \lambda_2 > \lambda_3 > ....\lambda n$ the trace test is defined as:

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{n} \ln(1-\hat{\lambda}_i) \qquad (18)$$

and the maximum eigenvalue as:

$$\lambda_{\max}(r, r+1) = -T(\ln(1-\hat{\lambda}_{r+1}) \qquad (19)$$

The null hypotheses of the two differ in that the *trace* tests whether the number of distinct eigenvalues is ≤r; whereas the maximum eigenvalue tests the null of r cointegrating relationships against r+1. In a practical setting, one may find that the trace test results may find 'more evidence' of a (or several) significant cointegrating relationships than the maximum eigenvalue test. Therefore in analysing and assessing

empirical research, one should be careful to consider which test(s) have been reported and whether (if both trace and eigenvalue are reported) they support or contradict.

3.6.3 Panel-based cointegration methods

Pedroni (1999, 2004) develops a number of *panel-based tests* for cointegration. Pedroni (1999) allows for cross-sectional interdependence with different individual effects. If the panel data follow an I(1) process, the Pedroni (1999 and 2004) panel cointegration model can be applied to ascertain whether a cointegration relationship exists. Pedroni (1999) suggests the following time series panel expression:

$$y_{it} = \alpha_i + \gamma_i t + X_{it}\beta_{it} + e_{it} \qquad (20)$$

Where y_{it} and X_{it} are the observable variables with dimension of $(N * T) \times 1$ and $(N * T) \times m$, respectively. He develops the asymptotic and finite-sample properties of the test statistics to examine the null hypothesis of *non-cointegration* in a panel. The tests allow for heterogeneity among individual members of the panel, including heterogeneity in both the long-run cointegration vectors and in their dynamics.

Pedroni develops two types of *residual-based* tests. For his first type, four tests are distributed as standard Normal asymptotically and are based on pooling the residuals of the regression for the within-group. The four tests in this class are:; the panel υ-statistic; the panel ρ-statistic; the panel PP-statistic (or t-statistic, non-parametric) and the panel ADF-statistic (or t-statistic, parametric). For the group of second type tests, three are also distributed as standard Normal asymptotically, but are based on pooling the residuals for the between-group. The three tests in this class are: the group ρ-statistic; the group PP-statistic (or t-statistic, non-parametric) and the group ADF-statistic (or t-

statistic, parametric). Pedroni (1999) presents the following heterogeneous panel cointegration statistics:

$$\text{Panel } \upsilon\text{-statistic: } Z_{\hat{\upsilon}} = (\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{L}_{11i}^{-2}\hat{e}_{it-1}^{2}) \qquad (21)$$

$$\text{Panel } \rho\text{-statistic: } Z_{\hat{\rho}} = (\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{L}_{11i}^{-2}\hat{e}_{it-1}^{2})^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{L}_{11i}^{-2}(\hat{e}_{it-1}\Delta\hat{e}_{it} - \hat{\lambda}_i) \qquad (22)$$

$$\text{Panel } t\text{-statistic (non-parametric): } Z_t = (\hat{\sigma}^2\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{L}_{11i}^{-2}\hat{e}_{it-1}^{2})^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{L}_{11i}^{-2}(\hat{e}_{it-1}\Delta\hat{e}_{it} - \hat{\lambda}_i) \qquad (23)$$

$$\text{Panel } t\text{-statistic (parametric): } Z_t^* = (\hat{s}^{*2}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{L}_{11i}^{-2}\hat{e}^{*2}{}_{it-1})^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{L}_{11i}^{-2}\hat{e}^{*}{}_{it-1}\Delta\hat{e}^{*}{}_{it} \qquad (24)$$

And the following heterogeneous group-mean panel cointegration statistics:

$$\text{Group } \rho\text{-statistic: } \tilde{Z}_{\hat{\rho}} = \sum_{i=1}^{N}(\sum_{t=1}^{T}\hat{e}_{it-1}^{2})^{-1}\sum_{t=1}^{T}(\hat{e}_{it-1}\Delta\hat{e}_{it} - \hat{\lambda}_i) \qquad (25)$$

$$\text{Group } t\text{-statistic (non-parametric): } \tilde{Z}_t = \sum_{i=1}^{N}(\hat{\sigma}_i^2\sum_{t=1}^{T}\hat{e}_{it-1}^{2})^{-1/2}\sum_{t=1}^{T}(\hat{e}_{it-1}\Delta\hat{e}_{it} - \hat{\lambda}_i) \qquad (26)$$

$$\text{Group } t\text{-statistic (parametric): } \tilde{Z}_t^* = \sum_{i=1}^{N}(\sum_{t=1}^{T}\hat{s}_i^2\hat{e}_{it-1}^{*2})^{-1/2}\sum_{t=1}^{T}(\hat{e}_{it-1}^{*}\Delta\hat{e}_{it}^{*}) \qquad (27)$$

Where $\hat{e}_{it}$ is the estimated residual from equation (20) above and $\hat{L}_{11i}^{-2}$ is the estimated long-run covariance matrix for $\hat{e}_{it}$. Similarly, $\hat{\sigma}_i^2$ and $\hat{s}_i^2$ ($\hat{s}_i^{*2}$) are, respectively, the long-run and contemporaneous variances for individual *i*. The other terms are defined in Pedroni (1999) with the appropriate lag length determined by the Newey-West method. All seven tests are distributed as standard Normal asymptotically. This requires standardization based on the moments of the underlying Brownian motion function. The panel υ-statistic is a one-sided test where large positive values reject the null of no cointegration. The remaining statistics diverge to negative infinitely, which means that

large negative values imply rejection of the null. The critical values are also tabulated in Pedroni (1999).

The statistics above are based on estimators that simply average the individually estimated coefficients for each member, and each of these tests is able to accommodate individual specific short-run dynamics, individual specific fixed effects and deterministic trends, as well as individual specific slope coefficients (Pedroni, 2004). The number of observations available is greatly increased in a panel framework when testing the stationarity of the residual series in a levels regression and this can substantially increases the power of the cointegration tests see, Rapach and Wohar, (2004).

3.7 Granger-type causality

3.7.1 Conventional Granger-type tests

Many tests of causality have been derived and implemented, including Granger (1969, 1988), Sims (1972), and Geweke *et al.* (1983a). We are not going to debate here whether Granger causality tests 'causality', but note that Granger-type causality tests are tests of temporal ordering. Granger showed via his *Representation Theorem* that a bivariate co-integrated system must have causal ordering in at least one direction. These inherent links between Granger-causality and co-integration have been exploited to formulate the current suite of tests for causality used in time series econometrics. The tests are all based upon the estimation of autoregressive or vector autoregressive (VAR), models involving (say), the variables *X* and *Y*, together with significance tests for sub-sets of the variables.

Although it is quite common to test for the direction of causality, the conclusions drawn in some studies are fragile for two important reasons. Firstly, the choice of lag lengths in the autoregressive or VAR models is often *ad hoc*, see for example, Jung and Marshall (1985), Chow (1987), and Hsiao (1987), although the length of lag chosen will

critically affect results. Secondly, in the absence of evidence on cointegration, "spurious" causality may be identified.

Engle and Granger (1987), show that if two series are individually I(1), and cointegrated, a causal relationship must exist in at least one direction. Furthermore, the Granger Representation Theorem demonstrates how to model cointegrated I(1) series in the form of a VAR model. In particular, the VAR can be constructed either in terms of the levels of the data, the I(1) variables; or in terms of their first-differences, the I(0) variables, with the addition of an error-correction term (ECM) to capture the short-run dynamics. If the data are I(1), but not cointegrated, causality tests cannot validly be derived unless the data are transformed to induce stationarity which will typically involve tests of hypotheses relating to the growth or first-difference of variables (if they are defined in logarithms), and not their levels. To summarise, causality tests can be constructed in three ways, two of which require the presence of cointegration. The three different approaches are defined below.

The *first stage* in testing for causality involves testing for the order of integration. Conditional on the outcome of such tests, the *second stage* involves investigating bivariate cointegration utilising the Johansen maximum likelihood approach. If bivariate cointegration exists then either uni-directional or bi-directional Granger causality must also exist, although in finite samples there is no guarantee that the tests will identify it. On the basis of the bivariate cointegration results, a multivariate model of cointegration may then be investigated to examine interaction effects, taking the error term from this cointegrating regression as a measure of the ECM term to capture the short run dynamics of the model. The *third stage* (or second if bivariate cointegration is rejected), involves

constructing standard Granger-type causality tests, augmented where appropriate with a lagged error-correction term, see Giles *et al.* (1993).

The *three-stage procedure* leads to *three* alternative approaches for testing causality. In the case of cointegrated data, Granger causality tests may use the I(1) data because of the superconsistency properties of estimation. With two variables *X* and *Y*:

$$X = \alpha + \sum_{i=1}^{m} \beta i \; X_{t-i} + \sum_{j=1}^{n} \gamma_j Y_{t-j} + u_t \cdots\cdots\cdots\cdots (28)$$

$$Y = a + \sum_{i=1}^{q} bi Y_{t-i} \sum_{j=1}^{r} c_j X_{t-j} + \nu_t \cdots\cdots\cdots\cdots\cdots (29)$$

where u_t and ν_t are zero-mean, serially uncorrelated, random disturbances and the lag lengths m,n,q and r are assigned on the basis of minimising some form of Information Criteria.

Secondly Granger causality tests with cointegrated variables may utilise the I(0) data, including an error-correction mechanism term, i.e.,

$$\Delta X = \alpha + \sum_{i=1}^{m} \beta i \; \Delta X_{t-i} + \sum_{j=1}^{n} \gamma_j \Delta Y_{t-j} + \delta ECM_{t-1} + u_t \cdots\cdots\cdots\cdots (28')$$

$$\Delta Y = a + \sum_{i=1}^{q} bi \Delta Y_{t-i} + \sum_{j=1}^{r} c_j \Delta X_{t-j} + dECM_{t-1} + \nu_t \cdots\cdots\cdots\cdots (29')$$

Where the error-correction term, derived from the cointegrating relationship, is denoted ECM.

Thirdly if the data are I(1), but not cointegrated valid Granger-type tests require transformations to induce stationarity. In this case the tests deploy formulations like (28') and (29') above, but without the ECM term, i.e., (28") and (29") below.

$$\Delta X = \alpha + \sum_{i=1}^{m} \beta i \ \Delta X_{t-i} + \sum_{j=1}^{n} \gamma_j \Delta Y_{t-j} + u_t \cdots\cdots\cdots\cdots\cdots (28")$$

$$\Delta Y = a + \sum_{i=1}^{q} bi \Delta Y_{t-i} + \sum_{j=1}^{r} cj \Delta X_{t-j} + \nu_t \cdots\cdots\cdots\cdots\cdots (29")$$

Granger causality tests based upon equations (28) and (29) involve the following:

Y Granger causes (GC), X if, $H_0: \gamma_1 = \gamma_2 = \gamma_3 = \ldots\ldots = \gamma_n = 0$ is rejected against the alternative $H_1:=$ at least one $\gamma_j \neq 0, \quad j = 1,\ldots,n$.

X GC Y if, $H_0: c_1 = c_2 = c_3 = \ldots\ldots = c_r = 0$ is rejected against the alternative $H_1:=$ at least one $c_j \neq 0, \quad j = 1,\ldots,r$. .

For equations (28') and (29') Granger causality tests involve the following:

ΔY Granger causes (GC), ΔX if, $H_0: \gamma_1 = \gamma_2 = \gamma_3 = \ldots\ldots = \gamma_n = 0$ is rejected against the alternative , $H_1:=$ at least one $\gamma_j \neq 0, \quad j = 1,\ldots,n$, or $\delta \neq 0$, (see Granger 1986).

ΔX GC ΔY if, $H_0: c_1 = c_2 = c_3 = \ldots\ldots = c_r = 0$ is rejected against the alternative $H_1:=$ at least one $c_j \neq 0, \quad j = 1,\ldots,r.$, or $d \neq 0$, (see Granger 1986).

Notice in this case however, with the possibility of causality being inferred from the significance of *d* or δ alone that the causal nexus is altered i.e., causality runs from the past *level* to the current rate of *change* without any lagged change effects.

For *non-cointegrated* data (*X* and *Y*, I(1)), Granger causality tests involve tests based upon equations (28") and (29"), in particular:

ΔY Granger causes (GC), ΔX if, $H_0: \gamma_1 = \gamma_2 = \gamma_3 = ... = \gamma_n = 0$ is rejected against the alternative, $H_1:=$ at least one $\gamma_j \neq 0, \quad j = 1,...,n$.

ΔX GC ΔY if, $H_0: c_1 = c_2 = c_3 = ... = c_r = 0$ is rejected against the alternative. $H_1:=$ at least one $c_j \neq 0, \quad j = 1,...,r$.

Given the inclusion of lagged dependent variables in (28) and (29), (28') and (29') and (28") and (29"), tests of the hypotheses utilising OLS results require the modified Wald statistics, nF_1 and rF_2, distributed (asymptotically) as χ^2 with n and r degrees of freedom, where F_1 and F_2 are the "normal" F statistics of the joint significance of the γ's and c's respectively. Furthermore, in the case of equations (28) and (29) we invoke the results of Lutkepohl and Leimers (1992), and Toda and Phillips (1991), which show that in bivariate non-stationary cointegrated models the Wald test will have the usual asymptotic χ^2 distribution.

In addition to the Wald test of zero restrictions, and '*t*' tests on *d* and δ where appropriate, the Final Prediction Error (FPE) can be used as an additional *indication* of causality, i.e. if FPE(m*,n*)<FPE(m*), it implies *Y Granger-causes X* (or ΔY Granger causes ΔX where appropriate), likewise for r* and q*, see Giles *et al.* (1993) for more details. All three criteria are used in the empirical section of the paper.

3.7.2 Toda and Phillips (1991) - type tests

Under conditions of cointegration, the ECM based tests discussed above involve some form of two-step process, i.e., test for cointegration and retain the residuals as the ECM term and utilise this variable in the second stage either as a direct test of causality following Granger (1986) and Engle and Granger (1987), or as part of the modelling strategy when testing the significance of the VAR terms.

In Toda and Phillips (1991), - henceforth TP - the authors consider a different, single-stage estimation, but (potentially), sequential testing, framework as well as a critical review of previous tests.

Consider the n-vector time series y_t generated by the k-th order VAR:

$$y_t = J(L)y_{t-1} + u_t \qquad t = -k+1,....,T \tag{30}$$

where L is the lag operator defined as, $J(L) = \sum_{i=1}^{k} J_i L^{i-1}$ and u_t an n dimensional random vector. Making sufficient assumptions to ensure that y_t is cointegrated, CI(1,1) with r cointegrating vectors (r≥1), see TP for details, rewrite (30) in the equivalent ECM form:

$$\Delta y_t = J^*(L)\Delta y_{t-1} + \Gamma A' y_{t-1} + u_t \tag{31}$$

where J^*(L) is defined analogous to the expression above.

3.7.2.1 Causality Tests

Following Sims, Stock and Watson (1990), consider a test of whether the last n_3 elements of y_t cause the first n_1 elements of the vector where y_t is partitioned as:

$$y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix} \begin{matrix} n_1 \\ n_2 \\ n_3 \end{matrix} \tag{32}$$

1. Levels VAR

The null hypothesis of non-causality based upon equation (32) would be:

$$H \ : \ J_{1,13} = ... = J_{k,13} = 0 \tag{33}$$

and $J_{13} = \sum_{i=1}^{k} J_{i,13} L^{i-1}$ is the n1 x n3 upper-right submatrix of J(L). Denoting A_3 as the last n3 rows of the matrix of cointegrating vectors A, if rank(A_3) = n3 , then via TP, Corollary 1, under the null hypothesis from (33):

$$F \xrightarrow{d} \chi^2_{n1n3k}$$

However, the rank condition on the sub-matrix (A_3), based upon OLS estimates, suffers from simultaneous equation bias, such that there is no valid statistical basis for determining whether the required sufficient condition applies. When the condition fails, the limit distribution is more complex than that shown above and involves a mixture of a χ^2 and a non-standard distribution and generally involves nuisance parameters.

2. Johansen-type ECM's

Based now upon equation (34), the null hypothesis of non-causality becomes:

$$H^* : J^*_{1,13} = ... = J^*_{k-1,13} = 0 \text{ and } \Gamma_1 A'_3 = 0 \qquad (34)$$

and $J_{13} = \sum_{i=1}^{k} J_{i,13} L^{i-1}$ is the n1 x n3 upper-right submatrix of J*(L), and Γ_1 are the first n1 rows of the loading coefficient matrix Γ. If rank Γ_1 = n1 *or* rank(A_3) = n3 , then under the null hypothesis (34)

$$F^* \xrightarrow{d} \chi^2_{n1n3k}$$

Again, if neither of these conditions are satisfied, causality tests based upon χ^2 will not in general be valid. However, unlike the case above, tests of such conditions are relatively easy to construct and constitute the *sequential testing strategy* of TP.

Consider for the moment either n1=1 or n3=1, (or n1=1 and n3=1), such that Γ_1 is a scalar denoted γ_1 as is A_3 denoted α_3, then define the following null hypotheses:

$$H^*: J^*_{1,13} = ... = J^*_{k-1,13} = 0 \text{ and } \gamma_1 \alpha'_3 = 0$$
$$\boldsymbol{H^*_{\uparrow} : J^*_{1,13} = ... = J^*_{k-1,13} = 0}$$

$$H^*_1: \gamma_1 = 0$$
$$H^*_3: \alpha_3 = 0$$
$$H^*_{13}: \gamma_1 \alpha'_3 = 0$$

The TP *sequential testing strategy* involves :

$$(P1) \text{ Test } H^*_1: \begin{cases} \textit{If } H^*_1 \textit{is} \text{ rejected test } H^* \\ \text{Otherwise test } H^*_{\uparrow} \end{cases}$$

$$(P2) \text{ Test } H^*_3: \begin{cases} \textit{If } H^*_3 \textit{is} \text{ rejected test } H^* \\ \text{Otherwise test } H^*_{\uparrow} \end{cases}$$

and when n1=n3=1:

$$(P3) \text{ Test } H^*_{\uparrow} \begin{cases} \textit{If } H^*_{\uparrow} \text{ is rejected, reject the null hypothesis of noncausality} \\ \text{Otherwise, test } H^*_1 \text{ and } H^*_3 \begin{cases} \text{If both } H^*_1 \text{ and } H^*_3 \text{ are rejected test } H^*_{13} \text{ if } \hat{r} > 1 \\ \text{or reject the null if } \hat{r} = 1 \\ \text{Otherwise, accept the null of non causality} \end{cases} \end{cases}$$

(where $\hat{r}$ is an estimate of r).

Having established this *theoretical hierarchy of testing*, based upon their Monte Carlo results, TP make the following observations/recommendations:

> *1. (P1) generally performs better than (P2) and should be preferred over (P2)*
>
> *2. When n1=n3=1, (P1) and (P2) are less vulnerable to size distortions than (P3) which should be avoided*
>
> *3. None of the sequential procedures (or conventional tests), performed well for sample sizes below 100, at least with systems of three or more variables*
>
> *4. The sequential tests outperform the conventional VAR tests which suffer considerable size distortions where tests are not valid asymptotically* χ^2.

Furthermore, consideration of their Monte Carlo results reveals that for many cases considered, "our testing procedures do not have much power unless the lag length *k* is specified correctly. This is not surprising because if *k>1* the coefficients of the lagged differences of y_3 are all zero." For other cases, "If we choose 22% critical values for those sub tests (H_1^*, H_3^*), then we would have approximately 5% significance level for the overall causality test ..... but of course we cannot do so without allowing large upward size distortions in other cases...." - TP. More generally, under many plausible cases it seems that the sequential procedures involve the potential to introduce large size distortions for relatively small deviations from assumed theoretical values, i.e., lag length, coefficient values and properties of the error term.

3.7.3 Toda and Yamamoto (1995)

The Toda and Yamamoto (1995) method can also be utilised to ascertain the direction of causality and involves using the levels of the variables irrespective of their order of

integration. The test involves adding additional lags based upon the potential order of integration ie., one additional lag if one assumes the data is I(1), two if I(2), etc. It has the advantage that it can be used when the order of integration is ambiguous or uncertain, however, the cost is in terms of efficiency.

3.8 The Structural Time Model (STM) approach of Harvey (1989)

In a series of papers, Andrew Harvey has argued that *structural time-series models* provide the most useful framework within which to consider 'stylised facts' about time series data. Nested within the general models he proposed, one can consider tests of TS v DS; the Hodrick-Prescott (1977) filter (and many other more general filters); etc. In a particularly useful paper in this respect, Harvey and Jaeger (1993) examine some of the consequences of the 'mechanical detrending methods of Hodrick and Prescott' and show that their uncritical use can lead to potentially spurious cycle detection. Instead of H-P, they propose the structural time series approach.

Below we will present the basics of the approach – those interested could usefully consult Harvey (1989).

3.8.1 The trend plus cycle model

Consider the following representation:

$$y_t = \mu_t + \psi_t + \varepsilon_t, \qquad t = 1,...T \tag{35}$$

where y_t is the series of interest; μ_t is the trend; ψ_t is the cycle and ε_t is the irregular component. The local linear trend model is defined as:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \qquad \eta_t \sim NID(0, \sigma_\eta^2) \tag{36}$$

$$\beta_t = \beta_{t-1} + \zeta_t \qquad \zeta_t \sim NID(0, \sigma_\zeta^2) \tag{37}$$

where β_t is the slope and independence is assumed between the white noise errors ζ_t and η_t.

The stochastic cycle is modelled as:

$$\begin{aligned} \psi_t &= \rho \cos \lambda_c \psi_{t-1} + \rho \sin \lambda_c \psi^*_{t-1} + \kappa_t \\ \psi^*_t &= -\rho \sin \lambda_c \psi_{t-1} + \rho \cos \lambda_c \psi^*_{t-1} + \kappa^*_t \end{aligned} \tag{38}$$

where ρ represents the damping factor, $0 \le \rho \le 1$, and λ_c is the frequency of the cycle in radians and κ_t and κ^*_t are both $NID(0, \sigma_\kappa^2)$. The disturbances in all three equations are treated as independent and the irregular component is also assumed to be $NID(0, \sigma_\varepsilon^2)$. Estimation of the hyperparameters, $(\sigma_\eta^2, \sigma_\zeta^2, \sigma_\kappa^2, \sigma_\varepsilon^2, \rho, \lambda_c)$ can be undertaken by maximum likelihood methods and separate estimates of the trend, cycle(s) and irregular components obtained. *STAMP* (Structural Time Series Analysis Modeller and Predictor) 8.3 is a powerful, flexible and very user friendly, Windows 'drop-down' menu-based software package written specifically for estimation and prediction of such models and methods.

This structural time series modelling approach is also very flexible as it nests within its general structure a range of special cases. For example:

i) The trend is a *random walk with drift and the cycle is an autoregressive component AR2* with the irregular either white noise or zero. In this case the level is 'stochastic' and the slope 'fixed'.

ii) The model is assumed to be a *local linear trend*. Here the level is 'stochastic' as is the slope. A stationary trigonometric cycle is included as is an irregular.

iii) Assume a *smooth trend* with level 'fixed' and slope 'stochastic'. The cycle may be a generalised version of the simple case considered above as in Harvey and Trimbur (2003).

Consider some useful properties of for example, the *local level* model:

$$y_t = \mu_t + \varepsilon_t, \qquad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2), \quad t = 1, \ldots T$$
$$\mu_t = \mu_{t-1} + \eta_t, \qquad \eta_t \sim NID(0, \sigma_\eta^2) \qquad (39)$$

Maximum likelihood estimates of σ^2_η and σ^2_ε and their relative variance $\sigma^2_\eta / \sigma^2_\varepsilon$ allow the calculation of for example, the 'signal to noise' ratio. *STAMP* can also create an H-P filter for example, assume a *local linear trend* (see equation 38 above) with the level 'fixed' the slope 'stochastic' and include an irregular component. Fix $\sigma_\eta = 0$, $\sigma_\xi = 0.000625 = 1/1600 = \lambda^{-1}$ will create the classic Hodrick Prescott filter calibrated for quarterly (US) data.

4.0 Empirical applications

4.1 Overview

The use of modern time series-based methods in cliometrics research is growing. Tables 1 and 2 below collate and summarise the methods used in empirical, time-series based papers published in the two main cliometrics journals, *Explorations in Economic History* (2000-2009) and *Cliometrica*, 2007-2009. The creation of *Cliometrica* itself is testimony

to the growing number of quantitative economic history papers being written. Furthermore, it is clear from the Tables, that there are a growing number of papers using modern time series methods in cliometrics more generally.

Tables 1 and 2 near here

The most common methods applied in the papers identified are simple unit root tests, typically of a Dickey-Fuller or Augmented Dickey-Fuller form. Cointegration, mainly Johansen-based and Granger causality are also commonly undertaken methods. Simple Vector- AutoRegressive (VAR) (see Sims, 1980) methods are also used by several authors. Testing for structural breaks; Kalman filter estimation; Auto Regressive Distributed Lag (ARDL) (see Pesaran, Shin and Smith, 2001); Generalised Auto Regressive Conditional Heteroskedasticity GARCH, (see Bollerslev, 1986); TAR (Threshold AutoRegression) and STAR (see Tong 1990, and Terasvirta, 1998); are less commonly used methods reflecting the particular applications under investigation. Overall, however, it is clear that many in the cliometrics group have adopted appropriate methods to consider the time-series questions they have chosen to consider, however, it is also clear from reading some applications by other authors, that more people need to consider their data, methods used and hence conclusions more carefully. This is not unique to cliometrics, but we will not take this point further here.

Table 1
Papers using Time Series methods: *Explorations in Economic History*, 2000-2009.

YEAR	AUTHOR(S)	TITLE	METHODS
2000	Greasley and Oxley	British Industrialization, 1815–1860: A Disaggregate Time-Series Perspective	Unit root tests//Kalman filter/ Stochastic trends
2000	Weidenmier	The Market for Confederate Cotton Bonds	VAR/unit roots/Granger causality
2000	Lew	The Diffusion of Tractors on the Canadian Prairies: The Threshold Model and the Problem of Uncertainty	Unit roots
2000	Keay	Scapegoats or Responsive Entrepreneurs: Canadian Manufacturers, 1907–1990	Unit roots/ trends
2000	Baten and Murray	Heights of Men and Women in 19th-Century Bavaria: Economic, Nutritional, and Disease Influences	Unit root tests
2000	Yousef	The Political Economy of Interwar Egyptian Cotton Policy	ARDL/ unit roots
2001	Fratianni and Spinelli	Fiscal Dominance and Money Growth in Italy: The Long record	Unit root/Granger /VAR
2001	Greasley, Madsen and Oxley	Income Uncertainty and Consumer Spending during the Great Depression	GARCH
2001	Allen and Keay	The First Great Whale Extinction: The End of the Bowhead Whale in the Eatern Arctic	AR/Structural break/trends
2002	Temin	Price Behavior in Ancient Babylon	Random walk
2002	Waldenstrom	Taxing Emerging Stock Markets: A Beneficial Policy? Evidence from the Stockholm Stock Exchange, 1907–1939	ADF/ Johansen-cointegration
2002	Goodwin, Grennes and Craig	Mechanical Refrigeration and the Integration of Perishable Commodity Markets	Unit root tests/ Engle-Granger regression/cointegration
2002	della Paolera and Taylor	Internal versus external convertability and emerging market crises: lessons from Argentine history	Unit Root tests/ VAR
2003	Antras and Voth	Factor prices and productivity growth during the British Industrial Revolution	ARMA/Crafts and Harley
2003	Tattara	Paper money but a gold debt: Italy on the gold standard	Cointegration/ unit root/ Johansen

2003	Toniolo, Conte and Vecchi	Monetary Union, institutions and financial market integration: Italy, 1862–1905	ARMA/ Kalman Filter/ Structural break
2003	Grant	Globalisation versus de-coupling: German emigration and the evolution of the Atlantic labour market 1870–1913	Cointergration/Unit root tests
2004	Greasley and Oxley	Globalization and real wages in New Zealand 1873–1913	Unit root tests/Cointegration/ADF
2004	Crafts and Mills	Was 19th century British growth steam-powered?: the climacteric revisited	AR(2)/Johansen/Cointegrating Vectors
2004	Grubb	The circulating medium of exchange in colonial Pennsylvania, 1729–1775: new estimates of monetary composition, performance, and economic growth	AR(2)/Dickey Fuller Tests
2004	Keay and Redish	The micro-economic effects of financial market structure: evidence from 20th century North American steel firms	Structural break/AR(3)
2005	Rousseau and Sylla	Emerging financial markets and early US growth	Granger causality/VAR-Sims/Dickey Fuller/Johansen
2005	Klug, Landon-Laneand White	How could everyone have been so wrong? Forecasting the Great Depression with the railroads	ARIMA
2005	Officer	The quantity theory in New England, 1703–1749: new data to analyze an old question	AR(1) /Dickey Fuller Test
2005	Lazaretou	The drachma, foreign creditors, and the international monetary system: tales of a currency during the 19th and the early 20th centuries	ARCH(1)/ARCH(2)
2005	Temin and Voth	Credit rationing and crowding out during the Industrial Revolution: evidence from Hoare's Bank, 1702–1862	AR(1)/VAR
2005	Sanz-Villarroya	The convergence process of Argentina with Australia and Canada: 1875–2000	Structural Break/AR(1)
2005	Hickson, Turner and McCann	*Much ado about nothing*: the limitation of liability and the market for 19th century Irish bank stock	Structural Break/ARIMA/AR(1)
2006	Van Nieuwerburgh, Buelens and Cuyvers	Stock market development and economic growth in Belgium	Cointegration/Johansen/Granger causality/Dickey-Fuller

2006	Burhop	Did banks cause the German industrialization?	Causality/Engle and Granger/Toda and Yamamoto
2006	Ogren	Free or central banking? Liquidity and financial deepening in Sweden, 1834–1913	VAR
2006	Sabate, Gadea and Escario	Does fiscal policy influence monetary policy? The case of Spain, 1874–1935	VAR Model/Granger Causality
2006	Jacks	What drove 19th century commodity market integration?	AR/random walk
2006	Mattesini and Quintieri	Does a reduction in the length of the working week reduce unemployment? Some evidence from the Italian economy during the Great Depression	Time series/heteroskedasticity
2006	Kling	The long-term impact of mergers and the emergence of a merger wave in pre-World-War I Germany	VAR
2007	Crafts, Mills and Mulatu	Total factor productivity growth on Britain's railways, 1852–1912: A reappraisal of the evidence	random effects/stochastic cost frontier model
2007	Ciccarelli and Fenoaltea	Business fluctuations in Italy, 1861–1913: The new evidence	cycle/random walk/Kalman filter
2007	Herranz-Locan	Infrastructure investment and Spanish economic growth, 1850–1935	VAR/Cointegration/Granger Causality/Dickey Fuller
2007	Landon-Lane and Rockoff	The origin and diffusion of shocks to regional interest rates in the United States, 1880–2002	VAR/time series/Unit root/Dickey Fuller
2007	Eichengreen and Hataseb	Can a rapidly growing export-oriented economy exit smoothly from a currency peg? Lessons from Japan's high-growth era	Dickey Fuller
2007	Bodenhorn	Usury ceilings and bank lending behavior: Evidence from nineteenth century New York	Unit root /AR/ARMA
2007	Streb, Wallusch and Yin	Knowledge spill-over from new to old industries: The case of German synthetic dyes and textiles (1878–1913)	VAR/VECM/Unit root tests/Johansen/Granger causality
2007	Solomou and Shimazaki	Japanese episodic long swings in economic growth	Kalman filter/Stochastic shifts/Johansen cointegration/unit root tests

2007	Federico	Market integration and market efficiency: The case of 19th century Italy	TAR/unit root/MA
2007	Jacks	Populists versus theorists: Futures markets and the volatility of prices	GARCH/time series
2008	Craig and Holt	Mechanical refrigeration, seasonality, and the hog–corn cycle in the United States: 1870–1940	TV-STAR/random walk/Dickey Fuller/unit root/TVAR
2008	Siklos	The Fed's reaction to the stock market during the great depression: Fact or artefact?	VAR/GARCH/SVAR/Impulse response
2008	Diebolt and Parent	Bimetallism: The "rules of the game"	VAR & Granger causality
2009	Rousseau	Share liquidity, participation, and growth of the Boston market for industrial equities, 1854–1897	VAR & Granger causality/Unit roots
2009	Klovland	New evidence on the fluctuations in ocean freight rates in the 1850s	Cointegration/Johansen/Dickey-Fuller
2009	Grossman and Imai	Japan's return to gold: Turning points in the value of the yen during the 1920s	Non stationary process/random walk/structural breaks
2009	Esteves, Reis and Ferramosca	Market Integration in the Golden Periphery. The Lisbon/London Exchange, 1854–1891	random walk/stationary/AR1
2009	Sarferaz and Uebele	Tracking down the business cycle: A dynamic factor model for Germany 1820–1913	Hodrick Prestcott filter/AR

Table 2
Papers using Time Series methods: *Cliometrica*, 2007-2009.

YEAR	AUTHOR(S)	TITLE	METHODS
2007	Neilsen	UK money demand 1873–2001: a long-run time series analysis and event study	VAR/Johansen/Stochastic Trends
2008	Mills	Exploring historical economic relationships: two and a half centuries of British interest rates and inflation	GARCH/EGARCH/non stationary series/Dickey Fuller/AR(2)
2008	Ricciuti	The quest for a fiscal rule: Italy, 1861–1998	Non-stationary/Engle-Granger/Autoregressive model/Dickey-Fuller
2009	Prados de la Escosura and Sanz-Villarroya	Contract enforcement, capital accumulation, and Argentina's long-run decline	Hodrick-Prestcott/Granger Causality/Augmented Dickey Fuller test
2009	Weisdorf and Sharp	From preventive to permissive checks: the changing nature of the Malthusian relationship between nuptiality and the price of provisions in the nineteenth century	VAR/CVAR/structural break
2009	Ljunberg and Nilsson	Human capital and economic growth: Sweden 1870–2000	Granger causality/Toda and Yamamoto/VAR/ADF
2009	Franck and Krausz	Institutional changes, wars and stock market risk in an emerging economy: evidence from the Israeli stock exchange, 1945–1960	VAR/Regression/Break
2009	Mills	Modelling trends and cycles in economic time series: historical perspective and future developments	ARIMA/Cycles/Kalman Filter/HP Filter
2009	Baubeau and Cazelles	French economic cycles: a wavelet analysis of French retrospective GNP series	Cycles/Time Series/Non-Stationary

4.2 When did the Industrial Revolution begin? Some results utilising unit root tests, structural breaks and direct measures of persistence.

In a series of papers we used time series methods to identify the timing and potential causes of the British Industrial revolution see, Greasley and Oxley (1994a,b, 1996b, 1997c,e,f, 1998,a,b, 2000). In this section we will present some of these results to demonstrate how time series methods were used to consider such questions and also present some new results based upon the tests of Leybourne, Kim and Taylor (2007)

The data used in this series of papers relates to an extended version of Craft and Harleys (1992) "best guess" estimates of the index of British Industrial Production, extended from 1913 to 1992. In a series of papers, we consider structural breaks in the series utilising both the Dickey-Fuller (1981) approach and the extensions of Perron (1989) and Zivot and Andrews (1992). On the basis of the results repeated below as Table 3, we identify an alternating TS/DS/TS characterisation of the data for the period 1700-1913 and present a case for dating the British Industrial Revolution as 1780-1851, see Greasley and Oxley (1994a,b, 1996b, 1997c,f). Furthermore, Greasley and Oxley (1996a, and 1997b,) use Perron (1989) and Zivot and Andrews (1992) methods to identify crashes and breaks in the post 1913 data coinciding with World War 1; the post (WW1) decline; a 1973 trend break and a 1979 crash, see Table 4 below. This leads to an alternating TS/DS/TS characterisation for the whole sample period 1700-1992.

Table 3
Testing for unit roots: *Levels data*

	1700-1913	1700-1780	1780-1851	1851-1913
ADF(2)	-1.13	-3.66*	-1.16	-4.55*
LM(SC)	0.20	2.51	0.84	0.81

ADF(2) denotes 2 augmentations: LM(SC) is a Lagrange Multiplier test of first-order serial correlation: * denotes significant at the 5% level based upon MacKinnon (1991). Results not presented here indicate that for the periods 1700-1913 and 1780-1851, the data is I(1) and not I(2).

Table 4
Perron-type unit root tests - 1922-92

Year	*Crash*	*Trend*	*Crash & Trend*
1929	*-2.362*	*-2.084*	*-2.046*
1939	*-2.758*	*-2.550*	*-2.600*
1945	*-3.191*	*-2.672*	*-2.741*
1973	*-2.948*	*-4.827**	*-4.809**
1979	*-4.362**	*-4.419**	*-4.718**

** denotes significant at the 0.05 level using Perron (1989) critical values.*

Furthermore, on the basis of ADF test results we concluded that the period 1700-1992 comprises several distinct epochs of industrial growth, in particular: *1700-1780; 1781-1851; 1852-1913; 1922-1973* and *1973-1992.* The defined periods are supported by the rich economic historiography and by the statistical results.

However, as discussed above, the characterisation of the time-series properties of a series as either DS or TS is an extreme one. In contrast, the results presented as Table 5 consider the Cochrane measure of persistence over a number of periods, including Greasley-Oxley epochs. This is crucial as Cochrane (1988) demonstrates that measures of persistence constructed for periods (segments) of differing growth rates, will tend to bias the results in favour of finding *too much persistence.*

If we consider the results presented as Table 5, and limit discussion initially to the column "Chatfield" which gives the Chatfield (1989) $2\sqrt{T}$ criteria for the choice of window width (where T is the effective sample size), a number of features emerge. Firstly, the periods identified by Greasley and Oxley (1994a,b) have markedly different measures of persistence which lend support to the results based upon ADF tests. As can be seen, the persistence measures for Greasley-Oxley epochs (pre-WW1) are respectively 0.132, (1700-1780), 0.607, (1781-1851) and 0.449 (1851-1913) showing that the Industrial Revolution exhibited a marked difference in persistence from the earlier and later periods.

Table 5
Cochrane (1988) measure of persistence

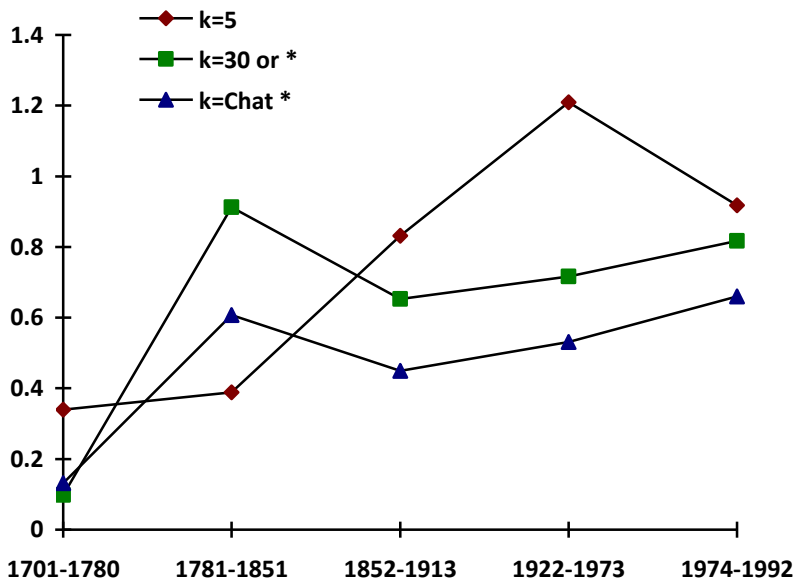
	k=5	k=10	k=15	k=20	k=30	Chatfield	k
1701-1992	0.521	0.469	0.465	0.544	0.659	0.712	34
	(0.077)	(0.096)	(0.116)	(0.153)	(0.218)	(0.248)	
1701-1780	0.340	0.265	0.164	0.135	0.098	0.132	18
	(0.092)	(0.094)	(0.066)	(0.058)	(0.043)	(0.056)	
1781-1851	0.389	0.464	0.577	0.704	0.912	0.607	16
	(0.111)	(0.173)	(0.241)	(0.310)	(0.396)	(0.258)	
1852-1913	0.832	0.538	0.497	0.558	0.653	0.449	16
	(0.251)	(0.209)	(0.214)	(0.248)	(0.270)	(0.195)	
1852-1992	0.807	0.558	0.532	0.622	0.559	0.549	24
	(0.169)	(0.159)	(0.179)	(0.232)	(0.234)	(0.217)	
1922-1992	1.225	0.975	0.980	1.223	1.377	1.058	16
	(0.349)	(0.362)	(0.411)	(0.543)	(0.597)	(0.449)	
1922-1973	1.209	0.789	0.593	0.774	0.716	0.531	14
	(0.391)	(0.195)	(0.172)	(0.343)	(0.265)	(0.233)	
1974-1992	0.918	0.504	0.817	-	-	0.660	8
	(0.362)	(0.192)	(0.172)			(0.279)	

k denotes the window size for the Bartlett estimator; figures in parentheses are asymptotic standard errors; a - denotes not calculated. All figures are corrected for small sample bias following Cochrane (1988).

Turning to the preferred Cochrane window width of 30, the results are even more pronounced with measures of 0.098, 0.912 and 0.653 for the respective periods.

Using the full sample period 1700-1992 and assuming no breaks in the series implies a high degree of persistence, i.e. 0.712 for the Chatfield rule, or 0.659 for k=30, reflecting the bias raised by Cochrane (1988) in favour of excessive persistence (or in favour of DS). A similar problem arises if the twentieth century is treated as a single epoch. In particular, the results for 1922-1992 imply a degree of persistence close to 1 i.e. 1.058 for k=14 or 1.377 for k=30. However, if the Greasley-Oxley epochs are considered, the pattern of persistence presented as Figure 1 (derived from Table 5), emerges.

Figure 1.
Measures of Persistence over various epochs



*=15 in the case of 1974-1992

For either k=30 or the Chatfield rule, persistence rises during the Industrial Revolution from the very low levels of pre-industrial Britain. It then declines pre-WW1, recovering

only slowly to (or approaches, based upon k=30) its Industrial Revolution level. However, some caution needs be expressed about the period 1974-1992 given the small sample size. The results based on k=30 or k=Chatfield are qualitatively the same, however, because of the quantitative differences the interpretation differs in important ways. In particular based upon k=30, the Industrial Revolution represents an historical high point in terms of persistence. Twentieth century persistence levels are moderately high and higher than the mid-late nineteenth century, but lower than the period 1780-1851. This result is not as clear-cut based upon k=Chatfield, although it depends crucially upon the small sample results of the period 1974-1992. On these basis, the Industrial Revolution period identified by Greasley and Oxley represents a unique period of high persistence.

Notice, however, that using a low value window, k=5, which is similar to a low order ARMA measure, such as Campbell and Mankiw (1987), suggests a much different picture, see Table 5 and Figure 1. Here, the Industrial Revolution seems unremarkable and the tendency for persistence to rise to very high levels into the twentieth century seems to emerge. Treating the period 1922-1992 as a valid era would exaggerate the position even further. The k=5 results help explain why some other measures exaggerate the degree of persistence experienced during the twentieth century. This coupled with the treatment of the post WW1 period as a single era explain why some results, i.e. Mills (1991) and Capie and Mills (1991) find in favour of a switch from TS to DS at this juncture. However, on statistical grounds, see Cochrane (1988), we would suggest that results based upon low valued k are in fact spurious.

Apart from providing a measure of persistence, interpretation of the normalised spectral density function gives a measure of the proportion of total variance of the process accounted for by cycles of various lengths, $l=2\pi/\omega$ where ω is the frequency, see Priestley (1981).

Table 6, below, presents estimates of the cycle lengths contributing most to the explanation of the variance of industrial production based upon the Chatfield rule for choosing window width.

Table 6
Cycle lengths, $l=2\pi/\omega$

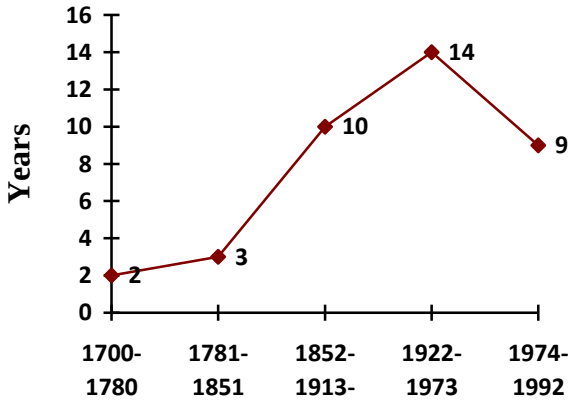
Years	*k*	*Cycle*
1701-1780	18	2
1781-1851	16	3
1852-1913	16	10
1922-1973	14	15
1974-1992	9	9

k=Chatfield window width

For the period up to 1913 it can be seen how pre- and post-industrial production cyclical elements differ. Short cycles (around 2-3 years), or a low cyclical element to the data, explain most of the variance in output pre-1780. Whereas post 1851 cycles of around 10 years, or a high cyclical element to the series, explain most of the variance in production. The cyclical nature of the Industrial Revolution period is similar to the pre-revolution period - in both cases the data suggests an economy not best classified as cyclical. However, as discussed below, for different reasons. These measures are qualitatively similar (though with a somewhat different interpretation given the methods of analysis) to those presented by Crafts, Leybourne and Mills (1989a) where they consider the periods pre-1783, with cycles of around 4 years, and post 1815, with cycles of around 7 years. Plots of the spectrum also indicate the following phenomena. For the periods 1701-1780,

1781-1852, the spectrum tends to rise continuously from low to high frequencies indicating the contribution of short cycles in explaining the variance of output. However, the shape of the spectrum reverses post 1852, including the twentieth century, indicating the contribution of longer cycles. Taking the post 1922 period as a whole suggests that cycles of 14-15 years contribute most to the explanation of output variance. These conform closely to the results of Leung (1992). However, as with the persistence measure, treating the twentieth century as a single epoch can be misleading. Taking the Greasley-Oxley epochs produces the following description, presented as Figure 2 below, of the cycle lengths which contribute most the explanation of the variance of industrial production.

Figure 2.
Implied cycle lengths



The amount of variance explained by longer cycles (or a more cyclical characteristic to the data), increases after the Industrial Revolution. However, cycle length appears to decline as persistence appears to increase. It does seem however, that there is strong

quantitative and historical evidence in favour of distinct periods of growth, as argued by Greasley and Oxley.

4.2.1 When did the Industrial Revolution begin: Revisited. The Multiple Persistence Change Model of Leybourrne, Kim and Taylor (2007).

In the approach presented above, it is relatively easy to identify a change in persistence stemming from a shift from a trend stationary (TS) process to one that is difference stationary (DS). However, it is less simple to identify a process whose time series properties appear to alternate, TS-DS-TS etc., as the properties of the data in the apparent DS section would tend to dominate the TS sections. Leybourne, Kim and Taylor (2007), consider testing for multiple changes in persistence model, in particular, they propose a test for the presence of multiple regime shifts and how to consistently estimate the associated change-point fractions. They partition *yt*, $t = 1, 2, ..., T$, into its separate $I(0)$ and $I(1)$ regimes and show that a test statistic appropriate for this purpose is based on a doubly-recursive application of a unit root statistic where they employ the local GLS detrended ADF unit root testing methodology of Elliott *et al.* (1996), - discussed above - used for detecting a single change in persistence. For further details of the test see Leybourne, Kim and Taylor (2007).

Utilising the same data as above on the log British industrial production for the period 1700-1992, the Leybourne, Kim and Taylor *M* test (with a maximum of 4 lags) implies that the series were:

> I(1) from 1700-1775; I(0) from 1775-1816; I(1) from 1816-1853; I(0) from 1853-1913; I(1) from 1913-1949; I(0) from 1949-1973; I(1) from 1973-1992.

The detailed test results are provided below in Table 7 below.

Table 7

Sample period	1700-1992	1700-1853	1913-1992
M statistic	-4.813 *	-6.252 *	-5.605*
I(0) start - end	1853-1913	1775-1816	1949-1973

* Denotes significant at the 5% level

Although the subsamples are slightly different to those proposed by Greasley and Oxley (1996b), the notion of alternating time series properties of the data are supported as is the macro-based timing of the Industrial Revolution.

4.3 Testing for Causality

4.3.1 What caused the British Industrial Revolution- disaggregate data?

In section 3.7 above, we presented a number of approaches to testing for Granger-type causality and stressed the need to consider the order of integration of the variables under consideration as this will crucially affect the validity of the inferences drawn.

In this section we will present some results taken from Greasley and Oxley (2000) where we consider the question which sectors 'caused' the British Industrial Revolution by utilising measures of disaggregated British industrial production, 1815-1851.

Causality in the context of these data and this question concern the linkages among the industries whose output defined early British industrialization. Of particular interest are those industries which are also ascribed in the historiography with key roles

in leading industrialization. Specifically we consider the importance of cotton, see Rostow (1963), utilizing the data for cotton cloth; iron and steel goods, see Hirschman (1958); coal, see Wrigley (1988), and elements of the food processing sector, see Horrell, Humphries, and Weale (1994), paying particular attention to beer and sugar. We show below that each of these industries have non-stationary output data for the period 1815-1860, and were part of wider groupings of industries which shared stochastic common trends. Any causal links among the industries with non-stationary data may be long-term, since their output movements have permanent effects. At issue is whether or not particular industries within the cotton, mining and metals, or the food and drink groups played leading roles within their sector, or had causal linkages which spilled across the common trend groupings.

Table 8 below presents unit root tests of the individual series from which those series identified as I(1) and I(0) can be identified. Given that total industrial production is deemed to be a unit root process over this period, if we are interested in ascertaining which series (variables) 'caused' this non-stationary outcome, we need to consider only those that are individually I(1). As discussed in 3.7, the form of the causality testing depends crucially on the order of integration of the univariate series. Of the 15 series deemed I(1), several combinations were used to consider robustness amongst the I(1) series.

Table 8
Unit Root Tests 1815-1860 (Augmented Dickey-Fuller Statistics+)

	ADF(1)	Trend (% p.a.)
A. Minerals		
1. Coal	-0.597	
2. Copper ore	-3.005	
B. Metals		
1. Pig iron and steel	-1.061	
2. Iron and steel products, machines and tools	-1.248	
3. Copper		
4. Tin	-1.491	
	-3.180	
C. Textiles		
1. Cotton yarn	-0.603	
2. Cotton cloth	-0.831	
3. Woolen and worsted yarn	-3.984*	1.77
4. Woolen and worsted cloth	-5.117*	1.40
5. Silk thread	-4.552*	2.95
6. Silk goods	-5.299*	3.61
7. Linen yarn	-4.203*	1.50
8. Linen goods	-4.265*	0.60
9. Hemp products	-3.385*	1.61
D. Food drink and tobacco		
1. Wheaten flour	-5.847*	0.80
2. Bread and cakes	-5.706*	0.96
3. Sugar	-1.538	
4. Beer	-2.577	
5. Malt	-2.471	
6. Spirits	-1.715	
7. Tobacco	-0.881	
E. Miscellaneous		
1. Shipbuilding	-2.617	
2. Paper	-0.281	
3. Leather	-3.843*	1.60
4. Leather goods	-3.937*	1.60

+ All the results are for ADF(1).
* Denotes significant at the 5% level according to MacKinnon (1991) critical values.

Table 9 below presents Johansen-based test results for cointegration using a 12 variable group which comprises the I(1) variables: coal, copper, cotton yarn, cotton pieces, pig-iron, malt, paper, shipbuilding, spirits, sugar, tobacco products, and beer. The

discontinuity in the spirits data around 1823 would tend to promote an idiosyncratic trend for this industry, and thereby reduce by one the number of common trends. r=the number of cointegrating vectors where the VAR lag length is chosen to 2 on the basis of Information Criteria tests.

Table 9
Johansen cointegration test results
For 12 I(1) series1815-1860

H_0:	H1:	Maximal eigenvalue	Trace
r=0	r=1	171.7*	683.0*
r≤1	r=2	113.3*	511.2*
r≤2	r=3	91.60*	397.9*
r≤3	r=4	85.67*	306.2*
r≤4	r=5	64.51*	220.6*
r≤5	r=6	44.60	156.1*
r≤6	r=7	37.95	111.4*
r≤7	r=8	26.89	73.54
r≤8	r=9	20.78	46.64
r≤9	r=10	14.36	25.86
r≤10	r=11	8.18	11.49
r≤11	r=12	3.31	3.31

*denotes rejects the null at the 5% level.

Table 10
Johansen cointegration test results
Mining and Metals Group, 1815-1860

H_0:	H1:	Maximal eigenvalue	Trace
r=0	r=1	85.81*	254.6*
r≤1	r=2	72.47*	168.7*
r≤2	r=3	42.21*	96.32*
r≤3	r=4	25.68*	54.11*
r≤4	r=5	20.74*	28.42*
r≤5	r=6	7.68	7.68

The Group comprises I(1) variables: coal, copper, copper ore, iron-steel goods, pig iron, and tin; I(0) variables: none. r=the number of cointegrating vectors where VAR lag length is 3, and * denotes rejects the null at the 5% level.

Table 11
Johansen cointegration test results
Textiles Group, 1815-1860

H_0:	H1:	Maximal eigenvalue	Trace
r=0	r=1	21.95*	27.73*
r≤1	r=2	5.78	5.78

The Group comprises I(1) variables: cotton pieces, cotton yarn. I(0) variables: linen yarn, linens, silk products, silk thread, woolens, worsted, hemp products, leather, and leather products. r=the number of cointegrating vectors where VAR lag length is 2, * denotes rejects the null at the 5% level.

Table 12
Johansen cointegration test results
Food, Drink and Tobacco Group, 1815-1860

H_0:	H1:	Maximal eigenvalue	Trace
r=0	r=1	102.2*	206.8*
r≤1	r=2	49.56*	104.6*
r≤2	r=3	34.64*	55.02*
r≤3	r=4	19.41*	20.37*
r≤4	r=5	0.96	0.96

The Group comprises I(1) variables: beer, malt, sugar, spirits and tobacco products. I(0) variables: bread, wheaten flour. r=the number of cointegrating vectors where VAR lag length is 4, * denotes rejects the null at the 5% level.

Tables 10-12 present cointegration results for groupings of the I(1) variables into sectors. The results from Table 9 identify 7 significant cointegrating relationships and hence 5 stochastic trends; Table 10, 5 significant cointegrating relationships and 1 stochastic trend; Table 11, 1 significant cointegrating relationships and 1 stochastic trend; and Table 12, 4 significant cointegrating relationships and 1 stochastic trend.

The results of tests for bi-variate causality between cotton pieces, iron and steel goods, coal, beer, sugar, and the other industries with non-stationary output series are shown in Table 13, below using the Toda and Yamamoto (1995) method:

Table 13
Toda and Yamamoto-type tests of causality, I(1) variables
1815-1860

Variable 1	Variable 2	p value 1 does not cause 2	p value 2 does not cause 1
Coal	Copper	0.253	0.626
	Copper ore	0.342	0.662
	Cotton pieces	0.429	0.003*
	Cotton yarn	0.340	0.008*
	Sugar	0.765	0.049*
	Malt	0.256	0.336
	Spirits	0.338	0.966
	Tobacco products	0.597	0.004*
	Iron and steel	0.738	0.011*
	Tin	0.672	0.994
	Pig iron	0.614	0.091**
	Paper	0.122	0.000*
	Shipbuilding	0.012*	0.230
	Beer	0.265	0.367
Cotton pieces	Copper	0.236	0.601
	Copper ore	0.508	0.487
	Cotton yarn	0.001*	0.003*
	Sugar	0.030*	0.252
	Malt	0.256	0.185
	Spirits	0.355	0.275
	Tobacco products	0.245	0.690
	Iron and steel	0.002*	0.003*
	Tin	0.657	0.811
	Pig iron	0.001*	0.001*
	Paper	0.000*	0.117
	Shipbuilding	0.064**	0.822
Sugar	Copper	0.681	0.536
	Copper ore	0.032*	0.198
	Cotton yarn	0.351	0.021*
	Malt	0.865	0.991
	Spirits	0.934	0.414
	Tobacco products	0.670	0.869
	Iron and steel	0.170	0.156
	Tin	0.415	0.944
	Pig iron	0.159	0.225
	Paper	0.004*	0.164

	Shipbuilding	0.932	0.291
Iron and Steel	Copper	0.149	0.207
	Copper ore	0.524	0.192
	Malt	0.046*	0.922
	Spirits	0.463	0.922
	Tobacco products	0.785	0.933
	Tin	0.927	0.794
	Pig iron	0.493	0.385
	Paper	0.404	0.334
	Shipbuilding	0.368	0.057**
Beer	Copper	0.536	0.353
	Copper ore	0.171	0.903
	Malt	0.349	0.091**
	Spirits	0.205	0.226
	Tobacco products	0.898	0.519
	Tin	0.405	0.632
	Pig iron	0.785	0.374
	Paper	0.335	0.213
	Shipbuilding	0.170	0.293
	cotton piece	0.260	0.467
	cotton yarn	0.316	0.467
	iron and steel	0.588	0.482
	sugar	0.187	0.036*
	malt	0.349	0.091**

p value denotes the probability of a causal relationship.* denotes reject non-causality null in favor of causality at the 5% level.** denotes reject non-causality null in favor of causality at the 10% level.

On the basis of the causality tests, the industries with the most pervasive links to other industries are coal and cotton. Clearly, coal appears to be a follower, with its output determined by previous output levels of cotton yarn, cotton cloth, sugar, tobacco, iron and steel goods, pig iron, and paper. Only shipbuilding appears to have been led by coal, a finding which possibly reflects the importance of coal to the British coastal trade, especially between Newcastle and London, in the first half of the nineteenth century. A widening industrial demand for coal as a fuel and power source in the 1815-1860 period from the metals, foods, and textiles sectors, was the key to coal's expansion. Cotton

pieces, to the contrary, played a leading role in industrialization by stimulating coal, sugar, iron and steel goods, pig iron, paper, shipbuilding, and cotton yarn, though the latter is partly an artefact of data construction. The two industries with idiosyncratic stochastic output trends, paper and shipbuilding, were both led by the cotton industry. To some extent the cotton results highlight the separateness of the foods sectors, with a causal link found only in the case of sugar.

Iron and steel goods, identified by the input-output studies of Chenery and Watanabe (1958), and Hirschman (1958) as the industry with the widest transactions linkages, had fewer causal links than cotton in the 1815 to 1860 period. Iron and steel goods were led by shipbuilding, and had no significant links with paper, though like cotton pieces it led one foods industry, in its case, malt, and coal. However, the results do not provide a simple basis for favoring cotton pieces over iron goods as the key to early industrialization, as the results indicate bidirectional causality between the two industries**.** Thus, the cotton, mining, and metals sector appears jointly dominated by cotton and iron, while coal appears as a follower, and no causal links emerge for tin, copper, or copper ore.

The causal links surrounding the foods sector appear less widespread. Beer, malt, spirits, sugar, and tobacco form the key foods grouping with a single stochastic common trend. Of these, beer is awarded the largest weight at 2.6% in Hoffmann's industrial production index (excluding building) for 1831-1860. Beer does not lead any other industry according to the causality tests, but follows both malt and sugar. In the foods sector, sugar has the widest causal links leading coal, paper, and copper ore, but following cotton. Generally the results play-down the importance of the linkages

surrounding the foods sector, and its role in defining the profile of early British industrialization. In part this arises from the large bread and flour industries having trend stationary output, which excludes both from shaping the swings in aggregate output and from having long-run causal links with the other foods industries, and with the cotton, mining and metals grouping. Sugar did lead, along with several other industries, the expansion of coal output, but had no causal links with iron goods and was led by cotton.

The causality test results help to refine further the interpretation of early industrialization, which emerged from the common trends perspective. In the case of the cotton, mining and metals grouping, cotton and iron are revealed as the leading industries, and coal as a follower. Outside the sector, cotton had the wider linkages, statistically causing paper and shipbuilding production. However, cotton cannot be regarded as the more important lever to industrialization in the 1815-1860 period as bidirectional causality between iron goods and cotton output lies at the heart of the cotton, mining and metals sector. The findings for the foods sector show fewer within sector causal links, though both malt and sugar led beer. Only sugar appears to have causal links reaching outside the foods sector, to paper, coal, and copper ore. Together, the common trends and causality results point to cotton, iron goods, and possibly sugar, as the key industries promoting swings in British industrialization to 1860.

4.3.2 What caused the British Industrial Revolution- aggregate level data?

In section 4.3.1 above and Greasley and Oxley (2000) we consider testing for the causes of the Industrial Revolution using Hoffman's (1955) disaggregate industrial production data. However, debate also revolves around the possible macro-level causes. In Greasley and Oxley (1997e, 1998a), we utilise time series methods to identify the causes for the

extended period 1780-1851. Several candidates for causality exist in the literature including:

i) Export Led Growth (ELG): export growth *causes* growth in output. This view is supported by for example, O'Brien and Engerman (1991), and Hatton and Lyons (1983), and was tested utilising data on industrial production and exports.

ii) Technological factors: developments in technology *cause* a change in the productive process and/or efficiency of production leading to a discernible change in the pattern of output growth. This view is supported by Tsoulouhas (1992), and was tested utilising data on the number of patents registered and processes stemming from such patents, as measured by Sullivan (1989).

iii) Population growth: here growth in the population influences output by both providing a growing pool of workers and also a growing source of domestic demand. Supporters of this view include Komlos (1990) and Simon (1994).

iv) Domestic factors (general): other domestic factors including for example wages and the change in domestic demand, are seen as contributing to a *domestically determined* revolution. Clearly population growth could be included in this category, although it is generally assigned a separate potential route of influence. Supporters of the domestically determined growth include Deane and Cole (1969) and McCloskey (1981). Such authors' views are often contrasted with supporters of the ELG hypothesis,

and i) and iv) could be regarded as two of the main competing explanations of the Industrial Revolution. In testing iv) data on real wages taken from Crafts and Mills (1994) are utilised to test whether real wage levels, or rates of growth, caused industrial production or vice versa.

v) Subsidiary hypotheses - imports cause exports: Deane and Cole (1969), posit that imports lead exports in the 18th century as British trade shifts from Europe to the W. Indies and North America. Colonial economies, however, had limited spending power and as such needed to export to Britain if they were to buy imports from Britain. If this hypothesis were true and were coupled with the ELG hypothesis, the data may suggest that imports *cause* output growth .

vi) Other possible candidate hypotheses: given the current level of interest in endogenous growth models see Rebelo (1991), it may seem natural to test for the effects of for example, human capital on growth. However, for the period of interest, 1780-1851, the absence of annual data precludes formal investigation of the roles played by investment in education and even physical capital in the Industrial Revolution.

As such there are five feasible testable hypotheses, i)-v) above. The first stage of the causality testing procedure investigates the order of integration of the data. The results of the tests (not presented here, but can be found in Greasley and Oxley (1997e, 1998a), where definitions of the variables used can also be found) on the log levels of, total industrial production; real wages; exports; imports; patents; processes and population

show in all cases that the null hypothesis of non-stationarity is not rejected. Results for tests of bivariate and multivariate cointegration (also not presented here but can be found in Greasley and Oxley (1997e, 1998a) between industrial production and the other variables of interest, namely, real wages; exports; imports; patents; processes and population identify a single significant bivariate cointegrating relationship between industrial production and the variable of interest in all cases, but imports. Utilising these results a multivariate Johansen approach was adopted including all variables except imports and and one significant cointegrating vectors was identified. A test of the restriction that the coefficient on population equals unity was not rejected implying that a proportionate relationship between the level of population and the level of industrial production cannot be rejected. The results, therefore, demonstrate the existence of both bivariate and multivariate cointegration between the variables of interest. The only candidate variable for which cointegration was not identified was imports. This constrains tests of causality to the I(0) representation of the import data, i.e. in this case, first difference or growth rates, and does not rule-out the identification of a spurious relationship. The potential for cointegration between exports and imports was discussed in Greasley and Oxley (1997e, 1998a) and is not considered further here, nor are the standard Granger-type causality tests except to summarise the results of such methods being the overall assessment is of unidirectional causality from processes to output or that *technological change caused changes in industrial production - the Industrial Revolution.*

4.3.3 Development Blocks, Innovation and Causality in New Zealand 1861-1939

The concept of a 'development block', where innovations in leading industries promote complementary activities, has been utilized widely to understand economic development see Dahmen (1988) and Rostow (1963). Early work, including that of Hirschman (1958) and Chenery and Watanabe (1958) measured input-output transactions to identify leading industries, highlighting the strategic importance of the linkages from iron and steel industries. In Greasley and Oxley (2010a) we gauge which, if any, New Zealand industry or groups of industries led her economic development.

A variety of approaches have been used to identify development blocks. Horrell, Humphries and Weale (1994) constructed an input-output table for 1841 to gauge the leading industries of the British Industrial Revolution, but the input-output method offers only a static perspective. Moser and Nicholas (2004) used historical patent citations to assess to the impact of electricity as a general purpose technology.

Enflo, Kander and Schon (2008) also consider how *development blocks* formed around electricity by using a combination of cointegration and causality analysis. They define a development block as consisting of a number of sectors that share a common long run trend (i.e. are cointegrated) and linked to each other by mutually reinforcing Granger causality, where the latter ensures short term complementarities among industries. The method of using modern time series methods to identify leading industries was earlier used by Greasley and Oxley (2000) to gauge the leading sector groups of the British Industrial Revolution.

In Greasley and Oxley (2010a) we firstly investigated the existence of common long run trends among the conventional industry sectors; Pastoral, Agriculture,

Manufacturing, Minerals and a Miscellaneous group. The unit root tests showed New Zealand's economic development was driven by 18 industries with the non stationary output trends. At issue is how many of these trends were common to more than one industry, pointing to the existence of development blocks. If trends were common to groups of industries then the possible sources of growth are simplified, as the effects of output innovations, including those from new technology will spill across industries. The cointegration tests showed (see Greasley and Oxley, 2010a, Table 3) that a small number of stochastic common trends drove output in most sector groups. Both the pastoral and agricultural sectors have two stochastic common trends, and the manufacturing sector only one. A cointegrating relationship was not observed for the mineral sector, and gold and kauri gum have individual output trends.

These findings show a small number of industry groups were central to the long run development of New Zealand economy to 1939, with, for example the 8 manufacturing industries forming a unified group. Interestingly though the pastoral sector did not form a singular development block. The existence of two common trends in the pastoral sector shows the dairy and the meat industries were not simply connected by the opportunities of refrigeration, but that different forces shaped their output. Possibly the dichotomy stems from much of the frozen meat trade originating in the South Island corporate enterprises, whereas North Island co-operatives dominated dairying. Dairying expansion also required the clearing and cultivation of wetter, forested North Island land, and different technology, most especially that connected to cream separation. In the case of agriculture the existence of two stochastic trends probably relates to differences between the output drivers of potatoes and the two grain crops (wheat and oats).

The existence of development blocks also requires short term complementarities between industries, and these are evaluated using tests of Granger causality. Additionally the list of possible leading industries will be made clearer by considering the direction of causal relationships between the industries. For example, sectors or development blocks which shared common trends may have been led by one particular industry. Of special interest is whether or not the impact of any industry spanned beyond its sector group to lead other sectors and overall commodity output.

Table 14
Summary of Granger Causality Tests (number of causal links)

	Leading	Following	Bi-directional
Pastoral Industries			
Meat	8	0	3
Cheese	6	1	3
Butter	8	1	1
Manufacturing Industries			
Wool Cloth	2	8	5
Beer	4	10	1
Grain milling	5	2	1
Biscuits	4	4	1
Saw mills & Doors	3	7	2
Foundry & Engineering	3	7	2
Printing & Publishing	7	1	2
Shoes & Boots	4	4	2
Agriculture			
Wheat	0	2	0
Oats	3	5	1
Potatoes	0	3	0
Mining & Other Industries			
Gold	2	0	4
Gas	4	3	2
Construction	0	11	2

The causal links between the 18 industries with non-stationary output are shown in Table 14, above. Generally, manufacturing, the pastoral industries and construction have the most causal links. However, construction and most manufacturing industries were followers, while pastoral industries' output often led the output of other industries. Thus beer has 15 causal links, including that with total commodity output, but in 10 cases beer followed, and in one other the causality was bi-directional. Similarly wool cloth has 15 causal links, with 13 of these as a follower or bi-directional. Construction has 13 causal links, but was unambiguously led by output in other industries. Several other manufacturing industries, including saw mills and doors, and foundry and engineering have multiple causal links, but principally these show bi-directional causality within manufacturing or a follower relationship with the pastoral sector. Within manufacturing, printing and publishing is the industry with the most leading causal links. Printing and publishing accounted for around 23% of (non-pastoral) manufacturing from 1915, becoming the largest element of the sector by 1935. The results show printing and publishing had bi-directional causality with all commodity output and wool cloth, and led beer, saw mills and doors, foundry and engineering, shoes and boots, potatoes, kauri gum and construction.

In the mining sector, gold was of principal importance and still contributed around 7% of commodity output in 1905. Gold had bi-directional causal links with all commodity output and the pastoral sector, and led beer output. The pastoral sector dominates the leading causal links with other industries. Meat and butter are the only industries which led all commodity output. Meat led the output of 9 industries, and had bi-directional causality with two others. Interestingly though, no evidence was found of

causality between the meat and dairy industries. In addition to all commodity output butter also led 7 industries, and cheese led the output of 6 industries, including butter.

The pastoral sector dominated economic development in New Zealand, but the meat and dairy sectors each had individual driving forces, and formed separate development blocks. Gold was important, at least until the early years of the twentieth century, and made a contribution to stimulating pastoral and all commodity output. The manufacturing sector (other than the manufacture of pastoral goods) did form a unified block which shared a single stochastic common trend, but most linkages of the sector, with one exception, were bidirectional or following. The exception is printing and publishing, which comprised a sizeable element of New Zealand manufacturing, and led four other manufacturing industries as well as potatoes, kauri gum and construction.A small number of key industries, specifically, meat, cheese, butter, gold, and printing and publishing, shaped the directions of New Zealand's economic development.

4.4 Time series-based test for Convergence

The economic underpinnings of the 'convergence hypothesis', the view that poorer countries, in terms of GDP per capita, tend to grow faster than richer countries and as a result, economies should 'converge' in terms of per capita income, arises naturally within the standard or augmented Solow neoclassical growth model. Here differences in initial endowments are seen to have no long term effects on economic growth with deficient countries able to catch-up to the leaders who suffer from diminishing returns. In contrast, Rebelo-type models of economic growth imply leadership can be maintained with non-convergence the likely outcome. As such, not only are tests of convergence

interesting in their own right, but they emerge as one natural testable implication of alternative models of growth. However, convergence is but one implication of such models and does not in itself represent a full test of the competing approaches. In order to test for convergence some form of clear definition and some appropriate form of time series data are required where, as we will see, the crucial feature to be exploited are the time series properties of the data.

Bernard and Durlauf (1995) utilise the Dickey-Fuller unit root testing procedure and cointegration as time series based tests of convergence. Here convergence implies output innovations in one economy should be transmitted internationally. The absence of transmission implies that per capita output differences between countries contains a unit root, since output shocks generating relative GDP movement infinitely persist causing economic divergence - an implication of the endogenous growth models of Rebelo.

Bernard and Durlauf (1995) define two types of convergence and two types of 'common trend'

Definition 2.1. Convergence in output

Countries i and j converge if the long-term forecasts of output in both (countries) are equal at a fixed time t:

$$\lim_{k\to\infty} E(y_{i,t+k} - y_{j,t+k} \mid I_t) = 0 \qquad (40)$$

Definition 2.1'. Convergence in multivariate output.

Countries p=1,...., n converge if the long-term forecasts of output for all (countries) are equal at a fixed time t:

$$\lim_{k\to\infty} E(y_{i,t+k} - y_{p,t+k} \mid I_t) = 0 \qquad \forall p \neq 1 \qquad (41)$$

Definition 2.2. Common trends in output

Countries i and j contain a common trend if the long-term forecasts (of output) are proportional at a fixed time t:

$$\lim_{k\to\infty} E(y_{i,t+k} - \alpha y_{j,t+k} \mid I_t) = 0 \qquad (42)$$

Definition 2.2'. Common trends in multivariate output

Countries p=1,...., n contain a single common trend if the long-term forecasts of output for all (countries) are equal at a fixed time t.

Letting $\bar{y}_t = \left[y_{2t}, y_{3t}, ... y_{pt} \right]$ then

$$\lim_{k\to\infty} E(y_{1,t+k} - \alpha_p^{'} \bar{y}_{t+k} \mid I_t) = 0 \qquad (43)$$

In terms of estimation and testing of the various types of convergence and common trend models, the main factor to note is that convergence implies that long-run forecasts of, in the case of output convergence, output differences, tend to zero at t→∞. If y_i and y_j etc, are I(1), which seems to be the empirical observation for most countries, it means that there is a natural way to test for convergence in the framework by invoking the properties and testing frameworks of unit roots and cointegration we have discussed earlier. In terms of Definition 2.1, *i* and *j* converge if their outputs are cointegrated with a restriction on the coefficients in cointegrating vector being [1, -1]. Alternatively in this bivariate case we can consider a simple unit root test on the differences in output. Note that if y_i and y_j are TS, then we can re-think of the definitions requiring that the time trends for each of *i* and *j* must be the same.

If the countries do not satisfy the strict requirement of convergence they may still be subject to the same permanent shocks. These are the cases relevant to Bernard and Durlauf's Definitions 2.2 and 2.2'. Testability in these cases can also invoke

cointegration, but in this case the requirements are not as strong – now the restrictions on the coefficients in cointegrating vector are [1, -α].

We can also consider slight variations on the Bernard and Durlauf (1995) definitions which can be illustrated via the concepts of *catching-up* and *long-run convergence.*

Definition: ***Catching-up:*** consider two countries *i* and *j*, and denote their per capita real output as y_i and y_j. *Catching-up* implies the absence of a unit root in their difference $(y_i - y_j)$.

This concept of convergence relates to economies *out of long run equilibrium* over a *fixed interval of time*, but assumes that they are sufficiently similar to make tests (and rejections), of the hypothesis non-trivial. In this case *catching-up* relates to the tendency for the difference in per capita output to narrow over time. Hence non-stationarity in $(y_i - y_j)$ must violate the proposition although the occurrence of a non-zero time trend, a deterministic trend, in the process in itself, would not.

Definition: ***Long-run convergence:*** consider two countries i and j, and denote their per capita real output as y_i and y_j . *Long-run convergence* implies the absence of a unit root in their difference $(y_i - y_j)$ or a time trend in the deterministic process, i.e., the absence of both a stochastic and deterministic trend.

Catching-up differs from long-run convergence in that the latter relates to some particular period T equated with long-run steady-state equilibrium. In this case the existence of a time trend in the non-stationary $(y_i - y_j)$ would imply a narrowing of the (per capita output) gap or simply that the countries though *catching-up* had *not yet* converged. Conversely,

the absence of a time trend in the stationary series implies that catching-up has been completed.

Clearly long-run convergence and catching-up are related in that both imply stationary $(y_i - y_j)$. However, long-run convergence relates only to (similar) economies in long-run equilibrium and therefore represents a much stronger version of the convergence hypothesis.

As defined above, tests of catching-up and long-run convergence hinge, therefore, on the time-series properties of $(y_i - y_j)$. The natural route for such tests involves Dickey-Fuller type tests based on the bi-variate difference in per capita output between pairs of countries, i and j, i.e.,

$$y_{it} - y_{jt} = \mu + \alpha(y_{i,t-1} - y_{j,t-1}) + \beta t + \sum_{\kappa=1}^{\nu} \delta\Delta(y_{i,t-k} - y_{j,t-k}) + \varepsilon_t \qquad (44)$$

where y indicates the logarithm of per capita output. If the difference between the output series contain a unit root, $\alpha=1$, output per capita in the two economies will diverge. The absence of a unit root, $\alpha<1$, indicates either catching-up, if $\beta\neq0$, or long-run convergence if $\beta=0$.

The main reservation surrounding the robustness of unit root tests in general, and therefore their application to tests of convergence in particular, concerns the possibility that structural discontinuities in the series may lead to erroneous acceptance of the unit root hypothesis.

Time series based tests of convergence, with and without breaks, were presented in Greasley and Oxley (1995, 1997a and 1998c) and a summary of some of those results are presented next.

4.4.1 Some results on convergence: Australia and Britain

Here we present pairwise tests for long run (steady state) convergence, and catching-up between Britain, Australia and the US for the period 1870-1992 using the unit root approach. On the basis of the results in Table 15 below, based on (43), neither version of the convergence hypothesis receives support, since a unit root cannot be rejected in the cross- country differences in GDP per capita.

Table 15
Unit root tests
UK-Australian, US-Australian and UK-US, GDP per capita
(without discontinuities)

Countries	Sample	ADF	LM(SC)
UK-Australia	1870-1992	*-3.250*	0.792
US-Australia +	1870-1992	*-2.301*	0.047
UK-US +	1870-1992	*-3.160*	0.994

** denotes significant at the 5% level based upon MacKinnon (1991), ADF denotes ADF(4) except those marked + which relate to ADF(2). # p value on included time trend = 0.627*

However the likelihood of structural discontinuities in the Australian (and the British) growth record, for example that associated with the crash of 1891, suggests their impact on the convergence process warrants investigation.

However, the failure of the time series approach to identify convergence may stem from discontinuities in the process generating the data and can be assessed by applying for example, the Zivot and Andrews' search procedure to the comparative series. The results in table 16 below report the maximum absolute ADF statistics obtained by searching over the period 1870-1992 for crash, trend, and joint crash and trend changes in a naturally extended version of equation (44).

All three pairwise results reject the existence of a unit root in some variant of the model and are supportive of some form of the convergence hypothesis. The UK-Australia results support long run convergence, with the 1891 crash marking a discontinuity in the process.

Table 16
Unit root tests - Differences in GDP per capita
Zivot and Andrews approach

Country	k	Year	Crash	Trend	Crash & Trend
UK-Australia	*4*	*1870-1992*	*-5.418**	*-4.192*	*-5.440**
			[1891]	*[1899]*	*[1891]*
US-Australia	*2*	*1870-1992*	*-4.085*	*-4.483**	*-4.585*
			[1891]	*[1943-44]*	*[1941]*
UK-US	*3*	*1870-1992*	*-4.303*	*-4.828*	*-5.342*
			[1966]	*[1950]*	*[1941]*

*k is the degree of augmentation * denotes significant at the 5% level based upon Zivot and Andrews (1992), [ ] denotes the year of the maximum absolute value of the ADF.*

The UK-Australia results do contain a significant joint crash and trend change discontinuity in 1891, but the absence of a significant individual trend break and the closeness of the crash and joint crash and trend break ADF statistics point to the dominance of the 1891 crash. Alternatively, both the US-Australia and UK-USA results contain significant trend discontinuities, and hence favour the weaker, *catching-up* version of the convergence hypothesis. Catch-up towards the USA's GDP per capita appears to date from 1950 for the UK and the years of World War Two for Australia

The results therefore show how the omission of significant discontinuities can lead to incorrect inferences being drawn regarding convergence and importantly the possible causes of economic growth. In particular, in the UK-Australia case where long run convergence is inferred, support for the Solow growth model emerges. Catching-up in the UK-USA and USA-Australia case is supportive of the augmented Solow model, while no case supports one important implication of the Rebelo model, i.e., long term

non-convergence. However, as stated earlier such implications do not constitute full tests of the growth models.

4.4.2 Alternative time-series based tests of convergence: St. Aubyn (1996)

St. Aubyn (1999), defines convergence as follows:
Two series , $y_i - y_j$, converge if:

$$(y_i - y_j) \rightarrow p\ \varepsilon_t \text{ as } t \rightarrow \infty \qquad (45)$$

where $\rightarrow p$, means converges in probability and ε_t is a random variable where:

$$E(\varepsilon_t) = D_{XY} \qquad (46)$$

$$\text{var}(\varepsilon_t) = \sigma < 0 \qquad (47)$$

Via (45)-(47), convergence implies that the difference between the two series converges in probability to a third series which is stationary, with constant mean D_{XY} and a constant variance σ. St. Aubyn (1999), relates these characteristics to previous notions of economic convergence, i.e.;

i). Point wise convergence $\rightarrow \text{var}(\varepsilon_t) = 0$,

ii). Unconditional convergence $\rightarrow \ D_{XY} = 0$

iii). Conditional convergence $\rightarrow \ D_{XY} \neq 0$.

4.4.2.1 Some results using the St.Aubyn approach

The results presented below as Tables 17 – 19 are taken from Oxley and Greasley (1997b) where they are particularly interested in establishing whether some groups of countries converge (but not others) via the idea of 'Convergence Clubs'.

The results treat France as the leader i.e., y_i in all instances where a 'European Club' is considered and likewise Sweden for the 'Nordic' group. For the full sample, 1900-1987, the null hypothesis of non-convergence is rejected in all cases with the

weakest result being between France and Germany where rejection is at the 10% level only. On the basis of the full sample period results the concept of a European Convergence Club comprising France, Belgium, Germany, Italy, The Netherlands and the United Kingdom cannot be rejected and likewise a Nordic Club of Sweden, Denmark, Finland and Norway.

St.Aubyn (1999) finds in his study of convergence, where he treats the United States as the leader, that in all cases that pre- and post- World War Two results differ. Table 18 and 19 below, taken from Greasley and Oxley (1997b), present results for the sample sub-periods 1900-1938 and 1946-1987. The results from Table 18 lead to non-rejection of the non-convergence null for all cases. The results from Table 19, however, confirm, with somewhat more significant, results from the full sample conclusions, i.e., rejection of the non-convergence null in all cases for the European Club, but not non-rejection in the Nordic case.

Table 17
Tests of Convergence: St. Aubyn test

Countries	1900-1987
France-Austria	*-3.359**
France-Belgium	*-2.681**
France-Germany	*-2.297***
France-Italy	*-9.004**
France-Netherlands	*-6.277**
France-UK	*-6.082**
Sweden-Denmark	*-3.143**
Sweden-Finland	*-5.474**
Sweden-Norway	*-2.847**

* denotes significant at the 5% and ** 10% level based upon St.Aubyn (1999)

Table 18
Tests of Convergence: St. Aubyn test

Countries	1900-1938
France-Austria	*0.200*
France-Belgium	*1.583*
France-Germany	*1.436*
France-Italy	*-0.154*
France-Netherlands	*0.107*
France-UK	*-0.025*
Sweden-Denmark	*1.339*
Sweden-Finland	*-0.217*
Sweden-Norway	*-0.601*

Table 19
Tests of Convergence: St. Aubyn test

Countries	1946-1987
France-Austria	*-10.24**
France-Belgium	*-4.336**
France-Germany	*-6.352**
France-Italy	*-2.651**
France-Netherlands	*-4.376**
France-UK	*-3.463**
Sweden-Denmark	*-0.119*
Sweden-Finland	*-1.142*
Sweden-Norway	*1.673*

* denotes significant at the 5% level based upon St.Aubyn (1999)

The sub-sample results seem to imply that, in the European case, convergence occurs most strongly post WWII . However, the full sample convergence implications for the

Nordic Club are not supported by either sub-period. This could be due to small sample estimation problems.

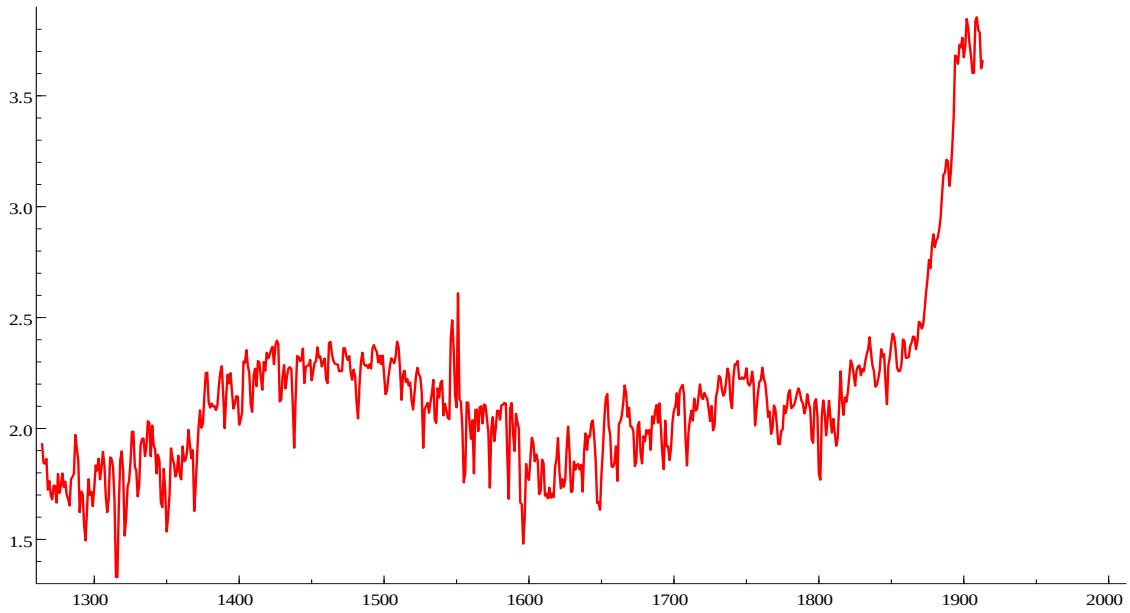
4.5 Application of STM model to English Real Wages data 1264 – 1913

In section 3.8 above some benefits of utilising the STM of Harvey (1989) were detailed. In the examples below we present some new results from applying these methods to the English (chiefly London) real wages data 1264 – 1913 using STAMP version 8.3. The real wages data are for building labour as discussed by Allen (2001). Long run real wages data have been utilized by Galor (2005) and by Crafts and Mills (2009) to consider the timing of the transition from the Malthusian era. The STM model provides an especially useful route to examining such issues by distinguishing between level and trend breaks and outliers in long run time series.

In the STM model the 'one-off' outliers are captured by interventions (for a discussion of the types of intervention see Harvey and Koopman (1992)), where outliers could be simply data driven measurement errors or 'obvious' one year only effects; the level shifts will typically pick up longer periods of level changes (possibly associated disease, including the Black Death, famine or wars); whereas a detected trend change will likely need an explanation via technological changes. Of course, if the underlying model structure is mis-specified then the interventions are just trying to handle the model misspecification. Thus the model we 'choose' should correspond with historical evidence of structural changes that suggest any dates identified for breaks are 'understandable and acceptable', most especially for any level and trend changes.

Figure 3 below presents a plot of the data, presented in logs.

Figure3.
Log of English Building Labour Real Wages, 1264-1913



As discussed in section 3.8 above there are alternative ways these data can be modelled within the STM approach. Some of the possible assumptions that can be made are shown and tested below, using the real wages data to illustrate. The key issues of the historiography concern the timing of the shift to higher trend real wages growth (associated either with the end of the Malthusian era, the Industrial Revolution or a post-Industrial Revolution demographic transition); and the existence or otherwise of a long period of constant real wages before any decisive shift to higher trend growth. Identifying trend breaks in the long run real wages data are complicated by particular demographic or monetary events, notably the near century of population decline associated with the Black Death from 1347 which reduced Europe's population by around one-third (Persson, 2010), and the debasements and re-coinages of the 16th century (Outhwaite, 1969). Real wages rose as population fell with the Black Death, but discerning the post-plague real

wages peak is further complicated by currency manipulations in the 1540s. English real wages show a peak around 1548. There were sharp spikes in London real wages around that date probably connected to currency manipulations, and post Black Death wage growth in England may not have been sustained beyond 1500.

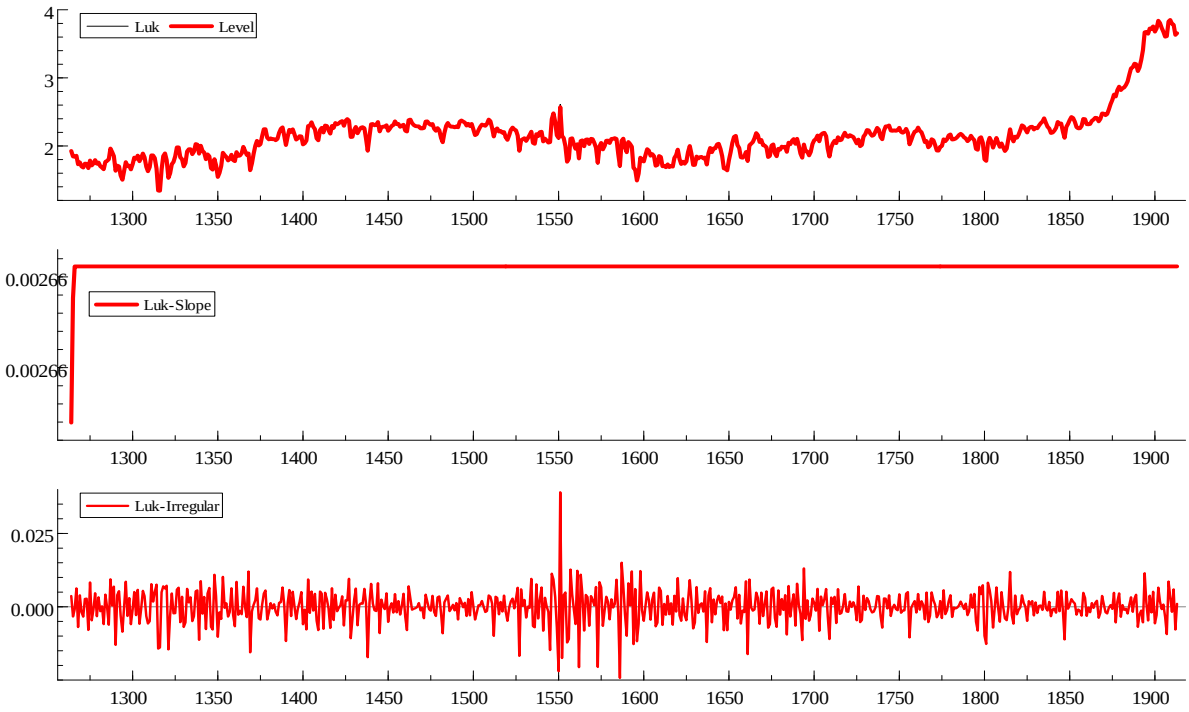
The ability of the STM model to reveal outliers as well as level and trend breaks offers the promise of disentangling the complex forces shaping long run English real wages. Some of the possible modelling strategies are listed 1-7, and a range of implications are illustrated below in sections 4.6.1-5.

1. The trend is a random walk with fixed drift (level stochastic and slope fixed) plus irregular (white noise), illustrated below as 4.6.1.
2. As above (1) but with a cycle which is AR2, not illustrated.
3. As above (2) but allowing for interventions to capture outliers. Therefore we have: the trend is a random walk with fixed drift (level stochastic and slope fixed) the cycle is AR2 plus irregular (white noise) + interventions (automatic), illustrated below as 4.6.2.
4. A local linear model: the level is stochastic; slope stochastic plus cycle plus irregular, not illustrated.
5. As above (4) but allowing for interventions, illustrated below, using a 5 year cycle, as 4.6.3 and 20-year cycle as 4.6.5.
6. As 4 above but fixed level; stochastic slope plus cycle plus irregular, not illustrated.
7. As above (6) but allowing for interventions, illustrated as 4.6.4 using a 5 year cycle.

4.5.1 Trend is a random walk with fixed drift

This model, without interventions to capture outliers, fits the data very poorly; the Durbin-Watson statistic=1.98 and R^2=0.004. The trend, shown in the middle segment of Figure 4 (the slope of the log of real wages) is constant by assumption; it would be clearly wrong to impose this formulation to represent the data.

Figure 4



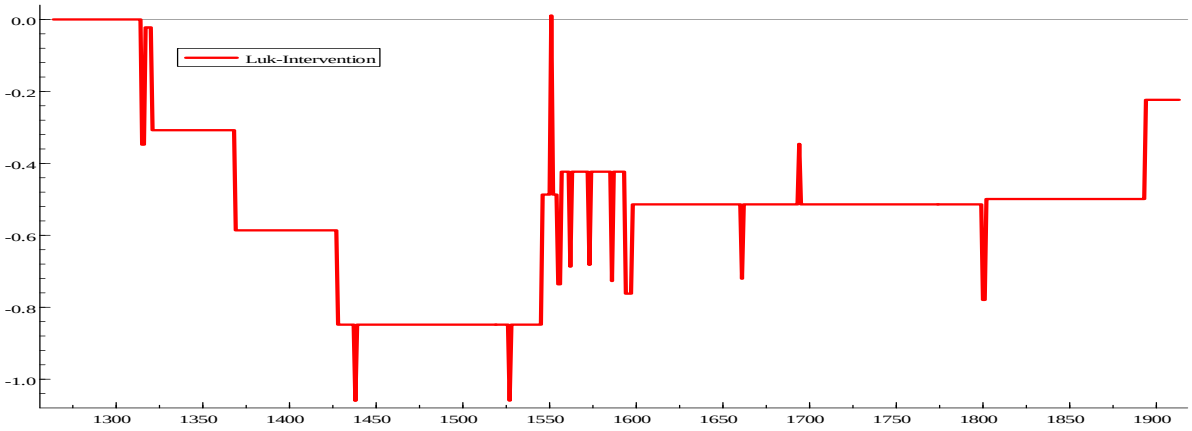
4.5.2. Trend is a random walk with fixed drift plus interventions

This formulation adds the use of interventions to the previous model to capture outliers. In STAMP interventions can be user defined or automated. Automation involves the use of the standardised smoothed estimates of the disturbances referred to as the *auxiliary residuals*. The auxiliary residuals are smoothed estimates of the irregular and level disturbances. Graphs of these residuals, in conjunction with normality tests, are used for detecting data irregularities such as *outliers*, *level changes* and *slope changes*. Here an outlier is an unusually large value of the *irregular* disturbance at a point in time which can be handled via an impulse intervention variable which takes a value 1 at that point, 0 otherwise. In contrast a *structural break in the level* is captured by a shift up or down as a

step intervention, 0 before the event and 1 subsequently, and a structural break in the slope as a staircase intervention, 1, 2, 3,….., starting after the detected break.

Figure 5 below plots the timing of the interventions as selected automatically by STAMP. The adding of the outliers leads to considerable improvement in goodness of fit; now the Durbin-Watson statistic=1.90; and R^2=0.40. It should be noted this model retains the assumption of a fixed trend, and the level and outlier interventions may reflect these types of interventions are attempting to fit the data when the underlying model choice is too constraining. The volume of level breaks shown by this model for the 16th century appears excessive, and casts doubt on whether or not outlier and levels interventions are appropriately identified.

Figure 5.
Timing and location of interventions



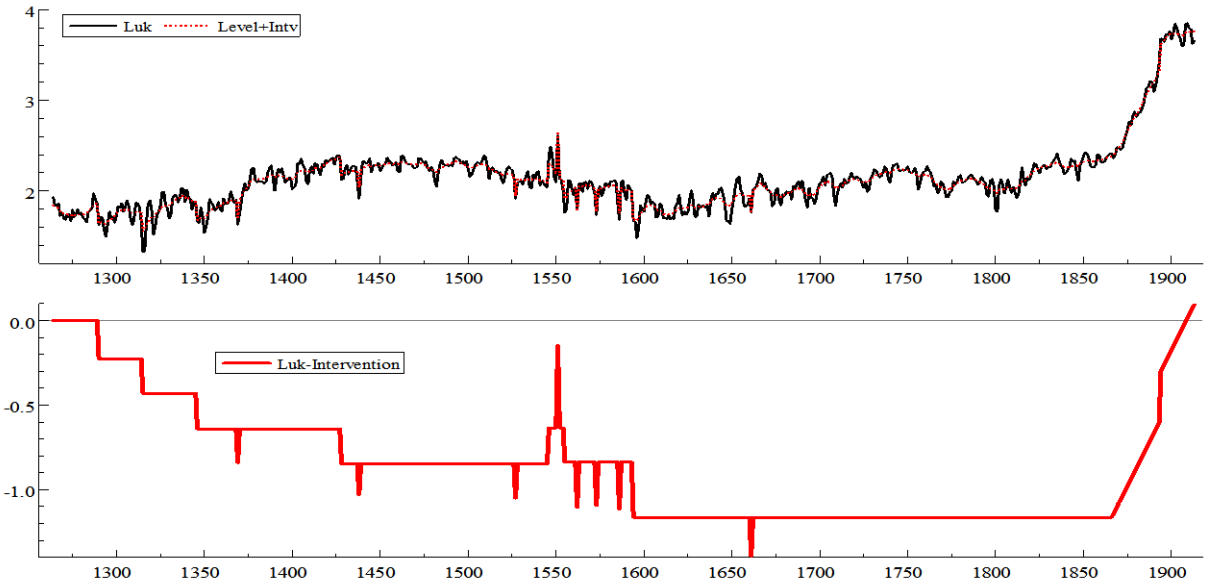
Notes: the outlier and level interventions are: Outliers 1438, 1527, 1551, 1562, 1573, 1586, 1661, 1694; Levels breaks 1315, 1317, 1321, 1369, 1428, 1546, 1555, 1557, 1594, 1598, 1800, 1802, 1894.

4.5.3. A local linear model: the level is stochastic; slope stochastic, plus 5 year cycle plus irregular plus interventions

This model relaxes the assumption of a fixed slope, and shows a further improvement in specification; the Durbin-Watson statistic=1.88; and R^2=0.43. Figure 6 below plots the

timing of the interventions as selected automatically by STAMP. They include a single trend break at 1867, which suggests that Industrial Revolution technology's impact of real wages was long delayed, and interrupted by a level shift around 1894. Overall the model shows fewer levels breaks than 4.6.2, and thus a representation of the long run data the accords more closely with the historiography. Between 1594 and 1867 the model shows no structural breaks in real wages and a single outlier in 1661. The long stagnation of English real wages following the higher levels during the century of population collapse to 1450 shown by this variant of the SMT model provides a plausible interpretation of English real wages history.

Figure 6



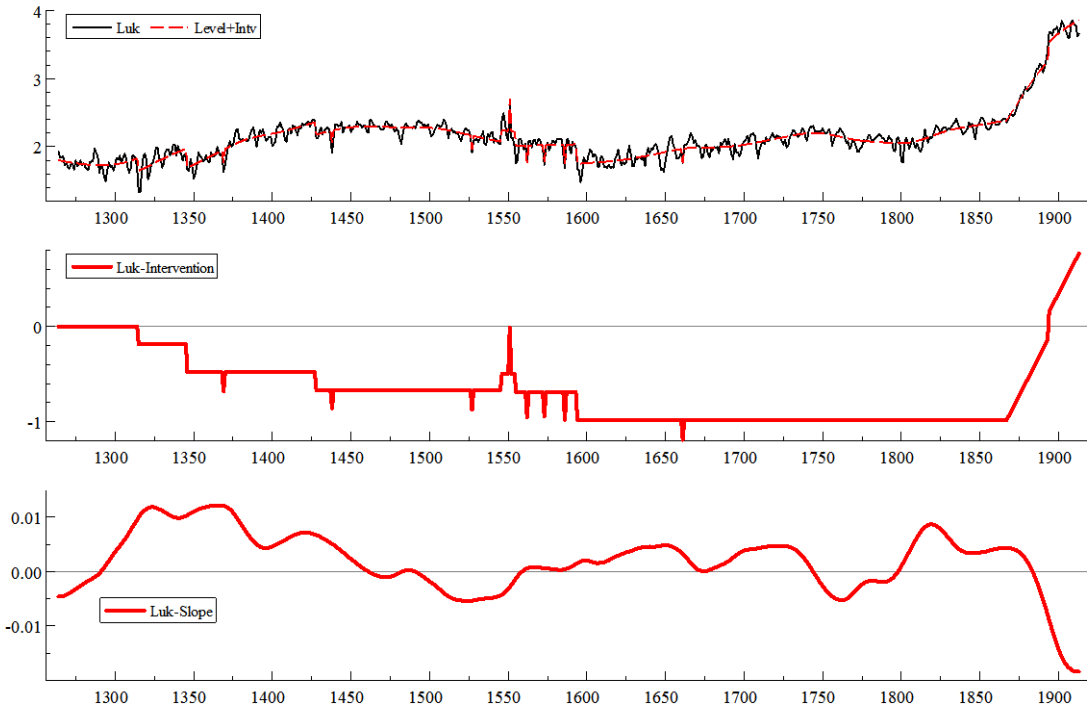
Notes: the interventions are: Outliers 1369, 1438, 1527, 1551, 1562, 1573, 1586, 1661; Levels breaks 1290, 1315, 1346, 1428, 1546, 1555, 1594, 1894; Slope breaks 1867.

4.5.4. Fixed level; slope stochastic, plus 5 year cycle plus irregular plus interventions

This variant retains the stochastic trend assumption, but adopts a fixed level representation. The model is shown inferior to that which incorporates a stochastic level,

with Durbin-Watson statistic=1.76; and R^2=0.40. However this representation also shows a long stagnation of real wages 1594-1868.

Figure 7

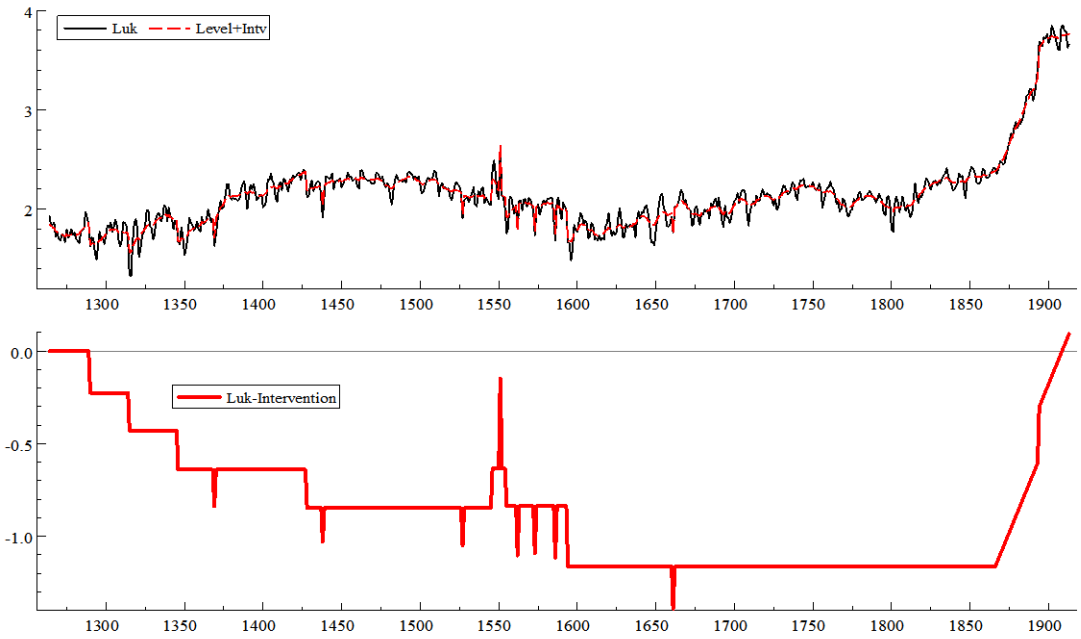


Notes: the interventions are: Outliers, 1369, 1438, 1527, 1551, 1562, 1573, 1586, 1661; Levels breaks 1315, 1346, 1428, 1546, 1555, 1594, 1894; Slope break 1868.

4.5.5 Stochastic level; Stochastic trend; 20 year cycle; irregular plus interventions

This variant replicates model 4.6.3 but adopts a 20-year cycle, which improves the goodness of fit, with Durbin-Watson statistic=1.87; and R^2=0.52. The timing of all interventions are the same as for model 4.6.3, including the single trend break in 1867.

Figure 8



Notes: the interventions are: Outliers 1369, 1438, 1527, 1551, 1562, 1573, 1586, 1661; Levels breaks 1290, 1315, 1346, 1428, 1546, 1555, 1594, 1894; Slope breaks 1867.

By using a sequence of SMT modelling strategies this section illustrates how English real wages 1264-1913 can be represented, by adopting varying assumptions surrounding trends, levels, irregular outliers, and cycles. The interventions identified by STAMP shed light on real wages history, but also in the plausibility of the alternative models. The flexible models which allow both the level and slope to be stochastic, and include cyclical elements provide the best representations of the data. The results from these simple illustrations conflict with elements of the published literature. For example the unified theory of Galor and Weil (2000) postulates a two break model where incomes per capita accelerate modestly around the Industrial Revolution of the late 18th century and more dramatically around the later 19th century demographic transition. The preferred

results here, in contrast, show a long period of stagnant real wages 1594-1867. Moreover Crafts and Mills (2009) utilizing Clark's (2005) alternative real wages find a trend break around 1800 in conflict with the results from the SMT model which show a later break of 1867.

4.6. Multiple Changes in Persistence and English Real Wages 1264-1913.

An alternative way to investigate long run trends in English real wages is to consider if the series have alternating stochastic properties. In section 4.3.1 above we considered the test of Leybourne, Kim and Taylor (2007) applied to British industrial production data, 1700-1992. The same test applied to the real wages data considered in section 4.6 suggests the following: I(0) 1264 – 1858, with effectively a zero trend then I(1) to 1913. The results, however, also suggest the potential for a short return to trend stationarity for the period 1881-1894, but the sample size is too small to confirm this categorically.

5.0 Some new developments/applications with potential for cliometrics

In Section 5 we introduce some new and emerging methods which we believe will be of importance to cliometrics research in the future. In particular we will discuss; the *mildly explosive processes* of Phillips and Yu (2009); *Graphical Modelling*; and although not totally new to cliometrics, a discussion of *fractionally integrated processes* and their *long memory* interpretations.

5.1 Mildly explosive processes – Phillips and Yu (2009)

Phillips and Yu (2009) have developed a new econometric methodology to test if and when bubbles emerge and collapse and apply it in various stock markets, real estate markets, mortgage markets, commodity markets, and the foreign exchange market over the period surrounding the subprime crisis.

The basis of their new approach is to consider *mildly explosive processes*. If we consider the typical DS v TS testing procedures for a unit root, we restrict our attention to regions of 'no more than' a unit root process – an autoregressive process where $\rho \leq 1$. Phillips and Yu (2009) model 'mildly explosive' behavior by an autoregressive process with a root ρ that exceeds unity, but still in the neighbourhood of unity.

The basic idea of their approach is to recursively calculate right-sided unit root tests of a standard form, eg., Dickey-Fuller-type, to assess evidence for mildly explosive behavior in the data. The test is a right-sided test and therefore differs from the usual left-sided tests for stationarity. More specifically, consider the following autoregressive specification estimated by recursive least squares:

$$X_t = \mu + \delta X_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim NID(0, \sigma^2) \qquad (48)$$

The usual H_0: $\delta = 1$ applies, but unlike the left-sided tests which have relevance for a stationary alternative, Phillips and Yu (2009) have H_1: $\delta > 1$ which with $\delta = 1 + c/kn$, where $kn \to \infty$ and $kn/n \to 0$ allows for their 'mildly explosive' cases. Phillips and Yu (2009) argue that their tests have discriminatory power because,

> "They are sensitive to the changes that occur when a process undergoes a change from a unit root to a mildly explosive root or vice versa. This sensitivity is much greater than in left-sided unit root tests against stationary

> alternatives.....Although a unit root process can generate successive upward movements, these movements still have a random wandering quality unlike those of a stochastically explosive process where there is a distinct nonlinearity in movement and little bias in the estimation of the autoregressive coefficient."

We believe these new approaches to identifying growing bubbles and their collapse will make a significant impact on this aspect of time series applied econometrics and because of the nature of the data and the questions on financial cliometrics.

5.2 Graphical modelling and implications for causality testing

Graphical Modelling (GM) is a relatively new statistical approach whose major development started in the 1970s. It is a very convenient interface to obtain and crucially present in a graphical way, pairwise relationships among random variables in a multivariate context. It has important links to causality and testing for causality and its manifest output makes it much easier to understand the linkages between variables of interest. Below, we will present some of the basis of the approach, some references to the technical details for those interested, some past applications and current developments. We believe the approach has enormous potential for cliometricans in the future, in particular, its ability to easily – graphically – present results on causality and exogeneity.

The initial step in the GM approach is the computation of the partial correlations among the variables in the multivariate system where this can be achieved by inverting and rescaling the covariance matrix as suggested by Whittaker (1990). With these computations complete, we can then distinguish between significant and non-significant

partial correlations using an opportune test. Finally we can present the results by a graph where the random variables are represented by *nodes* and a significant partial correlation between two random variables is indicated by a line that links them. The line in graph theory terminology is called and *edge*. If the variables in the graph are jointly distributed as a multivariate Gaussian distribution a significant partial correlation implies the presence of conditional dependence. For this reason the graph described by these conditions is called a *conditional independence graph* or CIG.

A more informative object is the *directed acyclic graph* or DAG. This is a directed graph where the arrows linking the nodes are where the joint distribution of the variables can be expressed as a sequence of marginal conditional distributions. For example, consider the graph in Figure 9 below.

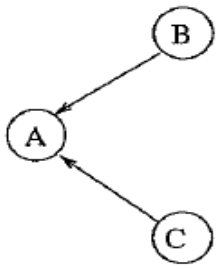


Figure 9: Directed Acyclic Graph

Its joint density function can be defined as: $f(a,b,c) = f(a \mid b,c) f(b) f(c)$

Although the DAG and the CIG represent a different definition of the joint probability, there is a correspondence between the two graphs which is embodied by the *moralisation rule*. The rule means that we can obtain the CIG from the DAG by transforming the arrow into lines linking unlinked parents. Consider by way of example Figure 4 below.

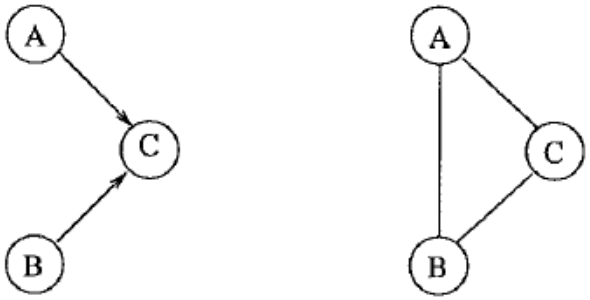


Figure 10: Moralisation of a Directed Acyclic Graph.

A and B are called the *parents* of C. The moralisation of the DAG on the left is obtained by transforming the existing arrows into edges and by adding an edge which links the parents. These edges are called *moral edges*.

Importantly, while the CIG represents the associations among the variables either in terms of conditional dependence or simply in terms of partial correlations, if the joint distribution is not Gaussian, the DAG has a natural interpretation in terms of *causality*. For those wishing to consider more in this area of graphs and causality we refer you to; Shafer 91996); Glymour and Cooper (1999); Lauritzen (2000); Pearl (2000) and Lauritzen and Richardson (2002).

The DAG is very attractive because of its causal interpretation, but all we can observe in practise is the CIG obtained by the sample partial correlation. Therefore, we need to perform the inverse to the moralisation, the *demoralisation*. While the transformation of a DAG to a CIG is unique, there are several DAGs which might give the same CIG. In this case we need to identify the moral links and remove them and to do that we need to use all the knowledge we have about the relationships among the random variables in the system.

In Oxley, Reale and Tunnicliffe-Wilson (2009), we apply these methods to identifying an interest rate transmission mechanism for New Zealand and the sort of graphical interpretations, which in the paper are compared to more traditional Structural VAR (SVAR) models look like the figures below, where the A's – H's relate to a range of increasing in maturity interest rates, both domestic and foreign (US):

Figure 11: Conditional Independence Graph
(a range of quarterly NZ and foreign interest rates, 1987-2001)

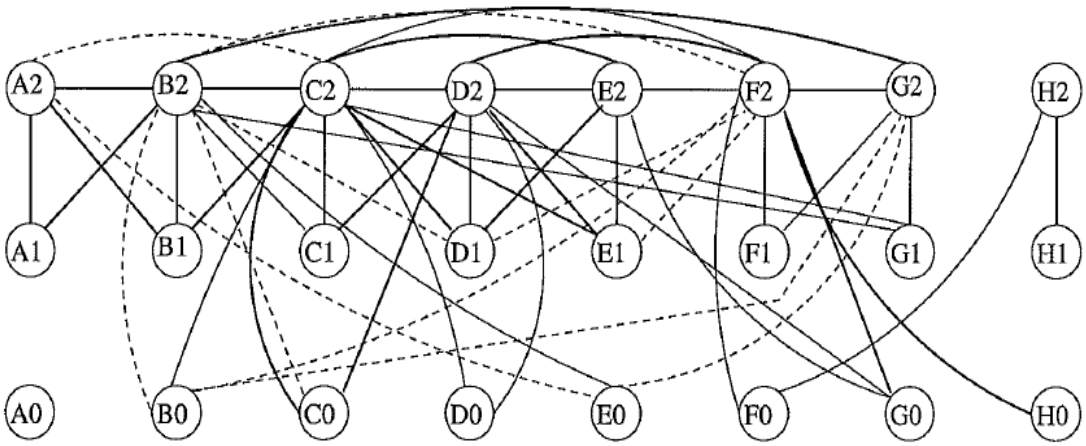
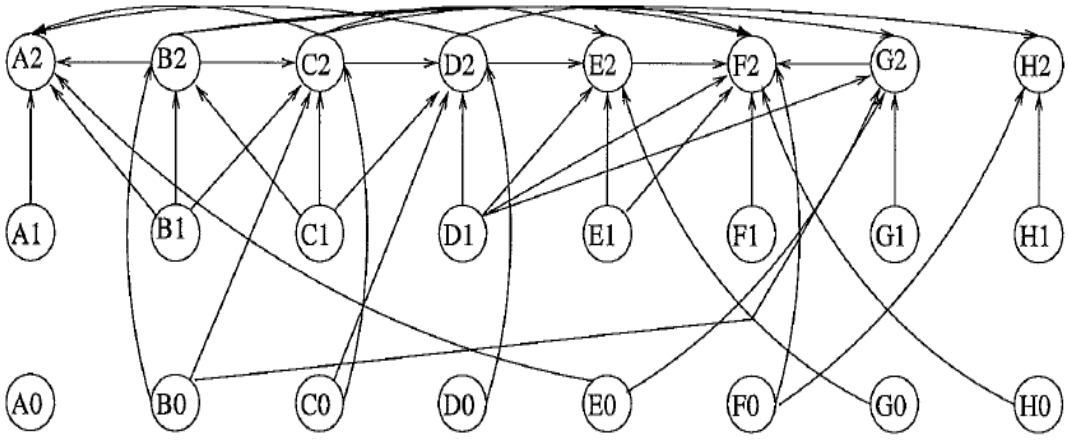


Figure 12: Chosen model representation.



5.3 Fractional integration and long memory

In section 3 above, we discussed the implications for the persistence of shocks in the DS and TS data generating worlds. In the former case, with a unit root, shocks to the I(1) process would have infinite persistence – in the later mean I(0) case with mean reversion, shocks have zero persistence.

However, what about cases where the order of integration is a fraction d >0 and <1? In this case we have a case of *fractional integration,* or in the discrete case, ARFIMA (*Auto Regressive, Fractionally Integrated Moving Average*) and the degree of persistence will depend on the size of *d*. With 0<d<1 we have processes that are typically described as having *long memory* or *long range dependence*. When -1/2<d<0 we describe the process as *anti-persistent.*

Some of the original work on such processes was undertaken by Hurst (1951) studying the Nile river, but in economics the processes were popularised by Granger and Joyeux (1980) with other major contributions by Robinson (1995, 2003), Beran (1992) and Hosking (1981). The importance of this class of processes derives from smoothly bridging the gap between short memory stationary processes and unit roots in an environment that maintains a greater degree of continuity (Robinson, 1994). For an up-to-date survey see Gil-Alana and Hualde (2009).

In economics and finance it appears common to see estimates of $d \cong 0.4$, implying significant long memory (but not infinite persistence) in the data considered. In economic history, Mishra, Prskawetz, Parhi and Diebolt (2009) argue that long memory in economic growth occurs because of stochastic memory in population growth. They estimate *d* using the Kim and Phillips (2000) modified log-periodgram estimator of Geweke and Porter

Hudak (1983b) to 'provide evidence of fractionally integrated population growth with non-mean convergent shock dynamics ... in 63 countries from 1950 – 2004'. Michelacci and Zalaroni (2000) also consider long memory issues, also in relation to economic growth. They present estimates of *d* based on the log-periodogram estimator of Geweke and Porter Hudak (1983b) and conclude that there is evidence of long memory and present a case for fractional (beta) convergence extending the ideas of the Solow-Swan growth model.

Several potential issues arise when considering long memory models. The first is whether long memory processes make sense. In finance, Rea *et. al.,* (2008a) argue it makes little sense to consider long-memory or long range dependence in finance theories such as, option pricing models. In the case of economic growth or population, however, long range dependence may make sense. Secondly, however, long memory and structural change are 'observationally equivalent' such that either assumptions can equally describe the data see Rea *et. al.,* (2008a). Finally and crucially, 'not all estimators of d are born equal', especially in small samples see Rea *et. al.,* (2008b).

Of the twelve estimators examined by Rea, W., Oxley, L. Reale, M. and Mendes, E. (2008b) the Whittle estimator and Haslett-Raftery (1989) estimators performed the best on simulated series. If we require an estimator to be close to unbiased across the full range of *d* values for which long memory occurs and have a 95 percent confidence interval width of less than 0.1 *d* units (that is 20 percent of the range for *d* values in which long memory is observed), then for series with fewer than 4,000 data points Whittle and Haslett-Raftery are the only two estimators worth considering.. For series with 4,000 or more data points, the Peng, *et. al.,* (1994) estimator gave acceptable performance. For series with more than 7,000 data points the periodogram estimator was a worthwhile

choice. For series with more than 8,200 data points the wavelet became a viable estimator. The remaining seven estimators they considered did not give acceptable performance at any series lengths examined and are not recommended. If you wish to conclude that long memory processes exist, be sure that the estimator used is 'fit for purpose' as the typical sample sizes, even in long run cliometric applications are very, very short compared to what is needed for efficient and unbiased estimation.

6.0 Epilogue

Keynes reminds us that 'Practical men, who believe themselves to be quite exempt from any intellectual influence, are usually the slaves of some *defunct economist*.' Keeping the slavishly following analogy, we wish to reiterate one quote and pass-on a message which represents a significant warning about delving into the world of time series econometrics and relying, slavishly, on the outcome of tests without recourse to history.

The quote is where we opened this paper: *"The power of a popular test is irrelevant. A test that is never used has zero power"*(McAleer, 1994, 2005). Although David Hendry might say, 'test, test, test', not all tests are born equal and both type I and type II (or even type III, see Kennedy 2002) pervade empirical work.

The message comes from the genius that was Sir Clive Granger. In Granger (unpublished) he considers a number of so called 'puzzles' in economics and in many cases links them to his famous notion of spurious regressions. However, he also discusses a case where he and a colleague considered the potential for *reintegration* and why relying on the results of statistical tests (especially those with known low power) can

be extremely dangerous. To demonstrate the point consider Granger's example from that paper.

In the case of cointegration, the sum of two I(1) series produces a linear combination that is I(0). The complete opposite case would have I(0)+I(0)=I(1). From a theoretical perspective this is impossible and Granger obviously knew this, however Granger and Jeon (unpublished) explored the case empirically. Assume that X_t is a zero mean white noise process with large variance and let $Y_t=Z_t+X_t$ where Z_t is a random walk with white noise inputs having a zero mean and small variance. In their empirical examples with 500 observations standard unit root tests concluded that X and Y were I(0), but their sum (Z_t) was I(1). The testing process, with standard tests, suggested it was possible to sum two I(0)'s and obtain an I(1) series! The moral here: never state 'the test proves/demonstrates the data are I(1) or I(0) or I(anything)'. Likewise never forget some of the messages of Kennedy (2002) i.e., Rule #1: Use common sense and economic theory; Rule #3: Know the context; Rule #7: Understand the costs and benefits of data mining and Rule #9: Do not confuse statistical significance with meaningful magnitude.

7.0 Conclusions

Cliometrics has been with us for half a century and its original hallmarks were the links between economics and economic history see Greasley and Oxley (2010b). However, the use of econometric methods has become more common in quantitative economic history. Understanding modern time series methods and their application, therefore becomes much more important in cliometrics than perhaps ever in the past. The developments in time series econometrics, which started around 1974, have radically changed what

cliometricians do, or should do. Granger and Newbold (1974), reminded us of the dangers of spurious regression; Dickey and Fuller (1979) began a research agenda on unit root testing which remains unresolved and provided tests which, although lacking power, are as popularly applied as ever. Engle and Granger (1987), gave us co- integration and the rest, as they say is history.

8.0 Acknowledgements

We wish to acknowledge the numerous, wide ranging and stimulating conversations on time series econometrics with the late Sir Clive Granger – his brilliance will be deeply missed. Section 4.1 benefited from the research assistance of Elizabeth Duston. The second author wishes to acknowledge financial support from the Marsden Fund.

9.0 References

Allen, R.C, and Keay, I. (2001) The First Great Whale Extinction: The End of the Bowhead Whale in the Eastern Arctic *Explorations in Economic History,* 38, 448-477.

Allen, R.C. (2001). The Great Divergence in European Wages and Prices from the Middle Ages to the First World War, *Explorations in Economic History*, 38, 411-447

Antras, P., and Voth, H-J. (2003) Factor prices and productivity growth during the British Industrial Revolution. *Explorations in Economic History,* 40, 52-77.

Baten, J., and Murray, J.E. (2000) Heights of Men and Women in 19th-Century Bavaria: Economic, Nutritional, and Disease Influences *Explorations in Economic History*, 37, 351-369.

Baubeau , P., and Cazelles , B. (2009) French economic cycles: a wavelet analysis of French retrospective GNP series *Cliometrica ,* 3, 275-300.

Ben-David, D., and Papell, D. (1998). Slowdowns And Meltdowns: Postwar Growth Evidence From 74 Countries, *The Review of Economics and Statistics,* 80, 561-571

Beran, J. (1992). A goodness-of-fit test for time series with long range dependence, *Journal of the Royal Statistical Society B* 54 , 749–760

Bernard, A., and Durlauf, S., (1995), Convergence in international output, *Journal of Applied Econometrics*, 10, pp. 97-108

Bernard, A., and Durlauf, S., (1996), Interpreting tests of the convergence hypothesis, *Journal of Econometrics,* 71, pp. 161-73.

Beveridge, S., and Nelson, C.R. (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, 151-174.

Bhargava, A. (1986). On the theory of testing for unit roots in observed time series, *Review of Economic Studies,* 52, 369-384.

Bodenhorn, H. (2007). Usury ceilings and bank lending behavior: Evidence from nineteenth century New York *Explorations in Economic History,* 44, 179-202.

Bollerslev, T. (1986). Generalised autoregressive conditional heteroskedasticity, *Journal of econometrics,* 31, 307-327.

Breitung, J. (2000). The local power of some unit root tests for panel data. *Advances in Econometrics, Volume 15: Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, ed. B. H. Baltagi, 161–178. Amsterdam: JAY Press.

Burhop, C. (2006). Did banks cause the German industrialization? *Explorations in Economic History,* 43, 39-63.

Campbell, J., and Mankiw, G. (1987). Are Output Fluctuations Transitory?, *The Quarterly Journal of Economics*, 102, 857-80

Capie, F., and Mills, T. (1991). Money and business output in the US and UK, 1870-193, *The Manchester School*, 59, 38-56.

Chatfield, C. (1989) *The Analysis of Time Series*, Chapman Hall: London.

Chenery, H.B., Watanabe, T., (1958). International comparisons of the structure of production. *Econometrica* 29, 487–521.

Cheung, Y-W and Lai, K S. (1995). Lag Order and Critical Values of the Augmented Dickey-Fuller Test, *Journal of Business & Economic Statistics*, 13, 77-280

Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Finance* 20: 249–272.

Chow, P.C.Y. (1987), Causality Between Export Growth and Industrial Performance: Evidence from the NIC's, *Journal of Development Economics,* **26**, 55-63.

Ciccarelli, C. and Fenoaltea, F. (2007) Business fluctuations in Italy, 1861–1913: The new evidence *Explorations in Economic History*, 44, 432-451.

Clark, G., 2005. The condition of the working class in England, 1209 to 2004. *Journal of Political Economy* 113, 1307–1340.

Cochrane, J. (1988). How big id the random walk in GNP, *Journal of Political Economy,* 96, 893-920.

Crafts, N. F. R and Harley, CK. (1992) Output growth and the British Industrial Revolution: A restatement of the Crafts-Harley view, *Economic History Review* 45, 703-30.

Crafts, N. F. R and T. C. Mills (2009), From Malthus to Solow: How did the Malthusian economy really evolve, *Journal of Macroeconomics,* 31, 68-93.

Crafts, N. F. R and T. C. Mills (2009), 'From Malthus to Solow: How did the Malthusian economy really evolve', *Journal of Macroeconomics*, 31, 68-93.

Crafts, N. F. R. and Mills, T. C. (1994) Trends in real wages in Britain, 1750-1913, *Explorations in Economic History*, **31**, 176- 94.

Crafts, N., Mills, T.C., and Mulatu, A. (2007) Total factor productivity growth on Britain's railways, 1852–1912: A reappraisal of the evidence *Explorations in Economic History,* 44, 608-634.

Crafts, N., and Mills, T.C. (2004) Was 19th century British growth steam-powered?: the climacteric revisited *Explorations in Economic History,* 41, 156-171.

Crafts, N.C.R., Leybourne, S.J., and Mills, T.C. (1989,a), Trends and Cycles in British Industrial Production 1700-1913, *Journal of the Royal Statistical Society*, Vol 152, No. 1, pp. 43-60.

Craig, L., and Holt, M. (2008). Mechanical refrigeration, seasonality, and the hog–corn cycle in the United States: 1870–1940 *Explorations in Economic History,* 45, 30-50.

Dahmen, E., (1988). Development blocks in industrial economics. *The Scandinavian Economic History Review* 36, 3–14.

Darne, O. (2009). The uncertain unit root in real GNP: A re-examination, *Journal of Macroeconomics*, 31, 153–166.

Deane, P. and Cole, W. A. (1969) *British Economic Growth*, 1688-1959, Cambridge University Press, Cambridge.

della Paolera, G., and Taylor, A.L. (2002) Internal versus external convertibility and emerging market crises: lessons from Argentine history. *Explorations in Economic History*, 39, 357-389.

Dickey, D.A. and Fuller, W.A. (1979), Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association,* 74, 427-31.

Dickey, D.A. and Fuller, W.A. (1981), Likelihood Ratio Tests for Autoregressive Time Series with a Unit Root, *Econometrica*, Vol. 49, No. 4, pp.1057-72.

Diebolt, C., and Parent, A. (2008). Bimetallism: The "rules of the game"*Explorations in Economic History,* 45, 288-302.

Eichengreen, B., and Hataseb, M. (2007). Can a rapidly growing export-oriented economy exit smoothly from a currency peg? Lessons from Japan's high-growth era *Explorations in Economic History,* 44, 501-521.

Elliott, G., Rothenburg, T., and Stock, J. (1996). Efficient tests of an autoregressive unit root. *Econometrica,* 64, 813-836.

Enflo, K., A. Kander and L. Schon (2009). Identifying Development Blocks – a new methodology implemented on Swedish industry. *Journal of Evolutionary Economics* 18, 57-76.

Engle, R., and Granger, C.W.J., (1987), Co-integration and Error Correction: Representation, Estimation and Testing, *Econometrica,* 55, 251-76.

Esteves, R.P., Reis, J., and Ferramosca, F. (2009). Market Integration in the Golden Periphery. The Lisbon/London Exchange, 1854–1891 *Explorations in Economic History,* 46, 324-345.

EViews, 6 and 7. QMS: Quantitative Micro Software, http://www.eviews.com/index.html

Federico, G. (2007). Market integration and market efficiency: The case of 19th century Italy *Explorations in Economic History,* 44, 293-316.

Franck, R., and Krausz, M. (2009). Institutional changes, wars and stock market risk in an emerging economy: evidence from the Israeli stock exchange, 1945–1960 *Cliometrica,* 3, 141-164.

Fratianni, M. and Spinelli, F. (2001). Fiscal Dominance and Money Growth in Italy: The Long record *Explorations in Economic History,* 38, 252-272.

Gaffeo E., Gallegati, M. And Gallegati, M. (2005). Requiem for the unit root in per capita real GDP? Additional evidence from historical data, *Empirical Economics*, 30, 37-63

Galor, O., (2005), 'From Stagnation to Growth: Unified Growth Theory', in P. Aghion and S. N. Durlauf eds., *Handbook of Economic Growth*, vol 1A, 171-285, Elsevier, Amsterdam.

Galor, O., and Weil, G. (2000) Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond, *American Economic Review*, 90, 806-828

Geweke, J., and Porter-Hudak, S. (1983a). The stimation and application of long memory series models, *Journal of Time Series Analysis,* Cambridge: Cambridge University Press.

Geweke, J., R. Meese and W.T. Dent (1983b), Comparing Alternative Tests of Causality in Temporal Systems: Analytic Results and Experimental Evidence, *Journal of Econometrics,* 21, 161-94.

Giles D.E.A., J.A. Giles and E. McCann (1993), Causality, Unit Roots and Export-Led Growth: The New Zealand Experience, *Journal of International Trade and Economic Development*, 1**,** 195-218

Gils-Alana, L and Hualde, J. (2009) Fractional integration and cointegration : An overview and empirical application, in Mills, T. and Patterson, K. Palgrave Handbook of Econometrics, Vol. 2.

Glymour, C., and Cooper, G.F. (1999). *Computation, causation and discovery*. MIT Press: Cambridge.

Goodwin, B.K., Grennes, T., and Craig, L. (2002) Mechanical Refrigeration and the Integration of Perishable Commodity Markets. *Explorations in Economic History*, 39, 154-182.

Granger C.W.J. (undated). Useful conclusions from surprising results. Typescript.

Granger, C.W.J. (1969). Investigating Causal Relations by Econometric Models and Cross Spectral Methods, *Econometrica,* 37, 424-38

Granger, C.W.J. (1986) Developments in the Study of Co-integrated Economic Variables, *Oxford Bulletin of Economics and Statistics,* 48, 213-28.

Granger, C.W.J. (1988) Some Recent Developments in a Concept: Causality, *Journal of Econometrics,* 39, 199-211.

Granger, C.W.J. and Jeon, Y. (unpublished), "untitled"

Granger, C.W.J. Hyung, N., and Yeon, Y. (2001). Spurious regression with stationary series, *Applied Economics,* 33, 899-904.

Granger, C.W.J. Joyeux, R. (1980). An introduction to long-range time series models and fractional differencing, Journal of Time Series Analysis 1, 15–30.

Granger, CWJ. and Newbold, P. (1974). Spurious regression in econometrics, *Journal of Econometrics,* 2, 111-120.

Grant, O. (2003) Globalisation versus de-coupling: German emigration and the evolution of the Atlantic labour market 1870–1913. *Explorations in Economic History,* 40, 387-418.

Greasley, D., Madsen, J. And Oxley, L. (2001) Income Uncertainty and Consumer Spending during the Great Depression. *Explorations in Economic History,* 38, 225-251.

Greasley, D., and Oxley, L. (1994a) Rehabilitation Sustained: The Industrial Revolution as a Macroeconomic Epoch, *Economic History Review*, 47, 760-68.
Greasley, D., and Oxley, L. (1994b). Structural change and unit root testing: British Industrial Production 1700-1913 *Applied Economics Letters*, 1994, 1, 29-30.
Greasley, D., and Oxley, L. (1995). A Time Series Perspective on Convergence: Australia UK and the US Since 1870. *Economic Record*, 71, 259-70.
Greasley, D., and Oxley, L. (1996a), Discontinuities in Competitiveness: The Impact of World War One on British Industry, *Economic History Review* XLIX, 1996, pp. 83-101
Greasley, D., and Oxley, L. (1996b), Technological Epochs and British Industrial Production, 1700-1992, *Scottish Journal of Political Economy*, 43. 258-74.
Greasley, D., and Oxley, L. (1997a). Time-series based tests of the convergence hypothesis: some positive results. *Economic Letters* 56, 143–147.
Greasley, D., and Oxley, L. (1997b). Unit roots and British industrial growth, 1922-1992, *Manchester School*, 65, 192-212.
Greasley, D., and Oxley, L. (1997c), Endogenous Growth or Big Bang: Two Views of the First Industrial Revolution, *Journal of Economic History*, 57, 935-49.
Greasley, D., and Oxley, L. (1997d). Shock persistence and structural change. *The Economic Record* 73, 348-62
Greasley, D., and Oxley, L. (1997e), Causality and the First Industrial Revolution, *Industrial and Corporate Change*, 7, 33-47.
Greasley, D., and Oxley, L. (1997f). Endogenous growth, trend output and the Industrial Revolution. *Journal of Economic History*, 957-960.
Greasley, D., and Oxley, L. (1997g). Convergence in GDP per capita and real wages: some results for Australia and the UK. *Mathematics and Computers in Simulation*, 43, 429-36.
Greasley, D., and Oxley, L. (1998a), Vector Autoregression, Cointegration and Causality: Testing for Causes of the British Industrial Revolution, *Applied Economics*, 30, 1387-97.
Greasley, D., and Oxley, L. (1998b). Causality and the First Industrial Revolution, *Industrial and Corporate Change* , 7, 33-47.
Greasley, D., and Oxley, L. (1998c). A Tale of two Dominions: Australia, Canada and the Convergence Hypothesis. *Economic History Review*, 51, 294-318.
Greasley, D., and Oxley, L. (1999). International comparisons on shock persistence in the presence of structural change and non-linearities. *Applied Economics*, 31, 499-507
Greasley, D., and Oxley, L. (2000). British Industrialization, 1815–1860: A Disaggregate Time-Series Perspective, *Explorations in Economic History*, 37, 98-119.
Greasley, D., and Oxley, L. (2004) Globalization and real wages in New Zealand 1873–1913. *Explorations in Economic History*, 41, 26-47.
Greasley, D., and Oxley, L. (2010a). Knowledge, Natural Resource Abundance and Economic Development: Lessons from New Zealand 1861-1939', forthcoming *Explorations in Economic History*
Greasley, D., and Oxley, L. (2010b). Clio and the Economist: Making Historians Count, *Journal of Economic Surveys*, this issue.
Greasley, D., Madsen J., and Wohar, M. (2010). Long-run Growth Empirics and Unified Theory, typescript.

Grossman, R., and Imai, M. (2009). Japan's return to gold: Turning points in the value of the yen during the 1920s *Explorations in Economic History,* 46, 314-323.
Grubb, F. (2004) The circulating medium of exchange in colonial Pennsylvania, 1729–1775: new estimates of monetary composition, performance, and economic growth. *Explorations in Economic History,* 41, 329-360.
Hadri, K. (2000). Testing for stationarity in heterogeneous panel data. *Econometrics Journal* 3, 148-161.
Hansen, B.E. (1992). Testing for parameter instability in linear models, *Journal of Policy Modeling,* 14, 517-533.
Harvey , A. and Trimbur, T. (2003). General Model-Based Filters for Extracting Cycles and Trends in Economic Time Series, *The Review of Economics and Statistics,* 85, 244-255.
Harvey, A. (1989). *Forecasting Structural Time Series Models and the Kalman Filter,* Cambridge: Cambridge University Press.
Harvey, A. and Jaeger, A., (1993), Detrending, Stylized Facts and the Business Cycle, *Journal of Applied Econometrics,* 8**,** 231-47.
Harvey, A.C. and Koopman, S.J., (1992). Diagnostic checking of unobserved components time series models, *Journal of Business and Economic Statistics*, 10, 377-389.
Haslett, J. and Raftery, A. E. (1989). Space-time Modelling with Long-memory Dependence: Assessing Ireland's Wind Power Resource (with Discussion). *Applied Statistics*, 38, 1-50.
Hatton, T. J. and Lyons, J. S. (1983) Eighteenth Century British trade: homespun or empire made, *Explorations in Economic History*, 20, 163- 82
Herranz-Locan , A. (2007) Infrastructure investment and Spanish economic growth, 1850–1935 *Explorations in Economic History,* 44, 452-468.
Hickson, C., Turner, J., and McCann, C. (2005). *Much ado about nothing*: the limitation of liability and the market for 19th century Irish bank stock *Explorations in Economic History,* 42, 459-476.
Hirschman, A.O., (1958). *The Strategy of Economic Development*. Yale University Presss, New Haven CT.
Hodrick, R., and Prescott, E. (1977). Postwar US business cycles: An empirical investigation. *Journal of Money, Credit and Banking,* 29, 1-16.
Hoffmann, W. H. (1955), *British Industry 1700-1950*, Oxford: Blackwell.
Horrell, S., Humphries, J. and Weale, M. (1994), 'An Input-Output Table for 1841', *Economic History Review*, 47, 545-66
Hosking, J.R.M., (1981). Fractional differencing, *Biometrika* 68, 165–176.
Hsiao, C. (1987), Tests of Causality and Exogeneity Between Exports and Economic Growth: The Case of the Asian NIC's, *Journal of Economic Development,* **12,** 143-59.
Hurst, H.E., (1951). Long-term storage capacity of reservoirs, *Transactions of the American Society of Civil Engineers* 116, 770–808.
Im, K.S., Pesaran, M.H., Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115, 53-74
Jacks, D. (2006). What drove 19th century commodity market integration? *Explorations in Economic History,* 43, 383-412.

Jacks, D. (2007) Populists versus theorists: Futures markets and the volatility of prices *Explorations in Economic History*, 44, 342-362.

Johansen, S. (1988), Statistical Analysis of Cointegrated Vectors, *Journal of Economic Dynamics and Control*, 12, 231-54.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford: Oxford university Press.

Jung, W.S. and P.J. Marshall (1985), Exports, Growth and Causality in Developing Countries, *Journal of Development Economics*, 18**,** 1-12

Keay, I. (2000) Scapegoats or Responsive Entrepreneurs: Canadian Manufacturers, 1907–1990, *Explorations in Economic History*, 37, 217-240.

Keay, I., and Redish, A. (2004) The micro-economic effects of financial market structure: evidence from 20th century North American steel firms. *Explorations in Economic History*, 41, 377-403.

Kendall, M. (1954). *Exercises in Theoretical Statistics* London: Griffin.

Kennedy, P. (2002) Sinning in the Basement: What are the Rules? The Ten Commandments of Applied Econometrics, *Journal of Economic Surveys*, 16, 569-589.

Kim, CS and Phillips, PCB. (2000) Modified log periodgram regression, Working Paper, Yale University.

Kling, G. (2006). The long-term impact of mergers and the emergence of a merger wave in pre-World-War I Germany *Explorations in Economic History*, 43, 667-688.

Klovland, J.T. (2009). New evidence on the fluctuations in ocean freight rates in the 1850s *Explorations in Economic History*, 46, 266-284.

Klug, A. and Landon-Lane, J. White, E. (2005) How could everyone have been so wrong? Forecasting the Great Depression with the railroads *Explorations in Economic History*, 42, 27-55.

Komlos, J. (1990) Nutrition population growth and the Industrial Revolution in England, *Social Sciences History*, 14, 69- 91

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54, 159-178

Landon-Lane, J. and Rockoff, H. (2007) The origin and diffusion of shocks to regional interest rates in the United States, 1880–2002 *Explorations in Economic History*, 44, 487-500.

Lauritzen, S.L. (2000). Causal inference from graphical models. In D.R. Cox and C. Klauppleberg (eds), *Complex Stochastic Systems*, Chapman & Hall, London.

Lauritzen, S.L. and T.S. Richardson, (2002). Chain graph models and their causal interpretations, *Journal of the Royal Statistical Society Series B* 64, 321-361

Lazaretou, S. (2005). The drachma, foreign creditors, and the international monetary system: tales of a currency during the 19th and the early 20th centuries *Explorations in Economic History*, 42, 202-236.

Lee, J., and Strazicich, M., (2001). Break Point Estimation and Spurious Rejections with Endogenous Unit Root Tests, *Oxford Bulletin of Economics and Statistics*, 63, 535-58

Lee, J., and Strazicich M.C., (2003). Minimum LM Unit Root Tests With Two structural Breaks, *The Review of Economics and Statistics*, 85, 1082-1089

Leung, S-K., (1992), Changes in the behaviour of output in the United Kingdom, 1856-1990, *Economics Letters,* 40, 435-44

Levine, A., Lin, C.F., Chu, C.S. (2002. Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics* 108, 1-24.

Lew, B. (2000) The Diffusion of Tractors on the Canadian Prairies: The Threshold Model and the Problem of Uncertainty. *Explorations in Economic History,* 37, 189-216.

Leybourne, S., Kim, T-H., and Taylor, A.M.R. (2007). Detecting Multiple Changes in Persistence, *Studies in Nonlinear Dynamics & Econometrics*, 11

Ljunberg, J., and Nilsson, A. (2009). Human capital and economic growth: Sweden 1870–2000 *Cliometrica,* 3, 71-95.

Lumsdaine, R. L. and D. Papell (1997), Multiple Trend Breaks and the Unit Root Hypothesis,*Review of economics and Statistics*, 79, 212-218.

Lutkepohl, H. and Leimers, H.-E. (1992) Granger-causality in cointegrated VAR processes: the case of the term structure, *Economics Letters,* 40, 263- 68.

MacKinnon, J. (1991), Critical Values for Cointegration Tests, in R.F. Engle and C.W.J. Granger, *Long-Run Economic Relationships*, Clarendon Press, Oxford.

Maddala, G. S., Wu, S.W., (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics,* 61: 631-652. *Manchester School,* 65, 192-212

Mattesini, F., and Quintieri, B. (2006). Does a reduction in the length of the working week reduce unemployment? Some evidence from the Italian economy during the Great Depression. *Explorations in Economic History,* 43, 413-437.

McAleer, M., (1994). Sherlock Holmes and the Search for Truth: A Diagnostic Tale, *Journal of Economic Surveys*, 8, 317-70

McAleer, M., (2005). Automated Inference and Learning in Modeling Financial Volatility, *Econometric Theory*, 21, 232-261.

McCloskey, D. N. (1994) 1780 - 1860, in *The Economic History of Britain Since 1700* (Eds) D. N. McCloskey and R. C. Floud, 2nd. edn, vol. 1, 242- 70, Cambridge University Press, Cambridge.

Michelacci, C, and Zalaroni, P. (2000) (Fractional) beta, convergence, *Journal of Monetary Economics* 45, 129- 153.

Mills, T.C. (1991), Are fluctuations in UK. output transitory or permanent? *The Manchester School,* 59, 1-11.

Mills, T.C. (2008). Exploring historical economic relationships: two and a half centuries of British interest rates and inflation. *Cliometrica,* 2, 213-238.

Mills, T.C. (2009) Modelling trends and cycles in economic time series: historical perspective and future developments *Cliometrica,* 3, 221-244.

Mishra, T., Prskawetz, A., Parhi, M., and Diebolt, C. (2009). A note on long memory in population and economic growth, *AFC Working Paper, No. 6*.

Moser, P., Nicholas, T., (2004). Was electricity a general purpose technology? Evidence from historical patents citations. American Economic Review 94, 388–394.

Muscatelli, V. A., & Hurn, S, (1992). Cointegration and Dynamic Time Series Models, *Journal of Economic Surveys*, 6(1), 1-43.

Neilson , H.B. (2007). UK money demand 1873–2001: a long-run time series analysis and event study. *Cliometrica* , 1, 45-61.

Nelson, C.R. and Plosser, C.I., (1982), Trends and Random Walks in Economic Time Series, *Journal of Monetary Economics*, Vol. 10, No. 1, pp.139-62.
Newey, W., West, K., (1987). A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 69 (6), 1519-1554.
Ng, S. And Perron, P. (2001). Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power, *Econometrica*, 69, 1519-1554
O'Brien, P. K. and Engerman, S. (1991) Exports and thegrowth and the British Economy from the glorious revolution to the peace of Amiens, in *Slavery and the Rise of the Atlantic System* (Ed.) B. Solow, Cambridge University Press, Cambridge, 177-209.
Officer, L.H. (2005). The quantity theory in New England, 1703–1749: new data to analyze an old question *Explorations in Economic History*, 42, 101-121.
Ogren, A. (2006). Free or central banking? Liquidity and financial deepening in Sweden, 1834–1913, *Explorations in Economic History*, 43, 64-93.
Outhwaite, R. B (1969). *Inflation in Tudor and Stuart England*, Macmillan, London.
Oxley, L. And Greasley, D. (1997b). Using the Kalman Filter in tests of convergence, paper presented at the Modelling and Simulation Society of Australia, MODSIM, conference, 1997.
Oxley, L., McAleer, M., (1993). Econometric Issues in Macroeconomic Models with Generated Regressors, *Journal of Economic Surveys*, 7(1), 1-40
Oxley, L., Reale, M., and Tunnicliffe-Wilson, G. (2009). Constructing structural VAR models with conditional independence graphs, *Mathematics and Computers in Simulation* 79, 2910-2916.
Pagan, A. (1984). Econometric Issues in the Analysis of Regressions with Generated Regressors, *International Economic Review*, 25, 221-247.
Park, J.Y. (1992). Canonical Cointegrating Regression, *Econometrica*, 60, 119-143.
Pearl, J. (2000). Causality, Cambridge University Press, Cambridge England
Pedroni, P. (1999) Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics* 61, 653– 670.
Pedroni, P. (2004) Panel cointegration: asymptotic and finite sample properties of fooled time series tests with an application to the PPP hypothesis. *Econometric Theory* 20, 597-625.
Peng, C. K., Buldyrev, S. V., Simons, M., Stanley, H. E., and Goldberger, A. L. (1994). Mosaic organization of DNA nucleotides. *Physical Review E*, 49, 1685-1689
Perron, P. (1989), The Great Crash, the Oil Price Shock and the Unit Root Hypothesis, *Econometrica*, Vol. 57, No.6, pp.1361-1401.
Persson, K. G. (2010), *An Economic History of Europe*, Cambridge University Press, Cambridge.
Pesaran, H., Shin, Y., and Smith, R. (2001) Bounds Testing Approaches to the Analysis of Level Relationships, *Journal of Applied Econometrics*. 16, 289- 26
Phillips, P.C.B and Perron, P. (1988) Testing for a unit root in time series regression, *Biometrika*, 75(2):335-346
Phillips, P.C.B. (1986). Understanding spurious regression in econometrics, *Journal of Econometrics*, 33, 311-340.
Phillips, P.C.B. (1998). New tools for understanding spurious regressions. *Econometrica*, 66, 1299-1325.

Phillips, P.C.B. (2005). Challenges of trending time series econometrics. *Mathematics and Computers in Simulation,* 68**,** 401-416.

Phillips, P.C.B., and Ouliaris, S. (1990). Asymptotic Properties of Residual Based Tests for Cointegration, *Econometrica*, 58, 165-93

Phillips, P.C.B., and Yu, J. (2009). Dating the Timeline of Financial Bubbles During the Subprime Crisis, typescript.

Prados de la Escosura, L. and Sanz-Villarroya, I. (2009) . Contract enforcement, capital accumulation, and Argentina's long-run decline. *Cliometrica,* 3, 1-26.

Priestley, M. (1981). *Spectral analysis and time series,* New York: Academic Press.

Rapach, D., and Wohar, M., (2004). Testing the monetary model of exchange rate d etermination: a closer look at panels, *Journal of International Money and Finance,* 23(6), 867-895

RATS (Regression Analysis of Time Series) 7.3. (2010). http://www.estima.com/

Rea, W., Oxley, L. Reale, M. and Mendes, E. (2008a). Long memory or shifting means? A new approach and application to realized volatility, Dept Economics WP No. 04/2008. http://www.econ.canterbury.ac.nz/RePEc/cbt/econwp/0804.pdf

Rea, W., Oxley, L. Reale, M. and Mendes, E. (2008b). The Empirical Properties of Some Popular Estimators of Long Memory Processes. http://www.econ.canterbury.ac.nz/RePEc/cbt/econwp/0815.pdf.

Rebelo, S., (1991), Long-Run Policy Analysis and Long-Run Growth, *Journal of Political Economy,* 99, pp. 500-21.

Ricciuti , R. (2008). The quest for a fiscal rule: Italy, 1861–1998. *Cliometrica,* 2, 259-274.

Robinson, P. (1994) Semi-parametric analysis of long memory time series, *Annals of Statistics*, 22, 515-539.

Robinson, P. (1995). Log-periodogram of time series with long range dependence. Annals of Statistics, 23, 1048-1072.

Robinson, P. (2003). *Time Series with Long Memory*, Oxford University Press.

Rostow, W.W (1963). *The Economics of Take-Off into Sustained Growth*. Macmillan, London.

Rousseau, P. and Sylla, R. (2005) Emerging financial markets and early US growth. *Explorations in Economic History,* 42, 1-26.

Rousseau, P.L. (2009). Share liquidity, participation, and growth of the Boston market for industrial equities, 1854–1897. *Explorations in Economic History*, 46, 203-219.

Sabate, M., Gadea, M., and Escario, R. (2006) Does fiscal policy influence monetary policy? The case of Spain, 1874–1935*Explorations in Economic History,* 43, 309-331.

Sanz-Villarroya, I. (2005). The convergence process of Argentina with Australia and Canada: 1875–2000, *Explorations in Economic History,* 42, 439-458.

Sarferaz, S., and Uebele, M. (2009). Tracking down the business cycle: A dynamic factor model for Germany 1820–1913 *Explorations in Economic History,* 46, 368-387.

Schmidt, P., Phillips, P.C.B, (1992). LM Tests for a Unit Root in the Presence of Deterministic Trends, *Oxford Bulletin of Economics and Statistics*, 54, 257-87

Schwert, G (1989). Tests for Unit Roots: A Monte Carlo Investigation, *Journal of Business & Economic Statistics,* 7, 147-59.

Shafer, G. (1996). *The Art of Causal Conjecture*, MIT Press, Cambridge MA.
Siklos, P. (2008). The Fed's reaction to the stock market during the great depression: Fact or artefact? *Explorations in Economic History*, 45, 164-184.
Simon, J. (1994) Demographic causes and consequences of the Industrial Revolution, *Journal of European Economic History*, 23, 141- 59.
Sims, C.A. (1972). Money, Income and Causality, *American Economic Review*, **62**, 540-52.
Sims, C.A. (1980). Macroeconomics and Reality, *Econometrica*, 48, 1-48.
Sims, C.A. Stock, J., and Watson, M. (1990). Inference in Linear Time Series Models with Some Unit Roots. *Econometrica* 58, 161–82.
Solomou, S., and Shimazaki, M. (2007). Japanese episodic long swings in economic growth *Explorations in Economic History*, 44, 224-241.
St. Aubyn, M (1999), Convergence Across Industrialized Countries (1890-1989). New Results using Time Series Methods, *Empirical Economics*, *24*, 23-44.
STAMP (**S**tructural **T**ime Series **A**nalyser, **M**odeller and **P**redictor) 8..3 for OxMetrics. (2010). http://stamp-software.com/
STATA 9, 10, 11. STATA Corporation. http://www.stata.com/
Streb, J., Wallusch, J., and Yin, S. (2007) Knowledge spill-over from new to old industries: The case of German synthetic dyes and textiles (1878–1913) *Explorations in Economic History*, 44, 203-223.
Sullivan, R.J., (1989), Englands "Age of Invention": The Acceleration of Patents and Patentable Invention During the Industrial Revolution, *Explorations in Economic History*, 26, 424-52.
Tattara, G. (2003) Paper money but a gold debt: Italy on the gold standard. *Explorations in Economic History*, 40, 122-142.
Temin, P. and Voth, H-J., (2005) Credit rationing and crowding out during the Industrial Revolution: Evidence from Hoare's Bank, 1702–1862, *Explorations in Economic History*, 42, 325-348.
Temin, P. (2002). Price Behavior in Ancient Babylon *Explorations in Economic History*, 39, 46-60.
Terasvirta T. (1998). Modelling economic relationships with smooth transition regressions. In *Handbook of Applied Statistics*, ed. A. Ullah, D. Giles. New York: Dekker.
Toda, H.Y. and Phillips, P.C.B., (1991) Vector Autoregression and Causality: A Theoretical Overview and Simulation Study, *Working Paper No. 977*, Cowles Foundation, New Haven, CT.
Toda, H. Y. and Yamamoto, T. (1995) Statistical Inference in Vector Autoregressions with Possibly Integrated Processes, *Journal of Econometrics*, **66**, 225-50.
Tong, H. (1990). *Non-linear Time Series: A Dynamical System Approach.* Oxford: Oxford university Press.
Toniolo, G., Conte, L., and Vecchi, G. (2003) Monetary Union, institutions and financial market integration: Italy, 1862–1905. *Explorations in Economic History*, 40, 443-461.
Tsoulouhas, T. C. (1992) A New Look at demographic and technological change England 1550- 1839, *Explorations in Economic History*, 29, 169 – 203.

Vahid, F. and Engle, R. F. (1993), Common Trends and Common Cycles, *Journal of Applied Econometrics*, **8**, 341-60.

Van Nieuwerburgh, S., Buelens, F., and Cuyvers, L. (2006). Stock market development and economic growth in Belgium *Explorations in Economic History*, 43, 13-38.

Vogelsang, T.J. (1997). Wald-type tests for detecting breaks in the trend function of a dynamic time series, *Econometric Theory*, 13, 818-849.

Waldenstrom, D. (2002) Taxing Emerging Stock Markets: A Beneficial Policy? Evidence from the Stockholm Stock Exchange, 1907–1939. *Explorations in Economic History*, 39, 29-45.

Weidenmier, M.D. (2000) The Market for Confederate Cotton Bonds *Explorations in Economic History*, 37, 76-97.

Weisdorf, J., and Sharp , P. (2009). From preventive to permissive checks: the changing nature of the Malthusian relationship between nuptiality and the price of provisions in the nineteenth century. *Cliometrica* , 3, 55-70.

Whittaker, J.C. (1990). *Graphical models in applied multivariate statistics*. Wiley: Chichester.

Wrigley, E. A. (1988), *Continuity, Chance and Change: the Character of the Industrial Revolution in England*, Cambridge, Cambridge Univ. Press.

Yousef, T.K. (2000). The Political Economy of Interwar Egyptian Cotton Policy. *Explorations in Economic History*, 37, 301-325.

Yule, G. (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, 1-64.

Zivot, E. and Andrews, D.W.K. (1992), Further Evidence on the Great Crash, the Oil Price Shock, and the Unit Root Unit Hypothesis, *Journal of Business and Economic Statistics*, 10, 251-70.