The methodological problem of unit roots: stationarity and its consequences in the context of the Tinbergen debate

Muhammad Ali Nasir[1] · Jamie Morgan[2]

Abstract

In this paper we highlight an important yet often neglected issue that arises within the context of the broader set of concerns set out in Keynes's seminal critique of Tinbergen's early work in econometrics and that is the problem of "trend" in the dataset. We use the example of conforming data to achieve stationarity to solve a problem of unit roots to highlight that Keynes concerns with the "logical issues" regarding the "conditions which the economic material must satisfy" still gains little attention in theory and practice. There is a lot more discussion of the technical aspects of method than there is reflection on conditions that must be satisfied when methods are applied. Concomitantly, there is a tendency to respond to problems of method by applying fixes rather than addressing the underlying problem. We illustrate various facets of the argument using central bank policy targeting and using examples of differencing, co-integration and Bayesian applications.

Keywords Tinbergen debate · Research methods · Causal inference · Policy targeting

JEL classification B16 · B23 · B41 · C11 · C18

1 Introduction

It is widely recognized but often tacitly neglected that all statistical approaches have intrinsic limitations that affect the degree to which they are applicable to particular contexts (in its most controversial form see Leamer 1983; 2010). John Maynard Keynes was perhaps the first to provide a concise and comprehensive summation of the key issues in his critique of Jan Tinbergen's book *Statistical Testing of Business Cycle Theories Part 1: A Method and its application to investment activity* (1939).[1] The book has a short first chapter highlighting

---

[1] Note, statistical methods adopted by Tinbergen were in use by other sciences long before his 1939 work. For example, Sir Francis Galton's seminal paper, "Regression towards Mediocrity in Hereditary Stature" (Galton, 1886). Moreover, even critics of Keynes's *General Theory* accepts that his work arose in times favourable to the development of economic analysis (see e.g. Luzzetti & Ohanian, 2011).

✉ Muhammad Ali Nasir
   m.a.nasir@leeds.ac.uk

[1] University of Leeds, Leeds, West Yorkshire, UK

[2] Leeds Beckett University, Leeds, West Yorkshire, UK

🖄 Springer

"logical issues", a longer second chapter explaining the method of multiple correlation and then three chapters of selected examples. According to Keynes it is "grievously disappointing" that so little attention is paid to the logical issues in the form of:

> explaining fully and carefully the conditions which the economic material must satisfy if the application of the method is to be fruitful… [According to Keynes, Professor Tinbergen] is much more interested in getting on with the job than in spending time in deciding whether the job is worth getting on with. He so clearly prefers the mazes of arithmetic to the mazes of logic, that I must ask him to forgive the criticisms of one whose tastes in statistical theory have been, beginning many years ago, the other way round. (Keynes, 1939, p. 559)

Tinbergen's original work was undertaken for the League of Nations, a relatively short-lived organization, and the methods Tinbergen did much to promote have outlived the League. The Cowles Commission (and later the Cowles Foundation), for example, did a great deal to promote the use of mathematics and statistics in economics first at University of Chicago and then at Yale (Christ, 1994).

Keynes's intervention has, of course, become the basis of the "Tinbergen debate" and is a touchstone whenever historically or philosophically informed methodological discussion of econometrics is undertaken.[2] It has remained the case, however, that Keynes's concerns with the "logical issues" regarding the "conditions which the economic material must satisfy" still gain little attention in theory and practice.[3] There is a lot more discussion of the technical aspects of method than there is reflection on conditions that must be satisfied when methods are applied. As David Hendry, a founding figure in the London School of Economics (LSE) approach to econometrics, notes, though few technically-minded econometricians are aware of the fact, Keynes provides proto-forms of most of the main concerns with linear regression models and method (Hendry, 1980; see also Garrone & Marchionatti, 2004). Today one might bullet these as:[4]

- The problem of "completeness" (missing variable bias);
- Lag treatment and unit of measurement problems (mis-specification),
- Issues focused on the distinction between cause and effect and the status of variables (simultaneity, endogeneity and exogeneity), as well as the problem of association between explanatory variables (multicollinearity) and "spurious correlation";
- The problem of measurability and issues focused on whether the parameters obtained are actually the true parameters (identification);
- Issues regarding changes in association over time and the "Acceleration principle" (Time-variation of coefficients);
- Assumptions of linearity (the possibility of nonlinearity).

On the problem of completeness, for example, Keynes asks, "is it assumed that the factors investigated are comprehensive and that they are not merely a partial selection out of all the factors at work?" (Keynes, 1973, p. 286-7). In terms of measurability he notes numerous

---

[2] For various discussions see Hendry and Morgan (1995), Leeson (1998) and Louçã (1999).

[3] It is still a matter of debate whether Keynes was entirely opposed to use of econometric methods. See, for example, comments such as "Newton, Boyle and Locke all played with alchemy. So let him continue" (Keynes, 1973, p. 320). He notes that had Tinbergen have been a private student he would have deserved every encouragement, however, it is very dangerous for a collection of reasonable economists to give his ideas any sort of imprimatur in its present stage….

[4] Note, we could have selected any of the following points as a focus for possible fixes (see later). For example, endogeneity. See, for example, Ullah et al. (2018); Ullah et al. (2021).

features relevant to what happens to an economy (and with reference to Tinbergen's work, within a business cycle) are unmeasurable–matters of expectations, confidence, innovations and progress, political determinations, labour disputes and so on–and as anyone familiar with Keynes understands, his list resonates with other work he undertook discussing change and degrees of uncertainty and what one can know. As Garrone and Marchinoatti (2004) point out, Tinbergen, however, does little to address such concerns in his reply and Keynes foreshadowed exactly this with his expectation that Tinbergen would likely react by seeking to "drown his sorrows in arithmetic". This is in fact what occurred and anyone familiar with how the vast majority of papers in econometrics are structured will recognise this is (no irony intended) a pattern repeated ever since.

It seems reasonable to suggest, furthermore, that Tinbergen's response either evades or misunderstands the underlying implications of Keynes's critique (see Lawson 1989). If Keynes's technical concerns follow from, rather than can be isolated from, his concern with conditions (i.e. the "logical issues"), then it seems reasonable to infer he has a conception of conditions that differs from those the methods typically require and which Tinbergen seems to rely on.[5] With this in mind, some critics, including well-known Keynesians and Post Keynesians argue that econometric use is limited insofar as its methods imply a "closed system", while most situations are "open system" and there has been some debate regarding what these terms means and whether the variability or irregularity of real world situations strictly limits the application of some methods (for a range of positions see Chick & Dow 2005; Lawson, 2015, 1997; Pratten, 2005; Fleetwood, 2017).[6]

The "Tinbergen debate", of course, was not solely focused on conditions nor was it always strictly focused on content rather than competency. For example, Samuelson (1946) questioned Keynes's technical skills and insight, Klein (1951) refers to Keynes's intervention as his "sorriest performance" and Stone (1978) combines the two–questioning both Keynes's understanding of econometrics and his temperament. As Garrone & Marchionatti (2004) make clear, though, many others at least partly concede that Keynes's concerns were well-founded and require at the very least careful response (for example, Patinkin 1976; Hendry, 1980; Pesaran & Smith, 1985; Rowley, 1988; McAleer, 1994; Dharmpala and McAleer, 1996; Keuzenkamp 2000).

The purpose of this paper, however, is not to reprise the "Tinbergen debate" but rather to highlight an important yet often neglected issue that arises within the context of the broader

---

[5] At times the two seem to assume different worlds. Consider, on the issue of completeness, Tinbergen responds to Keynes that many of "these external events will be, as a rule, of an unsystematic character, and may thus be part of unexplained residuals" (Tinbergen, 1940, p 141). This implies a framework of analysis in which everything which cannot be explained can be attributed to the residuals, but this in turn requires these to conform to definite conditions (homoscedasticity, no serial correlation, normality etc.) before one can draw any inferences. This presupposes a closed system. Similarly, on non-linearity, Tinbergen suggests in his reply to Keynes that "If we imagine the individual demand curve to be highly curvilinear, and if we imagine a great number of such demand curves to be added up, then the very fact that the place of maximum curvature will be different for most of these individual curves will already lead to a joint curve which is much more linear than any of the individual curves (p, 149)". This is a naïve example of aggregation, you cannot simply add up all the functions or estimates based on separate factors to obtain one consolidated composite function without assuming a world of strict homogeneities. Finally, in replying to Keynes's critique of quantifying a relationship which has already been explained by the economic theory, Tinbergen justifies his position by stating "For statistical testing there may be a difficulty of accuracy. Provided, however, the economist can guarantee what variables enter into each of these relations and reliable statistics exist for all these variables, the statistician is able to estimate the degree of uncertainty in the results in the ordinary way". If economists can "guarantee" these aspects, what is left to estimate and how does this reconcile doubt regarding accuracy?

[6] For further context see the discussion Lawson & Morgan (2021a, 2021b) and Olsen & Morgan (2005).

set of concerns we have set out and that is the problem of "trend" in the dataset.[7] In his original intervention Keynes writes:

> The treatment of time-lags and trends deserves much fuller discussion if the reader is to understand clearly what it involves… [And] The introduction of a trend factor is even more tricky and even less discussed. (Keynes, 1939, p. 565).

We would argue that Keynes is correct to suggest the problem of "trend" warrants "fuller discussion". According to Keynes, Tinbergen's approach to method did not pay much attention to the process of evolution of the trend, and Tinbergen's specific data choices were ad hoc and arbitrary, adopting a nine-year moving average for the pre-war period and rectilinear trend for the post-war period (Keynes, 1973). According to Keynes, this invites suspicion that the difference was a purposive convenience, a construction of statistical fact to service explanation (Keynes, 1939). Tinbergen's reply in the *Economic Journal* focuses on the attractions of a linear model–"for short periods there is not much difference between a straight line and moving average" (Tinbergen, 1940, p. 152)–and this is for all intents and purposes merely deflection, since it does not respond to Keynes's concerns regarding both the periodisation of data and the significance of how a trend might arise and evolve and how method deals with this. There is, however, more at stake here than a specific focus on problems arising in Tinbergen's early work might imply.

It was not until the 1970s that greater attention was paid to the "tricky" issue of trend in the dataset. In the *Journal of Econometrics* Granger & Newbold (1974) highlight, without reference to Keynes, that the presence of a trend in data can potentially lead to a unit root problem.[8] The presence of a unit root is often identified in macroeconomic data but the problem of unit roots for macroeconomic policy is rarely explored, despite Keynes's early reference to a problem and Granger and Newbold's separate recognition, and this is a notable omission since it has the potential to be a fundamental issue for policymaking (Nasir & Morgan, 2018).[9]

In the rest of this paper we explore some of the limits of unit root tests and stationarity using modern central banking as an illustration. However, the limits identified apply generally to any policy application where causal relationships are exploited to attempt to achieve some objective through "targeting".[10] As will become clear, targeting necessarily results in non-stationary datasets and, if successful, in unit roots. Problems of trends then follow and this creates a challenge for statistical analysis and thus, insofar as such analysis is used to inform policy decisions, policy analysis. At the very least, the existence of such problems requires

---

[7] Note, Keynes was not the first to note a problem of trend in econometric type methods. Yule (1926) who looked at the nonsense correlations that come with I(1) variables, though he did not call them I(1), and linked them back to what he had called spurious correlations in 1897. It is an accident that what Yule called nonsense correlations are now called spurious correlations. Keynes knew about Yule's work but it is more likely that Keynes concern was with the economic issue: what economic forces explained the arbitrary trends in Tinbergen's equations? This remains a problem with DSGE models which explain deviations from unexplained steady states.

[8] Note, one might argue that in particular cases a series could be considered as trend-stationary. That would, however, still invoke problems with the ability to draw inferences in the presence of a trend and problems identifying the factors leading to the evolvement of a trend in the first place.

[9] Nasir & Morgan (2018) explored the affinities between ergodicity and stationarity and discussed arising practical contradictions for central bank policymaking. The current paper elaborates, as stated, in the context of an underappreciated aspect of Keynes critique of Tinbergen i.e. the treatment of trends. The material also has implications for the relevance of open and closed system dependent approaches. See also Nasir & Morgan (2020). Section 2 draws on Nasir & Morgan (2018, 2020).

[10] For the importance of causality in particular to macroeconomics, see Hoover (2001).

careful consideration, as Keynes would have it, of the adequacy of conditions, and in more contemporary parlance familiar to econometricians, the link between what we infer about "data generating processes" and the application and management of models and statistical analysis.

2 Back to basics: unit roots and stationarity

Stationarity is a common requirement of statistical approaches to longitudinal data, specifically time series and panel data. A stationary data series does not possess a unit root. That is, it does not possess a trend value dominated by its own past values. If it does possess a unit root then the series will not revert to a mean and the series may be proportionally influenced by its internal innovations. This makes standard causal inference problematic because one cannot readily differentiate the effect of any other explanatory factor from the innovation in the series.[11] According to Granger & Newbold ([1974](#)), there is a spurious result and model problem. The same issue applies if the series "explodes". So, when stated in typical model form there are three possible cases.

Consider an AR (1) model:

$$Y_t = \varnothing Y_{t-1} + e_t \tag{1}$$

In Eq. ([1](#)) e_t is a random noise process. The coefficient $\varnothing$ can have one of 3 possible consequences for the series:

a. a. $|\varnothing|=1$ where the series contains a unit root and is non-stationary;
b. b. $|\varnothing|>1$ where the series explodes;
c. c. $|\varnothing|<1$ implying the series is stationary.

In order to "manage" the problem of causal inference it is standard practice to transform a non-stationary series into a stationary form. For example, one might seek to solve the problem by differencing.

If one subtracts Y_{t-1} from both sides of Eq. ([1](#)):

$$Y_t - Y_{t-1} = Y_{t-1} - Y_{t-1} + e_t \tag{2}$$

$$\Delta Y_t = e_t$$

As Eq. ([2](#)) highlights, taking the difference leaves the change in ΔY_t as white noise or e_t and so by integrating ΔY_t becomes stationary. More specifically, the series Y_t is integrated of order one and whilst $Y_t \tilde{I}(1)$ contains a unit root ΔY_t is stationary.[12]

What we want to emphasise is that though differencing and other transformations seem to be solving a problem often there are new problems created. This can be illustrated using central bank inflation targeting.

---

[11] This point requires acknowledgement of Hendry's ([2004](#)) point that despite the fact that causality and exogeneity are not required to be invariant characteristics of the data generating process (DGP), they remain relevant in non-stationary processes.

[12] Similarly, a series will be integrated of order d (denoted by $Y_t \tilde{I}(d)$) if Y_t is non-stationary, but $\Delta^d Y_t$ is stationary (See Asteriou & Hall [2016](#)). This is often described as integrated of a certain order of differencing, depending on how many times the difference (d) has been taken, i.e. I(d).

3 Central bank targeting: the tension between tests conducted for empirical adequacy and actual policy purpose

Modern central banking is focused on price stability and many central banks have formal or informal inflation targeting systems.[13] The Bank of England, for example, has an inflation target of 2% with a 1% range. If the allowable leeway is exceeded the Governor of the Bank must write a letter to the Chancellor of the Exchequer at the Treasury explaining why this is the case and, usually, what the Bank is intending to do to recover the situation. The 1998 Bank of England Act allows the Chancellor to develop and modify the remit of the Bank and its target. Since independence in 1997 HM Treasury has required the Bank to send a further letter 3 months after the first if the situation persists. Since March 2020 the Bank has been required to publish its open letter alongside the Bank's Monetary Policy Committee minutes.[14] Furthermore, after the global financial crisis and since the creation of a Financial Policy Committee with formal macroprudential powers, it has been accepted that there can be reasons to allow deviations to persist for some period in order to expedite other policy goals, such as financial stability and the Monetary Policy Committee and the Financial Policy Committee are expected to coordinate (Morgan, 2022). This is phrased as discretion based on discussion of "trade-offs" and a "time horizon".[15]

In any case, price stability remains at the heart of Bank independence and the "current cost of living crisis" in the UK makes this only too clear. It also makes it clear that independence is no guarantee a Bank can control its inflationary environment sufficiently to render high and persistent levels of inflation unlikely in all circumstances. Moreover, price stability at the Bank, as it is at similar central banks worldwide, is situated to a contextualising goal of furthering economic growth and the Bank seeks to model the relation between inflation and growth in a variety of ways (structural equation DSGEs etc.).

Clearly, both targeted inflation and desired growth take *positive* values. If we return to the material set out in Sect. 2 and keep this in mind then two potential consequences are immediately apparent.

First, if the inflation target *is* achieved then the data series will *not* conform to the statistical properties of stationarity. Consider this based on Eq. (1): $Y_t = \varnothing Y_{t-1} + e_t$ If we specify the model as inflation and an inflation target of 2% is consistently achieved then Y_t will become equal to Y_{t-1} and theta will become 1 and e_t zero. If theta becomes 1 there is a unit root. Moreover, if one takes the difference to remove the unit root then the net of the difference would be zero (a nonsensical result for the series).

Second, if the target *is not* achieved then the Bank faces a problem of continually aligning future outcomes based on previous actual positive values for inflation in a lead-lag process. This requires policy to focus above or below the actual target in order to achieve the target. However, the point of reference still remains the achievement of a long-term positive value that is unit root *creating*. Anyone familiar with the continual dance of monetary policy will know that the target is rarely achieved and if achieved is rarely achieved for long–success in reality amounts to relatively low and stable inflation as a background factor conducive to, rather than undermining of, economic activity (now and in the immediate future over which plans are made) through expectations formation and through interest rate policy. As such,

---

[13] The concept of a target was initially proposed by Meade (1978) when defining "The Meaning of "Internal Balance". For interesting insights into Nominal Income Targeting, see Tobin (1980) and Bean (1983).

[14] See: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/871684/MPC_RemitFINAL.pdf.

[15] Visit: https://www.gov.uk/government/collections/inflationary-targets.

though, the reality and very purpose of policy means that the initial dataset is highly unlikely to be stationary over time.

Now, setting out the above may seem trivial. It is important to understand, however, that stationarity is typically taken to be a condition of empirical adequacy, and as a consequence stationarity is usually tested for. Stationarity and thus either removing or managing unit root situations has become basic to empirical work. Testing and conforming data are standard practice for time series analysis and are part of how authority is claimed, but there is clearly a potential problem of data interpretation in relation to policy practice if stationarity is our point of reference.

The targetter, in this case the Bank of England, is, of course, not concerned with achieving stationarity. The targetter is concerned with achieving the target rate, but policy success leads to data series with non-mean reverting characteristics and hence a unit root. Put another way, the test is seeking something at odds with the consequences of policy and thus at odds with the purpose of policy. This is then exacerbated when data is transformed to conform to stationarity, since the data can take on new characteristics, which can in turn create problems of interpretation for policy (as we will show in the next section).

The important point, therefore, is that policy practice (economic objectives) seems to indicate a focus on stationarity is an undesirable starting point for empirical investigation. There is a limitation if application focuses on unit root tests and a limitation if the response is to conform the data series to stationarity. The combination of points implies that there is a problem if economic properties of the data are put aside in order to ensure that statistical properties hold. Our standard practice has for years encouraged us to opt for ensuring statistical properties hold, through various fixes. This is a complication for policymakers. It might be argued that the further a statistical fix takes us away from the conditions in which the data arises then the more problematic the fix becomes. This returns us to Keynes's original intervention on trends.

4 Differencing and problems of data transformation

In 1939 Keynes questioned Tinbergen's treatment of trends (1939). Tinbergen took the moving average and Keynes commented that:

> This seems rather arbitrary. But, apart from that, should not the trends of the basic factors be allowed to be reflected in a trend of the resulting phenomenon? *Why is correction necessary*? (Keynes, 1939: p. 565; emphasis added). [And in so far as choice of period of analysis has implications for parameters and inferences Keynes goes on to suggest] this looks to be a disastrous procedure. (Keynes, 1939: p. 566).

In his *Economic Journal* response Tinbergen states:

since a change in period will in most cases change the trend in the dependent variable in about the same extent as the trend in the agglomerate of explanatory variables. Therefore, the procedure followed here is not, as Mr. Keynes thinks, disastrous (Tinbergen 19,740, p, 152).

And yet three and half decades later, as we have previously alluded, Granger and Newbold (1974) show that it could be–albeit it was not their intent to comment on Tinbergen. In order to establish this they use a Monte Carlo approach and two artificially generated series; unit roots failed to reject significant association between them, leading to a "spurious

Table 1 Consumer price index (2000–2021)

Year	CPI(Y_t)	ΔYt	Year	CPI(Y_t)	ΔY_t
2000	0.8	–	2011	4.5	1.2
2001	1.2	0.4	2012	2.8	− 1.7
2002	1.3	0.1	2013	2.6	− 0.2
2003	1.4	0.1	2014	1.5	− 1.1
2004	1.3	− 0.1	2015	0.0	− 1.5
2005	2.1	0.8	2016	0.7	0.7
2006	2.3	0.2	2017	2.7	2
2007	2.3	0.0	2018	2.5	− 0.2
2008	3.6	1.3	2019	1.8	− 0.7
2009	2.2	− 1.4	2020	0.9	− 0.9
2010	3.3	1.1	2021	2.6	1.7

Source: Office for the National Statistics (ONS; 2022)

A unit root test is required to establish that the original data was not stationary and that its first difference is stationary. One of the more common tests is the Augmented Dickey and Fuller (ADF) test. ADF has particular limitations (see Perron 1989; Ranganathan & Ananthakumar, 2010), but our purpose is to highlight a generally applicable consequence arising implicitly from the purpose for which the test is undertaken and so we need not discuss these particular limitations. Table 2 sets out the results:

regression" problem. To reiterate, to avoid this problem stationarity has become accepted as a precondition of economic and statistical analysis. As we have already suggested in Sect. 2 contemporary practice to deal with non-stationarity is to take the difference and as previously stated differencing may transform a non-stationary series with unit roots into a stationary series.

Consider, however, the focus set out in Sect. 3 on inflation. Table 1 sets out UK Consumer Price Index (CPI) data and its first difference ΔY_t 2000 to 2021:

The results establish that the series was non-stationary and then *becomes* stationary after taking the difference. To expedite explanation of the arising issues the CPI series can also be represented graphically as Fig. 1:

The point we want to make here is that the patterns in the new treatment of the data are not identical to the old. Consider what that means, data is differenced in order to achieve stationarity and policymakers are being asked to rely on this data to make decisions.

Putting aside the Log CPI for the moment, both the CPI and differenced CPI follow similar patterns. However, there are differences in the pattern for the two for various periods. For example, between 2010 and 2011 and also 2012–2013. The differenced series also takes

Table 2 ADF test: consumer price index 2000–2021

ADF test Stats	With intercept	p value	Trend & Intercept	p-value	None	p value
At level I (0)	− 2.626	0.104	− 2.545	0.306	− 0.753	0.378
At first diff I(1)	− 4.543**	0.002	− 4.374*	0.013	− 4.655**	0.000

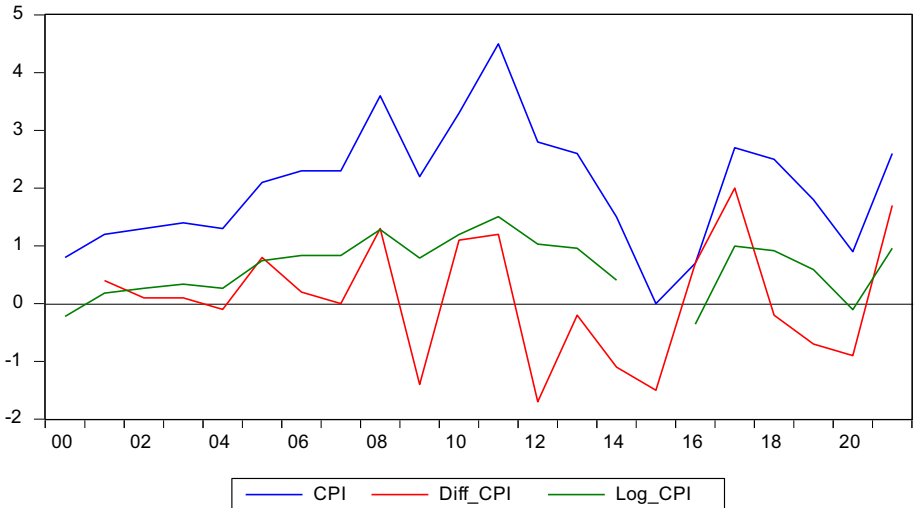*5% level of Significance, **1% level of Significance

Fig. 1 Consumer Price Index (2000–2021). Source: Office for National Statistics (2022). *Note* no value for Log-inflation (CPI) in 2015 as the annual CPI inflation was 0

negative values for some periods. Clearly, this creates problems of interpretation of data for policy purposes in so far as one takes stationarity to be a relevant point of reference. This is important because it is not just an issue of discovering a "well-behaved" dataset, such that the problem of a spurious result and model are resolved. Rather, it is a matter of constructing a dataset whose results may now be different in so far as one cannot just assume that the statistical results necessarily align with the real economic situation.

The outcome highlighted may not seem a particularly important problem because any policymaker is likely to have at their disposal both series. It is important though to think about context. Transforming the data series to achieve stationarity is used to address the problem of causal inference created by the existence of unit roots. Thus if one is approaching the problem of causation for policy purposes based on stationarity there is a potential problem of matching one data series to the other.[16] One need only refer to Fig. 1 to appreciate that similar issues arise if one opts for a logarithmic transformation.

To be clear, there is a potential mismatch because standard practice opts to begin from ensuring that statistical properties hold. Once the data series is transformed, as it is in Fig. 1, then a problem arises with common analytical statistical methods used to determine the degree of causality among variables of interest. The change in pattern for 2010–2011 and 2012–2013 *and* the negative values taken for some periods (which will express as deflation) are in this context extremely important. They will affect signs and magnitudes of relations and hence have implications for policy discussion. In Keynes's terms, key questions are, what is the importance of the trend and how does its treatment bear on the conditions that the statistical procedure is supposed to be providing insight into?

---

[16] To be clear, the problem of matching is that one is shifting between a non-stationary data series in which the actual causal processes are occurring and a transformed stationary data series used to explore causal factors. The danger is that by prioritising statistical properties one obfuscates in terms of actual economic properties.

The problem may also extend to forecasting. For example, the Bank of England uses ARIMA modelling, a condition of which is that data be stationary.[17] We noted earlier that pursuing a target might create non-stationarity. If targeting means the data series is non-stationary, but the model forecasting future inflation levels requires the inflation series to be stationary in order that ARIMA can be applied, then another area of mismatch arises.

In general, it might be argued then that relying on a statistical procedure, which removes the very feature that is central to and is realised through policy, is liable to complicate matters and may well be counterproductive. One might, of course, consider various existent "work-arounds". For example, quantitative macro models incorporate non-stationary shocks, and are referenced to non-stationary aggregate variables. However, for a work-around to be a reliable solution it surely must be able to make claims regarding its adequacy to the conditions of the economy and yet "work-arounds" are not posed as solutions to the basic problem of mismatch created by the very purpose of macroeconomic targeting.

It should also be noted that testing using or requiring stationarity is not just carried out by researchers in economics. Tests are common to the field of management, finance, and accounting, to name just a few. In any case, the underlying point is that there is a "data generating process" and seemingly more attention is paid to perpetuating statistical method rather than thinking about the "logical issues" of application. The greater part of intellectual resources are expended on managing problems of modelling rather than rethinking adequacy from the more fundamental point of view re the "conditions which the economic material must satisfy". This, moreover, is not just an issue invoked by differencing, every fix or work-around creates some set of problems, for example, in terms of causation, and so involve swapping one set of problems for another….

5 From differencing to co-integration?

Various technical problems arise if one uses differencing and these illustrate the issue of trade-offs when applying fixes. A commonly acknowledged problem is that differencing leads to loss of long-run information in the data, and this information is central to the adequacy of a theoretical model (see Geda et al., 2012)[18]. If one assumes, for example, a model is correctly specified as a relationship between two entities (say, X & Y) and both series are differenced then the error process is also being differenced, which then produces a non-invertible moving average error process. This leads to issues with estimation (Plosser and Schwert, 1977). If we difference the variables the model can no longer give a *unique* long-run solution: if we have the known value of x we cannot solve for y without knowing the past values of both x and y, hence a solution for x is not unique for any given x (Asteriou & Hall, 2016).

Recognizing these technical issues has led to attempts to combine short-run differenced equations with the ("level-based") long-run. The resulting co-integration models are then rendered via a stationary vector. However, in terms of limitations, this approach introduces new versions of existent problems. Co-integration can result in both the failure to identify any actual, and the tendency to accept spurious, causal factors (Carmen, 2001). For example, the currently preferred version of co-integration (the Johansen method) has attracted particular criticism in that it exhibits a high probability of generating outliers, as well as high variance

---

[17] For details, see http://www.bankofengland.co.uk/research/Documents/workingpapers/2013/wp471.pdf.

[18] One might note here that Keynes's critique of the treatment of trend data also implies the problem of information loss. In a different context, Martínez-Rivera & Ventosa-Santaulària (2012) also showed that the spurious regression problem cannot always be solved by using standard procedures.

and sensitivity to the lag length selected for the Vector Error Correction Model (see Brooks 2008).[19]

To be clear, in drawing attention to these problems, our purpose is not to pursue detailed explanation and discussion of the many potential technical issues. The point we wish to emphasise is that attempts to resolve problems of non-stationarity by either differencing or finding a stationary co-integrating process do not necessarily resolve the underlying problem. That is, whether conforming non-stationary data to stationarity is an appropriate starting point or foundation for the analysis of causal relationships. "Loss of information" is an anodyne statement but it places question marks against the adequacy of the treatment of data and this very obviously creates a subsequent interpretation problem for policymakers. Put another way, facts are not "speaking for themselves" and one has a variant of the problem identified in the Quine-Duhem thesis–auxiliary assumptions and background hypotheses are being invoked in order to apply fixes–and this is a question mark against any tacit empiricist position.

These problems, of course, are not restricted to central banking, and issues like information loss are generalisable.[20] In general terms, loss of information creates "de-trending", which leads to problems of information absences for a policymaker (e.g., Sargan 1964; Hendry & Mizon, 1978; Granger & Joyeux, 1980; Juselius, 1991; Fuleky and Bonham, 2015). In particular, the differencing leads to loss of information on common behaviour of two or more *separate* series which, subject to the adequacy of the methods with reference to the data generating process, could have provided insight regarding causal inferences (for detailed discussion see Hendry, 1995; Reiss, 2015).[21]

It is important too to understand that a primary issue here is whether the nature of causation sought from the data generating process is one the methods used are set up to identify. Arguably this is obscured by a further technical problem. In the case of co-integration, the absence of a co-integrated causal relation implies a lack of causal relation in the data series. It may, however, *still be the case* that there is some causal relation in the original data series and vice versa.[22] Transforming the data in order to explore this creates a problem based on the transformation (as already established in the previous section) because of the policy purpose, which, if achieved, creates a unit root. So, neither the co-integration nor the transformation to stationarity resolve the problem *of stationarity* in the original dataset, *but at the same time* a basic question of adequacy of theory of causation is invoked (a challenge that technical discussion is not really set up to meet since it is–in Keynes's terms–a "logical issue", or in more modern philosophical parlance, one wold say the problem is "ontological").

To fully appreciate what co-integration is *not* doing we provide an example below. In the example, we reuse the CPI series and combine it with the short-term nominal interest rate set by the Bank of England, termed Bank Rate. As previously we adopt a common or standard

---

[19] Moreover, the co-integration is not accurate in terms of predicting co-integrating relationships in the future, since the basis of the co-integration is static. The assumption of stability in the long-run relationships is not necessarily warranted either as "linkages may be time-varying and episodic". One proposal to account for the potential time-variance in co-integration studies has been suggested by Gregory & Hanson (1996), whose method detects structural breaks in the data, revealing support of long-run relationships not captured by static tests. According to Maddala & Kim (1998), Johansen's procedure is also highly sensitive to the assumption that errors are independent and normally distributed. If indeed errors are not normally distributed, it has been found that it is more likely that the null hypothesis of no Cointegration is rejected, despite this being a possibility (Huang & Yang, 1996).

[20] For example, transformation may remove any unit root, but may also result in important information loss regarding any bubble, which could affect policy prescriptions (see Reiss 2015).

[21] The differencing can also lead to inefficient and inconsistent inferences (Mizon, 1995).

[22] On co-integration and the "principle of common cause" contrast Sober (2001) and Hoover (2003).

Table 3 Johansen cointegration test summary

Selected (0.05 level*) Number of Cointegrating Relations by Model

Data trend:	None	None	Linear	Linear	Quadratic
Test type	No Intercept	Intercept	Intercept	Intercept	Intercept
	No Trend	No Trend	No Trend	Trend	Trend
Trace	0	0	0	0	0
Max-Eig	0	0	0	0	0

Information criteria by rank and model

Data trend:	None	None	Linear	Linear	Quadratic
Rank or No. of CEs	No Intercept	Intercept	Intercept	Intercept	Intercept
	No Trend	No Trend	No Trend	Trend	Trend
	Log Likelihood by Rank (rows) and Model (columns)				
0	− 55.346	− 55.346	− 54.985	− 54.985	− 54.914
1	− 53.915	− 51.262	− 50.901	− 49.219	− 49.153
2	− 53.176	− 49.890	− 49.890	− 47.150	− 47.150
	Akaike Information Criteria by Rank (rows) and Model (columns)				
0	5.934*	5.934*	6.098	6.098	6.291
1	6.191	6.026	6.090	6.021	6.115
2	6.517	6.389	6.389	6.315	6.315
	Schwarz Criteria by Rank (rows) and Model(columns)				
0	6.133*	6.133*	6.397	6.397	6.689
1	6.589	6.474	6.588	6.569	6.712
2	7.115	7.086	7.086	7.111	7.111

*Critical values based on MacKinnon-Haug-Michelis (1999)

approach for illustrative purposes, in this case the Johansen method of co-integration. For completeness we ran available alternative specifications. The results are presented in Table 3 below:

As Table 3 indicates no evidence of co-integration between CPI and the Bank Rate was found for the period of analysis. As such, one might then infer, if this were the method adopted to overcome problems of stationarity in separate datasets in order to explore causation, no causation existed. However, the only *warrantable* finding is that no causation has been found between the CPI and the Base Rate. The use of methods, moreover, creates new methodological problems based on the underlying assumptions of co-integration. Co-integration tests typically employ the assumption that the co-integration vector is constant during the period of analysis. The method requires that the two series remain regular and "well-behaved" during the period. It is, however, possible and arguably likely that the relationships change and do so more frequently than the periodization chosen for the analysis. This requires yet another fix. One might, for example, incorporate single break (Hansen & Gregory, 1996) or "two unknown break" solutions (Hatemi, 2008).[23] There is, here, a subtle difference between

---

[23] One might suggest this issue parallels the Frequentist vs. Bayesian parameter problem as well as the issue of discrete break vs. smooth shift in parameters debate. For the latter see Primiceri (2005) and Nasir (2017).

additional complication and model sophistication–the adjustment is a complication of methods rather than an improvement in the sophistication of the model as a direct response to the data generating process. In any case, the complication creates further problems if there are *multiple* breaks.

Finally, if we return to Table 3, a loss of information is also evident when considering alternative criteria (AIC, SIC LL). Different criteria set out in the columns show different information loss. Taking all of this into account, what co-integration is *not* doing is resolving the original and thus primary problem i.e., the presence of non-stationarity; and it is not, therefore, resolving this problem for policy insight. Co-integration merely shifts the terms and focus of argument and analysis regarding causation. It does not resolve the problem of causation inhering in the original dataset.

6 Bayesian estimation as an occasional solution?

Another possible way forward is to shift to a Bayesian framework. A key claim of Bayesian approaches is that they do not require stationarity (Sims, 1989; Fanchon & Wendel, 1992). However, use of Bayes creates problems of commensuration of, and dispute based on, methods and the assimilation of findings. For example, the Johansen identification procedure makes use of the observed data and, one might argue, cannot be applied under a pure Bayesian scheme. There is, though, some dispute here. Engle and Yoo (1987) argue that Bayesian estimation is not appropriate for co-integrated data because the standard priors imply a model that approaches the classical VAR model estimated with the differenced data. According to Engle and Yoo, this model is mis-specified because it does not include the error correction term. By contrast, Fanchon and Wendel (1992) argue that Bayesian models are useful even in the case of non-stationarity and co-integration. According to Warne (2006), since the full rank matrix used in these operations is arbitrary, the individual parameters of the co-integration relations and the weights are not uniquely determined, and one must make some further assumptions. However, the Johansen procedure ensures that all the parameters of the co-integrated VAR are identified. Significantly, the core issue of mis-specification remains a dividing line and this fundamentally affects how one views the problem of stationarity and ranges of methods.

The core concerns we have raised in this paper relate to unit roots, stationarity and the appropriateness of methods for policy insight. Here, some of the more fundamental critiques of Bayesian estimation techniques also seem relevant. A great deal of policy work depends on testing a null hypothesis at some arbitrary significance level. This typically relies on a classical confidence approach. The great advantage of Bayes is that it lacks a sharp null hypothesis. However, this is also a weakness in so far as the incorporation of priors requires one to make particular assumptions about changes to behaviour over time (for the fundamental issues see Efron, 1986). The problem becomes practical in terms of choosing appropriate or adequate priors (Efron, 2015)[24]. There are numerous corollary problems that arise. Bayesian updating, for example, involves four problems:

> Firstly, if a possibility was not included as a prior *ex ante*, it must have zero posterior probability. Secondly, notwithstanding this technical problem, even if it were possible within the Bayesian framework to add more possibilities over time as they are revealed, this would anyway at some point require the downgrading of the probability

---

[24] In the absence of relevant prior experience, the informed priors may lead to superficially appropriate looking estimates and credible confidence intervals. However, they lack the logical force associated with experience based priors.

of *ex ante* included possibilities due to the finite amount of likelihood available and the requirement that probabilities sum to unity (the so-called problem of "additivity"). This downgrading of probability may contribute to highly impactful and extreme outcomes being overlooked or dismissed by policymakers. Thirdly, reserving a portion of probability for these possibilities revealed over time, which is then "consumed" as these possibilities are revealed so that downgrading of the probability of *ex ante* included possibilities is avoided, does not help because the amount of reserve needed is unknown *ex ante*. Fourthly and finally, an eventual convergence with objective reality through successive updating based on information revealed over time is of no use in circumstances in which decisions must be taken now to head off what might be an extreme outcome in the future. If the convergence to objective probabilities eventually reveals that an extreme outcome has a high probability it is at that point already too late to prevent it. (Derbyshire & Morgan, 2022, p. 6)

While the issue here is not stationarity or unit roots, it takes no great perspicacity to understand that these four problems all speak to the "logical issues" problem i.e. conditions of application. In any case, if we think of instances where Bayes is posed as a fix or alternative then Stock notes that "the posteriors reported by the Bayesian unit rooters shed no new light on the question of whether there are unit roots in economic time-series" (Stock, 1991, p. 409).

As a final example, in assessing Patrick Suppes' probabilistic theory of causality and Bayesian networks, Julian Reiss states:

While Bayesian networks arguably constitute a development and generalisation of his probabilistic theory of causality, they ignore that the theory made causal judgements always against a framework given by a scientific theory. Purely empirical applications, such as those intended by the developers of Bayesian networks, fall outside its scope. And this is for good reason: unless a scientific theory justifies assumptions such as the CMC (Causal Markov Condition) and the FC (Faithfulness Condition) and assumptions about the probability model, there is little reason to believe that they should hold (Reiss, 2016: p. 295).

Given the four problems itemised in the previous quote there is clearly a major challenge involved in opting for Bayesian methods if one is to make inferences about, for example, causation.

There is more that might be said here but we have provided sufficient examples to illustrate that Bayesian analysis adds as much to problems as it does to solutions for policymakers (contrast Marriott & Newbold, 1998; Stock, 1991; Schotman & Van Dijk, 1991).[25]

7 Conclusion

Econometrics has changed a great deal since Keynes wrote his assessment of Tinbergen's work. His comment, however, that there needs to be careful consideration of the "conditions

---

[25] Stock makes a subsequent point that communicating findings to a policy audience becomes problematic once it becomes clear that posteriors are highly sensitive to priors. Fundamental issues also arise in terms of the truth adequacy of belief expressed in priors. If priors become learning processes in what sense are they based on adequate information/knowledge and learning strategies (see also Phillips, 1987)? There are many further methodological issues here that could arise. For example, stating the weaknesses of Bayes is not an argument for the alternative, unless the alternative lacks weaknesses, or is demonstrably superior for the purposes posed, and in terms of which debate is engaged. Classical claims may be clearer, but if incorrect they are not necessarily to be preferred. The main point here is that the issues are not settled.

which the economic material must satisfy" remains as relevant as ever. We have explored a number of issues focused particularly on a subset of the problem of trends in data. Stationarity is a common requirement of statistical approaches. Transforming data to eliminate unit roots achieves stationarity but does so in ways that can create mismatches between policy purposes and statistical procedures and their properties. This point is, of course, basic and to the statistically sophisticated, obvious. However, there is a danger that as one becomes statistically sophisticated one simply suppresses the significance of the limits stationarity implies–one fails to attend to the "logical issues". Opting for stationarity can be a significant problem if the actual point of policy creates real-world consequences that have unit roots. Shifting back and forth between data series then involves a mismatch based on applied methods to explore causal inference and most of the "fixes" are not really fixing this.

Keynes is one of those figures who is constantly being claimed and interpreted and it is perhaps worth noting that while his critique was trenchant he appointed an econometrician, Dick Stone, to be head of the Department of Applied Economics in Cambridge. Stone was later to write that Keynes suggested of the 1944 White Paper on Employment Policy that "theoretical economics has now reached a point where it is fit to be applied" and foresaw "a new era of 'Joy through Statistics'" (Stone, 1978). Still, there is considerable difference between the positions of say Aris Spanos, who suggests Keynes's critique has been addressed and answered and Lars Syll who suggests it has not and that it cannot be answered (for examples of their work, see Spanos, 2015; Syll, 2016). There is also a broader issue here, as Tony Lawson, and to somewhat different ends and following Roy Weintraub, Paul Davidson and Donald Katzner argue, methods are not chosen and developed merely because they are the best we have. Developments acquire authority and exhibit path-dependency, and this is not a neutral or merely objective process of progress (see, for example, Lawson, 2015; Davidson, 2003; Katzner, 2003). Still, one can at least agree with Keynes that insufficient attention to the "logical issues" risks development of methods without due attention to the world they are applied to. This begs ontological questions and these are not simply answered through work-arounds and fixes.

Finally, there is great scope here for future research. One might, for example, explore similar issues and implications across the array of business school departments and disciplines where analytical statistics have become common and stationarity has become a fundamental requirement of data series.

References

Asteriou, D., & Hall, S. G. (2016). *Applied econometrics*. Palgrave Macmillan.

Bean, C. R. (1983). Targeting nominal income: An appraisal. *Economic Journal, 93*(372), 806–819.

Brooks, C. (2008). *Introductory econometrics for finance*. Cambridge University Press.

Carmen, G. M. (2001). Causality and cointegration between consumption and GDP in 25 OECD countries: Limitations of the cointegration approach. *Applied Econometrics and International Development, AEEADE*(1), 39–61.

Chick, V., & Dow, S. (2005). The meaning of open systems. *Journal of Economic Methodology, 12*(3), 363–381.

Christ, C. F. (1994). The cowles commission contributions to econometrics at Chicago: 1939–1955. *Journal of Economic Literature, 32*(1), 30–59.

Davidson, P. (2003). "Is "mathematical science" an oxymoron when used to describe economics? *Journal of Post Keynesian Economics, 25*(4), 527–545.

Derbyshire, J., & Morgan, J. (2022). Is seeking certainty in climate sensitivity measures counterproductive in the context of climate emergency? The case for scenario planning. *Technological Forecasting and Social Change*, *182*(e17), 121811.

Efron, B. (1986). "Why isn't everyone a bayesian? *The American Statistician American Statistical Association, 40*(1), 1–11.

Efron, B. (2015). Frequentist accuracy of bayesian estimate. *Journal of Royal Statistical Society B, 77*(3), 617–646.

Engle, R. F., & Yoo, B. S. (1987). Forecasting and testing in co-integrated systems. *Journal of Econometrics*, *35*(1), 143–159.

Fanchon, P., & Wendel, J. (1992). Estimating VAR models under non-stationarity and co-integration: Alternative approaches for forecasting cattle prices. *Applied Economics, 24*(2), 207–217.

Fleetwood, S. (2017). The critical realist conception of open and closed systems. *Journal of Economic Methodology, 24*(1), 41–68.

Galton, F. (1886). Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland, 15*, 246–263.

Garrone, G., & Marchionatti, R. (2004). Keynes on econometric method, *A reassessment of his debate with Tinbergen and other econometricians*, 1938–1943. University of Torino, Department of Economics. Working paper No 1/2004.

Geda, A., Ndung'u, N., & Zerfu, D. (2012). *Applied time series econometrics. A practical guide for macroeconomic researchers with a focus on Africa*. Nairobi: Univ. of Nairobi Press.

Granger, C. W., & Joyeux, J. R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis, 1*, 15–29.

Granger, C. W. J., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics, 2*(2), 111–120.

Gregory, A., & Hanson, B. (1996). Residual based tests for cointegration in models with regime shifts. *The Journal of Econometrics, 70*, 99–126.

Hatemi-J, A. (2008). Tests for cointegration with two unknown regime shifts with an application to financial market integration. *Empirical Economics, 35*(3), 497–505.

Hendry, D. F. (1980). Econometrics-alchemy or science? *Economica, 47*, 387–406.

Hendry, D. F. (2004). Causality and exogeneity in non-stationary economic time series, centre for philosophy of natural and social science, causality. Metaphysics and Methods Technical Reports CTR 18/04.

Hendry, D. F., & Mizon, G. (1978). Serial correlation as a convenient simplification Nota Nuisance: A comment on a study of the demand for money by the Bank of England. *Economic Journal, 88*, 349–363.

Hendry, D. F., & Morgan, M. S. (Eds.). (1995). *The foundations of econometric analysis*. Cambridge University Press.

Hoover, K. D. (2001). *Causality in macroeconomics*. Cambridge University Press.

Hoover, K. D. (2003). Nonstationary time series, cointegration, and the principle of the common cause. *British Journal for the Philosophy of Science, 54*, 527–551.

Huang, B., & Yang, C. (1996). Long-run purchasing power parity revisited: Monte carlo simulation. *Applied Economics, 28*(8), 967–975.

Juselius, K. (1991). Long-run relations in a well-defined statistical model for the data generating process: Cointegration analysis of the PPP and UIP relations between Denmark and Germany. In J. Gruber (Ed.), *Econometric decision Models: New methods of modelling and applications*. Springer Verlag.

Katzner, D. (2003). Why mathematics in economics? *Journal of Post Keynesian Economics, 25*(4), 561–574.

Keuzenkamp, H. A. (2000). *Probability, econometrics and truth. The methodology of econometrics*. Cambridge University Press.

Keynes, J. M. (1939). Professor Tinbergen's method. *Economic Journal, 49*(195), 558–577.

Keynes, J. M. (1973). The collected writings of J.M. Keynes, vol. XIV, The General Theory and After. Part II. Defence and Development. London: Macmillan (Royal Economic Society).

Klein, L. (1951). "The life of J. M. Keynes. *Journal of Political Economy*, *59*, 443–451.

Lawson, T. (1989). Realism and instrumentalism in the development of econometrics. *Oxford Economic Papers, 41*(1), 236–258.

Lawson, T. (1997). *Economics and reality*. Routledge.

Lawson, T. (2015). *Essays on the nature and state of economics*. Routledge.

Lawson, T., & Morgan, J. (2021). Cambridge social ontology, the philosophical critique of modern economics and social positioning theory: An interview with Tony Lawson, Part 1. *Journal of Critical Realism, 20*(1), 72–97.

Lawson, T., & Morgan, J. (2021). Cambridge social ontology, the philosophical critique of modern economics and social positioning theory: An interview with Tony Lawson, Part 2. *Journal of Critical Realism, 20*(2), 201–237.

Leamer, E. (1983). Let's take the con out of econometrics. *American Economic Review, 73*(1), 31–43.

Leamer, E. (2010). Tantalus on the road to Asymptopia. *Journal of Economic perspectives*, *24*(2), 31–46.

Leeson, R. (1998). The ghosts I called I can't get rid of now: The Keynes-Tinbergen-Friedman-Phillips critique of keynesian macroeconometrics. *History of Political Economy, 30*(1), 51–94.

Louçã, F. (1999). The econometric challenge to Keynes: Arguments and contradictions in the early debate about a late issue'. *The European Journal of the History of Economic Thought, 6*, 404–438.

Luzzetti, M., & Ohanian, L. (2011). "Macroeconomic paradigm shifts and Keynes's General Theory." https://voxeu.org/article/macroeconomic-paradigm-shifts-and-keynes-s-general-theory

Maddala, G., & Kim, I. (1998). *Unit roots, cointegration and structural change*. Cambridge University Press.

Marriott, J., & Newbold, P. (1998). Bayesian comparison of ARIMA and stationary ARIMA models. *International Statistical Review, 66*(3), 323–336.

Martínez-Rivera, B., & Ventosa-Santaulària, D. (2012). A comment on 'Is the spurious regression problem spurious? *Economics Letters, 115*(2), 229–231.

McAleer, M. (1994). Sherlock holmes and the search for truth: A diagnostic tale. *Journal of Economic Surveys, 8*, 317–370.

Meade, J. (1978). The meaning of 'Internal balance'. *Economic Journal*, *88*(351), 423–435.

Mizon, G. E. (1995). A simple message for autocorrelation correctors: Don't. *Journal of Econometrics, 69*(1), 267–288.

Morgan, J. (2022). Macroprudential institutionalism: The Bank of England's Financial Policy Committee and the contemporary limits of central bank policy. In P. Hawkins & I. Negru (Eds.), *Monetary Economics, Banking and Policy: Expanding economic thought to meet contemporary challenges. London:* Routledge.

Nasir, M. A. (2017). Zero Lower Bound & Negative Interest Rates: Choices for Monetary Policy. Feb 3, 2017. Available at SSRN: https://ssrn.com/abstract=2881926

Nasir, M. A., & Morgan, J. (2018). The unit root problem: Affinities between ergodicity and stationarity, its practical contradictions for central bank policy, and some consideration of alternatives. *Journal of Post Keynesian Economics, 41*(3), 339–363.

Nasir, M. A., & Morgan, J. (2020). Paradox of stationarity? A policy target dilemma for policymakers. *The Quarterly Review of Economics and Finance*. https://doi.org/10.1016/j.qref.2020.05.007

Olsen, W., & Morgan, J. (2005). A critical epistemology of analytical statistics: Addressing the sceptical realist. *Journal for the Theory of Social Behaviour, 35*(3), 255–284.

Patinkin, D. (1976). Keynes and econometrics: On the interaction between the macroeconomics revolutions in the interwar period. *Econometrica, 44*, 1091–1123.

Perron, P. (1989). The great crash, the oil price shock and the unit root hypothesis. *Econometrica*, *57*, 1361–1401.

Pesaran, H., & Smith, R. (1985). Keynes on econometrics. In T. Lawson and H. Pesaran (eds), cit.

Phillips, P. C. B. (1987). Towards a unified asymptotic theory for autoregression. *Biometrika*, *74*, 535–547.

Plosser, C., & Schvert, W. (1977). Estimation of a noninvertible moving average process: The case of over-differencing. *Journal of Econometrics, 6*, 199–224.

Pratten, S. (2005). Economics as progress: The LSE approach to econometric modelling and critical realism as programmes for research. *Cambridge Journal of Economics, 29*(2), 179–205.

Primiceri, G. E. (2005). Time varying structural Vector auto-regression and Monetary Policy. *Review of Economic Studies, 72*, 8214–8852.

Ranganathan, T., & Ananthakumar, U. (2010). Unit root test: Give it a break. In *30th International Symposium on Forecasting, International Institute of Forecasters*.

Reiss, J. (2015). *Causation, evidence, and inference*. Routledge.

Reiss, J. (2016). Suppes' probabilistic theory of causality and causal inference in economics. *Journal of Economic Methodology*, *23*(3), 289–304.

Rowley, R. (1988). The keynes-tinbergen exchange in retrospect. In O.F. Hamouda and J.N. Smithin (Eds.), cit.

Samuelson, P. (1946). Lord keynes and the general theory. *Econometrica, 14*, 187–200.

Sargan, J. (1964). Wages and prices in the UK: A study in econometric methodology. In P. Hart, G. Mills, & J. Whittaker (Eds.), *Econometric analysis for national planning.* Butterworths.

Sargan, J. (1964). "Wages and prices in the UK: a study in Econometric Methodology.". In P. Hart, G. Mills, & J. Whittaker (Eds.), *Econometric Analysis for National Planning*. London: Butterworths.

Schotman, P. C., & Van Dijk, H. K. (1991). On bayesian routes to unit roots. *Journal of Applied Econometrics, 6*(4), 387–401.

Sims, C. (1989). A nine variable probabilistic model of the US economy. *Institute for Empirical Macroeconomics, Federal Reserve of Minneapolis*, Discussion Paper, 14.

Sober, E. (2001). Venetian sea levels, british bread prices, and the principle of the common cause. *British Journal for the Philosophy of Science, 52*, 331–346.

Spanos, A. (2015). Revisiting Haavelmo's structural econometrics: Bridging the gap between theory and data. *Journal of Economic Methodology, 22*(2), 171–196.

Stock, J. H. (1991). Bayesian approaches to the 'unit root' problem: A comment. *Journal of Applied Econometrics, 6*(4), 403–411.

Stone, R. (1978). Keynes, Political Arithmetic and Econometrics. In Proceedings of the British Academy, 64, Oxford: Oxford University Press.

Syll, L. P. (2016). *On the use and misuse of theories and models in mainstream economics*. College Publications/WEA Books.

Tinbergen, J. (1939). *Statistical testing of business-cycle theories: II. Business cycles in the United States of America, 1919–1932*. League of Nations, Economic Intelligence Service.

Tinbergen, J. (1940). On a method of statistical business research. A reply. *Economic Journal, 50*, 141–154.

Tobin, J. (1980). Stabilization policy ten years after. Brookings Papers on Economic Activity, Volume, 11: pp. 19–90.

Ullah, S., Akhtar, P., & Zaefarian, G. (2018). Dealing with endogeneity bias: The generalized method of moments (GMM) for panel data. *Industrial Marketing Management, 71*, 69–78.

Ullah, S., Zaefarian, G., & Ullah, F. (2021). How to use instrumental variables in addressing endogeneity? A step-by-step procedure for non-specialists. *Industrial Marketing Management, 96*, A1–A6.

Wrane, A. (2006). Bayesian Inference in Co-integrated VAR model with Applications to the Demand for Euro Area M3. European Central Bank, Working Paper Series No 692 / November.

Yule, G. U. (1926). Why do we sometimes get nonsense correlations between time-series? *Journal of the Royal Statistical Society, 89*, 2–9.