On a Definition of Trend

Silva Lopes, Artur

14 April 2025

On a Definition of Trend

Artur Silva Lopes [*]
independent researcher

This version: April 14, 2025

Abstract

Several reasons explain the absence of a precise, complete and widely accepted definition of trend for economic time series, and the existence of two major disparate models is one of the most important. A recent operational proposal tried to overcome this difficulty resorting to a statistical test with good asymptotic properties against both those alternatives. However, this proposal may be criticized because it rests on a tool for inductive, not deductive, inference. Besides criticizing this recent definition, drawing heavily on previous ones, the paper provides a new proposal, more complete, containing several necessary but no sufficient condition(s).

Keywords: trend; time series; macroeconomy; statistical testing.
JEL codes: C12, C18, C22, B41.

[*]Email: arturslopes(at)proton.me.

1 Introduction

Since trends are often the dominant visual feature of many economic time series, the concept of trend is essentially an empirical, even graphical one. Any line joining the first and the last data points, progressing smoothly and monotonically throughout the sample will do the trick. Often, a simple straight line is deemed sufficient but sometimes, besides some sophistication, only a slowly bending curve provides the desired approximation to the data points. This is possibly the major reason explaining why there is no precise, complete and widely accepted definition of trend. Several difficulties contribute to this situation too and in the next paragraphs I point to some of them.

First, trends are unobservable phenomena that exist only in the minds of researchers using the framework of the traditional decomposition of economic time series, inspired in the work of the astronomers of the 17th and 18th centuries. As is well known, the decomposition reads as

$$y_t = \tau_t, + c_t + s_t + i_t,$$

τ_t denoting the trend, c_t the cycle, s_t the seasonal and i_t the irregular component, and its description may be found in, e.g., Persons (1919).

Second, economic research has been focusing mostly (almost exclusively) on business cycles, with trends considered a nuisance, as "that which trend filters remove" (White and Granger, 2011). Only recently research on trends has gained some importance under the threat of climate change.

Third, the radically opposing features that characterize the two major trend models, the linear deterministic and the random walk type trend, are very difficult to enclose in a single definition. The first corresponds to the traditional notion of trend, which appears to be still dominant, though the second, with its typical wandering behaviour, has been increasingly acknowledged.

Fourth, a lack of guidance from economic theory about what a trend represents and which should be its properties in relation to the other components of the decomposition above. Though it ceased to be dominant some time ago, the idea that trends should be orthogonal to cycles and should be driven by entirely distinct forces still persists.

In practice, several vague and partial definitions of trends coexist, none dominating the others and serving as solid ground to develop research. Phillips (2010) even blames this omission for some retardment of progression in knowledge in this area, questioning

> is it possible to measure and discuss with clarity any quantity that is undefined?

And indeed, although the dispersion of detrending methods, corresponding to distinct notions of trend, has served well the analysis of cycles, providing several distinct angles of view, it has probably delayed research in, e.g., growth theory and in economic history.

In an interesting paper addressing the presence of trends in local and global temperature data, Rivas and Gonzalo (2020, RG2020 hereafter) propose a "practical definition" of trend which ingeniously resorts to the asymptotic properties of the standard textbook test for a linear trend to wrap the two disparate models of trend into a single definition. Asymptotically, the test allows that any trending time series, whether evolving linearly and deterministically or similarly to a random walk trend, will be correctly classified as such. This is because the test is consistent against both alternatives. The problem is that resorting to a statistical test to base an unambiguous, undisputable and apparently error- and uncertainty-free statement about the true nature of a time series is not possible. Statistical tests are tools for inductive inference and do not permit making such type of statements. As a testing procedure, the method proposed in RG2020 possesses nice, well known asymptotic power properties. However, these properties are not sufficient to warrant such peremptory statements as "there is an increasing trend" (p. 153) or "there is a trend component in all distributional characteristics of interest" (p. 167). Type I errors (finding evidence for a trend when none is present) cannot be avoided even asymptotically and both type I and type II (not finding evidence for an existing trend) errors may well be present in real, finite sample situations, and therefore such strong declarations about the true nature of time series are not admissible. That is, Rivas and Gonzalo propose using a statistical test to assert peremptorily and unambiguously the true trending nature of a time series, neglecting any uncertainty and the possibility of error that such a tool for inductive inference entails.

In this paper I criticize this definition and provide a simple alternative proposal whose major purpose is to integrate various different previous proposals, gathering several necessary conditions that were previously dispersed. The goal is therefore to improve completeness of necessary conditions, making the definition more clarifying, not precision, as no single sufficient condition is provided. That is, the proposed definition is of a much different character from that in RG2020.

The remainder of the paper is organised as follows. The next two sections provide an overview of the literature which is necessary to understand the definition in RG2020. Section 2 briefly describes the two major trend models which are encompassed by the definition in RG2020 and section 3 reviews the theoretical results on inference for the linear trend model, which are key to understand the merits and the shortcomings of the definition. Section 4 presents and criticizes the definition in RG2020 and illustrates the criticisms with a Monte Carlo study whose results are not completely new but which serve to clarify the shortcomings that it entails. In section 5 an alternative definition is proposed and commented and section 6 concludes the paper.

2 The major types of trends

The most common notion of trend is associated with the well known linear deterministic trend model,

$$y_t = \alpha + \beta\, t + u_t, \quad t = 1, 2, \ldots, T, \tag{1}$$

u_t denoting a zero mean, weakly dependent and stationary, I(0), process. As is well known, y_t may denote a previously logarithmized time series, in which case the model for the original series is the exponential trend one, $x_t = \exp(\alpha + \beta t + u_t)$, and hence $y_t = \log(x_t)$. Model (1) is commonly known as defining a trend stationary process (TSP).

The rival trending model is the random walk type trend which, in the most simple case is generated by the simple, no drift, random walk model,

$$y_t = y_{t-1} + \epsilon_t, \quad \epsilon_t \sim iid(0, \sigma_\epsilon^2), \tag{2}$$

and which can be generalized replacing the ϵ_t process with, say, a process such as u_t of (1). As is well known, this is the most simple nonstationary process which is integrated of order 1, I(1), i.e., a difference stationary process (DSP).

While the first process consists almost only of a deterministic component, in the second the trend is purely stochastic as it is the accumulation of random shocks ($\sum_{i=1}^{t} u_i$ or $\sum_{i=1}^{t} \epsilon_i$ in the simplest case). To further complicate matters, a stationary stochastic component is present in (1) while a deterministic linear trend may coexist peacefully with the stochastic one of (2) provided a simple drift parameter is inserted in the right hand side of (2) (e.g., $y_t = \beta + y_{t-1} + \epsilon_t$). It is for this reason that the problem of selecting one of the two models is not simply one of a dichotomy of a deterministic *vs.* stochastic trend, as is sometimes stated in the literature[1]. It is also for this reason that I prefer to name the second type of trend as random walk type and not simply as stochastic.

3 Inference in the linear trend model

This section briefly reviews the properties of OLS estimation and testing for the (deterministic) linear trend model, as the definition proposed in RG2020 is based precisely on that model. In the next section that definition will be reviewed and commented. In this section, in the first subsection it is assumed that the data generation process (DGP) coincides with the model in what concerns the long-run properties (weak dependence and stationarity) and the presence of the linear trend. In the second subsection this assumption will be removed, the data assumed as generated by an I(1) process, and hence the linear trend model becomes a well known case of a spurious regression.

3.1 Coincidence between model and DGP

To scrutinize inference properties of the method underpinning RG2020's definition, I first consider the favourable case where the DGP is indeed a TSP, the linear trend model as assumed, so that there is coincidence between the model and the DGP. Moreover, in the most friendly and simplest TSP case, the error term is a(n independent) white noise process with finite fourth moment. Hence, both the DGP and the model are represented with $y_t = \alpha + \beta t + \epsilon_t$, $\epsilon_t \sim iid(0, \sigma^2)$.

[1]The statement that unit root tests perform the role of selecting between these two opposed views for macroeconomic time series belongs to common or popular wisdom but it is incorrect.

This is a case that I designate as the purely deterministic linear trend process because the stochastic component is such an unpredictable and uninformative process.

This is the most benign case of non-stationarity because although the regressor matrix does not comply with the standard assumption for stationary processes, of convergence in probability to a regular matrix[2], it can be shown that the OLS estimator for β, $\hat{\beta}$, is not only consistent but it is even hyperconsistent, as it converges to β at rate $T^{3/2}$:

$$T^{3/2}(\hat{\beta} - \beta) \xrightarrow{d} \mathcal{N}(0, 12\sigma^2), \tag{3}$$

see, e.g., Hamilton (1994, pp. 457-60).

On the other hand, the traditional standard error of $\hat{\beta}$ ($se(\hat{\beta})$) converges also at an unusual rate of T, resulting in t-tests that are asymptotically valid. In particular, for the t-ratio for β, i.e., to test $H_0 : \beta = 0$ *vs.* $H_1 : \beta \neq 0$ in model (1):

$$t_{\hat{\beta}} = \hat{\beta}/se(\hat{\beta}) \xrightarrow{d} \mathcal{N}(0, 1),$$

when H_0 is true. This is the usual or standard test for a (linear) trend that is presented in textbooks.

Moreover, in case the gaussianity assumption is added, i.e., if $\epsilon_t \sim iid\mathcal{N}(0, \sigma^2)$, asymptotic properties become pointless because the OLS estimator can be shown to be the best unbiased estimator (BUE) and usual inference methods become valid exactly, i.e., in finite samples. This is a case that is even more favourable then the classical linear regression model because, since the only regressor (time) is deterministic, it is indeed "fixed in repeated samples", as was assumed in old textbooks.

Consider now a still favourable but not so positive case, the assumed model the purely deterministic one (with white noise errors) but the DGP differing from it in the properties of the error term, now serially correlated though stationary and possibly non-Gaussian too. As one of the basic assumptions of the Gauss-Markov theorem is not satisfied, OLS is now not even BLUE (best linear unbiased). However, provided the error term is a very general I(0) process,

[2] This standard condition requires that $\text{plim}\left(\frac{1}{T}\mathbf{X}'\mathbf{X}\right)$, where $\mathbf{X}$ denotes the regressor matrix, must be a regular positive definite matrix.

according to a result obtained originally by Grenander and Rosenblatt[3], OLS maintains its hyperconsistency property and it is asymptotically equivalent to GLS and hence asymptotically efficient.

As in (3) asymptotic normality still holds but as the asymptotic variance now depends on nuisance parameters, its usual estimator is generally biased and inconsistent. Hence, inferences about β based on usual OLS formulas cease to be valid, even asymptotically. However, since it is common practice to accompany regression estimates with some diagnostic statistics, it is very unlikely that serial correlation will pass unnoticed and, e.g., a Breusch-Godfrey test statistic will in general be able to detect the problem. Provided there is some first order serial correlation, even the Durbin-Watson statistic will likely detect it as the only regressor is deterministic. And since heteroskedastic and autocorrelation consistent (HAC) estimation of the variance is also routine, it is very likely that its most popular version, a Newey-West corrected standard error is reported and the corresponding robust t-ratio is employed. This t-ratio may be written as

$$t^{HAC} = \frac{\widehat{\beta}}{\sqrt{\widehat{\sigma}_u^2 \left[\sum_{t=1}^{T}(t-\bar{t})^2\right]^{-1}}}, \tag{4}$$

where $\widehat{\sigma}_u^2$ denotes an estimator of the so called long-run error variance (see, e.g., section 5.3 in Lopes (2025)), and it is asymptotically standard normal under the null hypothesis.

Despite its popularity, this method is often unsatisfactory. Size problems of the tests are somewhat attenuated only, not completely eliminated in small samples, where rejections of the true null hypothesis far exceeding the nominal size may continue to subsist[4]. The problem may be particularly acute when first order serial correlation is positive and strong (though stationary). Kiefer and Vogelsang (2005) proposed a method allowing a much better control of the size of the tests, aiming directly to their performance and sacrificing consistent

[3]Grenander, and M. Rosenblatt (1957). *Statistical Analysis of Stationary Time series*, Wiley, New York. The general conditions are clearly stated in, e.g., Canjels and Watson (1997).

[4]Both OLS and even HAC estimation of the variance tend to overestimate the precision of $\hat{\beta}$, with reduced standard errors, and therefore t-ratios (and t-statistics in general) become inflated, rejecting too often a true null hypothesis.

variance estimation. Sometimes dubbed as HAR (Heteroskedasticity and Autocorrelation Robust) testing, it involves an elaborate procedure, which has not yet been widely adopted. Therefore, I assume that a representative practitioner will employ simply a popular HAC method.

3.2 The random walk trend as the DGP

In this case the DGP is the random walk type trend but the assumed model continues to be the linear deterministic trend model of equation (1), which is the base of the standard textbook test for a trend. I will start with the driftless case and later mention the case with drift.

This case was firstly addressed by Nelson and Kang (1984) in an influential and disturbing article showing, through Monte Carlo evidence, that the linear trend model fitted the random walk surprisingly "well", actually producing very often erroneous inferences supporting the presence of the deterministic trend. That is, despite the absence of the deterministic trend in the data, the standard test provides spurious, favourable evidence for its presence in a much higher proportion of the cases than the nominal size of the tests should allow. Table 1 illustrates this problem for the case of a simple random walk, with Gaussian white noise innovations.

Table 1 - Percentage rejections of the true null hypothesis that $\beta = 0$ in (1) when the DGP is a simple random walk

$T(L)$	T=30 (2)	50 (3)	100 (4)	200 (5)	500 (6)	1000 (8)
t_{OLS}	77.36	82.61	87.55	91.42	94.41	96.05
t_{HAC}	69.23	71.09	75.87	80.55	86.19	88.80

The DGP is $y_t = y_{t-1} + \epsilon_t, \epsilon_t \sim iid\mathcal{N}(0, \sigma^2)$ and in each replication the first 50 observations were discarded. The nominal size of the two-sided test ($H_1 : \beta \neq 0$) is 5% and the size estimates are based on 20000 replications. t_{OLS} and t_{HAC} refer to the percentage rejections with the standard OLS and with the HAC t-ratio using the Newey-West version with a Bartlett kernel and a lag truncation parameter denoted with L, respectively.

The first line of results in the table reveals a very serious, dramatic problem of over-rejections of a true null hypothesis, the standard test (denoted with t_{OLS}) very frequently and misleadingly detecting the presence of a deterministic linear trend which is really absent. Moreover, the problem only gets worse with growing sample sizes as a consequence of a test statistic that diverges (to infinity) as $T \to \infty$.

Indeed, Durlauf and Phillips (1988) demonstrated that the t-ratio statistic diverges. This occurs despite the convergence of the OLS estimator of β to its true value of zero, i.e., they also showed that

$$T^{1/2}\hat{\beta} \xrightarrow{d} \mathcal{N}(0, 6\sigma_{LR}^2/5),$$

where σ_{LR}^2 represents the long-run variance of the innovations of the I(1) process (in case it is a white noise process this is simply its variance), which contrasts with the inconsistency of $\hat{\alpha}$, whose distribution even diverges as T grows. What is most relevant here, however, is the spurious performance of the standard test, erroneously indicating the presence of a non-existent linear trend.

Durlauf and Phillips (1988) analysed the case with drift as well, and showed that the behaviour of both the OLS estimators and the test statistics is identical to the no drift case. In what concerns the standard test, the seriousness of the spurious problem now vanishes as the drift parameter implies that a linear deterministic trend is indeed really present in the data.

4 The "practical definition" of trend of Rivas and Gonzalo (2020)

Apparently trying to advance in the path of making the concept of trend more precise, Rivas and Gonzalo (2020, RG2020) proposed an operational, "*practical definition*" of trend (p. 158), their definition 7:

> a characteristic C_t of a functional stochastic process X_t contains a trend if in the [O]LS regression
>
> $$C_t = \alpha + \beta t + u_t, \quad t = 1, \ldots, T,$$
>
> $\beta = 0$ is rejected.

This equation is not assumed to represent the DGP. It is only instrumental for their definition, as it must be seen as the "*linear LS approximation of an* *unknown trend function h(t)*". Therefore, there is no assumption regarding the properties of the error term and since the dependent variable may be I(1), u_t may be I(1) as well[5].

As stressed in Lopes (2025), RG2020 are resorting to the asymptotic power properties of the standard test to aid in their definition of trend:

a) when $y_t \sim I(0)$, a TSP, standard asymptotic theory is valid and since the test is consistent it asymptotically classifies the series as trending;

b) when y_t is a driftless $I(1)$ process, since the statistic diverges, asymptotically the test will also always reject the null, thereby producing evidence for a trend, which actually exists but it is a stochastic one in this case, not the deterministic linear trend of the alternative hypothesis;

c) when $y_t \sim I(1)$ with drift, since the asymptotic behaviour of the t-statistic is identical to the driftless case, the trend will also be always detected in the limit as $T \to \infty$, and it is present in both forms, deterministic and stochastic, in this case.

Therefore, in the limit, as $T \to \infty$, any trending I(0) or I(1) time series will be considered as such. Note, however, that wrong classifications of non-trending time series (as trending) are not precluded, they will occur even asymptotically because the size of the tests is fixed, it does not vanish as T grows.

Inasmuch as it relies on an outcome of a statistical test, this definition cannot be considered an Aristotelian one. It does not provide the essence of a trend, it neither describes nor explains what a trend is. It can hardly be considered as complying with the principle stated in Burge (1993, p. 311)[6]:

> Definitions associated with concepts fix necessary and sufficient conditions for falling under the concept. They give the essence, or if not the essence at least the most fundamental individuating conditions of the entities that the concept applies to.

[5]Besides the linear deterministic trend and the random walk type trend, RG2020 further claim that their analysis is also valid for fractionally integrated, for near unit root and for local model trends. As in RG2020, I will keep the analysis confined to the two first cases.

[6]As will become more clear below, I am considering that the outcome of a statistical test cannot provide any " fundamental individuating condition".

From a strictly logical point of view the definition is also invalid, as b) above allows anticipating: when the null hypothesis is rejected because the process is a driftless I(1) one, it is correctly considered as trending; but the statement that $\beta \neq 0$ is not true because β exists only in the linear trend model. This is a case of a correct classification for a wrong reason because the parameter β does not even exists.

Supporting a definition on a statistical test violates the conservativeness criterion as well. According to this criterion, any definition should not only lead to any inconsistency, but it should also not lead to anything that was not obtainable before (see, e.g., Bellnap, 1993). Actually, and besides the previously mentioned cases of rejections for non-trending processes, one may also consider the cases of those trending processes for which the sample size is not sufficient, is not large enough to warrant that the rejection of the hypothesis is certain. That is, type II errors are indeed prevented but only asymptotically, and known trending processes may not be recognized as such in small samples with the new definition.

A more general but also deeper comment than the previous considerations was made in Lopes (2025):

> it does not appear to be admissible that a definition regarding the presence or the absence of a characteristic [in a time series] might depend entirely on the outcome of a statistical test (however powerful it might be). A finding of evidence can surely depend, but that is not the same as an undisputable, uncertainty-free statement about the true nature of a series. To say that there is evidence on the presence of a trend is not the same as asserting unambiguously that one is present.

This is indeed the most important shortcoming in the definition put forward in RG2020: it is based on a tool for statistical, inductive inference, not on a set of unambiguous necessary and sufficient conditions characterizing trends and allowing deductive reasoning. That is, an operational (or "practical") definition cannot rely on a statistical test, which is always susceptible to error and to uncertainty. It is as if there are neither errors nor uncertainty associated with the outcomes of statistical tests. Essential or fundamental individuating, defining conditions cannot depend on such fallible instruments. Two simple examples may help to clarify these statements:

a) a (possibly weakly[7]) trending time series observed over a sample whose size is insufficient for the test to reject the false null hypothesis is incorrectly *defined*, not only tentatively considered as is usual with statistical tests, as non-trending;

b) a non-trending stationary process, possibly observed also over a small sample and unfortunately slightly behaving as a trending one may be *defined* as trending in some cases where the test incorrectly rejects, in a much stronger statement than the one that is usually allowed in statistical inference, say, "some evidence for the presence of a trend was found with a 5% level test".

While in both cases inference is admissible and valid provided its statement is made in terms of the evidence that was found, using such evidence to unambiguously assert the true nature of the series is not admissible. Both examples demonstrate also that the definition in RG2020 is not "extensionally adequate" (see Gupta and Mackereth, 2023) in terms of descriptive adequacy because there are actual counterexamples to it.

The definition is ingeniously based on a statistical test with good asymptotic properties, better than many tests and consistent against the two most important and disparate alternative DGPs. Nevertheless, its capacity cannot transcend that of a statistical test. Rigour and precision are much more apparent than real.

Moreover, the definition is prone to some arbitrariness too, because inferences partially depend on the chosen significance level. Sticking to the usual 5% eliminates the variability of the decisions but requires a very strong but necessarily incomplete justification. Only a 0% significance level test would allow eliminating type I errors — always classifying non-trending series correctly —, but consistency would be lost and the inconsistency would be extreme as power would also be trivially 0%.

[7] In the common sense of a small slope coefficient, as in White and Granger (2011), i.e., I am not introducing any new concept.

Notice further that the problem does not lie in the unobservable nature of trends. Trends are components of economic time series that are not observed, as the cyclical, seasonal and irregular components, but even if they were observed the problem would remain unaffected. It lies in the inductive, and hence uncertain and error prone nature of the inference process, which is unavoidable and admissible at the stage of gathering evidence but not as a pillar of a definition. Actually, what Rivas and Gonzalo propose is a testing procedure for the detection of trends, not a real definition, as they present the usual simulation study over a wide variety of processes to assess the size and power properties of their method over finite samples. That is, they implicitly acknowledge that their method incurs on type I and type II errors, as any other statistical test, and hence it is incapable of performing the role of a criterion for deductive inference.

RG2020 further recommend using the robust HAC version of the standard test. Following this recommendation, already in Table 1 the line entitled with t_{NW} presents the percentages of findings of a trend for the driftless random walk case, i.e., of correct classifications of trending time series (though the trend is not really deterministic). Although in this case the method should not even be employed because it is not strictly valid in non-stationary environments, as shown in the table its impact is unsurprising: for all the sample sizes the percentage of rejections is reduced relatively to the standard OLS version, making the detection of the trend more difficult. That is, as HAC based inference attenuates the spurious rejection problem, an undesired effect of reducing the evidence for the presence of a trend occurs in this case.

This same effect can be observed in most of the cases of Table 2, when the DGP is a TSP, in particular for all the cases where the error term is serially correlated, that is, for the last four blocks of cells of the table. Basing inference on the HAC version of the statistic decreases the power of the trend test, worsening the performance of RG2020's definition. Summarizing these two cases, one can state that regardless of the order of integration, provided there is some kind of serial correlation in the data, the HAC version of the test does not improve the performance of the classification method, instead it further worsens it. On the other hand, as shown in the first block of the same table, somewhat surprisingly the HAC test does improve the power of the test in small samples when there is no serial correlation.

Table 2 - Percentage rejections of the false null hypothesis that $\beta = 0$ in (1) when the DGP is a trend stationary process

$T(L)$	30 (2)	50 (3)	100 (4)	200 (5)	500 (6)	1000 (8)
	DGP: $y_t = 3 + 0.005t + \epsilon_t,\ \epsilon_t \sim iid\mathcal{N}(0,1)$					
t_{OLS}	6.63	8.78	30.68	98.20	100.0	100.0
t_{HAC}	10.44	12.37	34.73	98.38	100.0	100.0
	DGP: $y_t = 3 + 0.005t + \epsilon_t + 0.4\epsilon_{t-1},\ \epsilon_t \sim iid\mathcal{N}(0,1)$					
t_{OLS}	15.61	16.51	32.95	92.08	100.0	100.0
t_{HAC}	13.80	13.08	24.34	85.97	100.0	100.0
	DGP: $y_t = 3 + 0.005t + \epsilon_t + 0.8\epsilon_{t-1},\ \epsilon_t \sim iid\mathcal{N}(0,1)$					
t_{OLS}	18.93	19.22	29.67	81.13	100.0	100.0
t_{HAC}	14.71	12.85	18.91	67.61	100.0	100.0
	DGP: $y_t = 3 + 0.005t + u_t,\ u_t = 0.4u_{t-1} + \epsilon_t,\ \epsilon_t \sim iid\mathcal{N}(0,1)$					
t_{OLS}	21.99	22.67	36.22	88.19	100.0	100.0
t_{HAC}	18.25	15.86	23.32	77.08	100.0	100.0
	DGP: $y_t = 3 + 0.005t + u_t,\ u_t = 0.8u_{t-1} + \epsilon_t,\ \epsilon_t \sim iid\mathcal{N}(0,1)$					
t_{OLS}	53.42	52.82	54.34	64.59	99.57	100.0
t_{HAC}	41.69	34.34	28.96	36.54	97.12	100.0

In each replication the first 50 observations were discarded. The nominal size of the two-sided test is 5% and the power estimates are based on 20000 replications. t_{OLS} and t_{HAC} refer to the percentage rejections with the standard OLS and with the HAC t-ratio using the Newey-West version with a Bartlett kernel and a lag truncation parameter denoted with L, respectively.

As shown in Table 3, when the data are generated by a stationary, trendless and weakly serially correlated process, as in the last four blocks of the table, the autocorrelation robust version of the test does improve the classification when there is really no trend, neither deterministic nor stochastic. A generalized picture of size distortions clearly emerges from that table, and as expected the autocorrelation robust test lessens the problem particularly in large samples. Nevertheless, even for samples as large as $T = 1000$ it does not succeed in eliminating the problem completely. And obviously, an adequate classification method, unattainable with any statistical test, should always "accept" the null hypothesis in this context, producing 0% type I errors.

Table 3 - Percentage rejections of the true null hypothesis that $\beta = 0$ in (1) when there is no trend in the DGP

$T(L)$	30 (2)	50 (3)	100 (4)	200 (5)	500 (6)	1000 (8)
DGP: $y_t = \epsilon_t,\ \epsilon_t \sim iid\mathcal{N}(0,1)$						
t_{OLS}	6.04	5.71	5.40	5.12	4.99	5.28
t_{HAC}	9.92	8.67	7.39	6.35	5.56	5.61
DGP: $y_t = \epsilon_t + 0.4\epsilon_{t-1},\ \epsilon_t \sim iid\mathcal{N}(0,1)$						
t_{OLS}	15.22	14.09	13.84	13.68	13.39	13.77
t_{HAC}	13.47	10.97	8.87	7.45	6.41	6.19
DGP: $y_t = \epsilon_t + 0.8\epsilon_{t-1},\ \epsilon_t \sim iid\mathcal{N}(0,1)$						
t_{OLS}	18.76	17.38	17.13	16.53	16.66	17.03
t_{HAC}	14.22	11.58	9.22	7.66	6.47	6.39
DGP: $y_t = 0.4y_{t-1} + \epsilon_t,\ \epsilon_t \sim iid\mathcal{N}(0,1)$						
t_{OLS}	21.69	20.53	20.30	20.09	20.18	20.54
t_{HAC}	18.00	14.06	11.04	8.88	7.64	7.09
DGP: $y_t = 0.8y_{t-1} + \epsilon_t,\ \epsilon_t \sim iid\mathcal{N}(0,1)$						
t_{OLS}	53.22	52.40	52.07	51.76	51.07	51.97
t_{HAC}	41.63	34.10	26.92	21.85	18.04	14.99

In each replication the first 50 observations were discarded. The nominal size of the two-sided test is 5% and the size estimates are based on 20000 replications. t_{OLS} and t_{HAC} refer to the percentage rejections with the standard OLS and with the HAC t-ratio using the Newey-West version with a Bartlett kernel and a lag truncation parameter denoted with L, respectively.

5 A simple alternative proposal

The definition in RG2020 was hardly proposed to fill a wide gap of some absence of a general definition. Actually, there are many general definitions, though none also sufficiently complete and/or precise to be widely adopted. However, this dispersion is perhaps more beneficial than detrimental because the variety of detrending methods that it originates leads to a corresponding variety of business cycle estimates. This may be useful to illuminate different angles of view of short- and medium-term economic fluctuations which may require, say, different types of measures of economic policy. On the other hand, since the focus of almost all research is on business cycles, there has been rarely (if ever) any interest in comparing different trend estimates.

Notwithstanding this, consistency with the previous criticisms requires that they are accompanied by a proposal of an alternative definition. The purpose is not to propose a set of conditions to select the best estimate of the trend, making the different estimates collapsing to a single one, as it may be advantageous to have a variety of estimates. For instance, this might be also helpful for a thorough discussion about the forces that drive trends. The purpose is also not one of enhancing precision so that more accurate estimates result. Actually, in the absence of a “true trend”, accuracy is impossible to measure. Rather, it is meant only to adapt to the most frequent usages of the term, respecting those usages while enclosing in a single definition the two most important notions of trends.

The proposed alternative is the following:

trends are non-stationary, non-oscillatory, slowly evolving processes corresponding to persistent, underlying long-term movements of economic variables, displaying only low-frequency variation. Estimated trends must provide some information about the past behaviour and the current position of the variables, and they must also indicate the likely direction of their future evolution.

Several remarks must be made about this proposal.

a) The proposal does not aim to establish “the definition” of trend, only “a definition”. Other definitions are admissible. However, at the current state of knowledge it seems unlikely that a more precise definition could be formulated. This also means that several of the limitations pointed to the definition in RG2020 are applicable to this proposal as well.

b) The proposal is meant to apply to economic trends but hopefully it may suit other types of trends as well (in e.g., climatology and demography).

c) The proposal does not aim to be original. It aims to collect and integrate a wide variety of different contributions from several sources, so as to establish the maximum number of characteristics that must be *common* to trend processes, and particularly to the two main rival trend processes. In particular, the last sentence heavily borrows from Phillips (2010), where an intuitive and graphical description of trends is provided which is simultaneously simple and enlightening. Phillips further stresses the importance of the predictive information that must be present in estimated trends:

> A trend line summarizes [these] primitive requirements. It summarizes where we have been, shows where we are now in relation to the past, and most of all, reveals a hint of where we are going.

However, though dominant, and dating back at least to Harvey (1989), this prevalence of the predictive content is not completely common to all available notions.

d) The proposal is intentionally general, vague and non-operational[8], containing only the information that is strictly necessary and nothing else, and respecting previous usages of the term, in a somewhat Wittgenstein fashion. Therefore, it complies with the conservativeness criterion and it tries to maintain the cohesion of this area of research preventing its fragmentation.

e) The only condition that appears to be uncommon, perhaps even original, is the one on non-stationarity. Its purpose is to strengthen the distinction from the cyclical component, which must be stationary. Moreover, if the trend is meant to represent the permanent but evolving component of a time series it must be non-stationary.

f) The proposed definition is mostly *nominal*, graphical or visual, not real, as is usual with the notion of trend, to distinguish it from other forms of variation of time series, but without accessing the underlying structure of trends, i.e., their essence. See also g) and j) below.

g) Despite sharing with the definition in RG2020 the criticism of a missing essence, the proposal complies with the minimum criterion stated in Burge

[8]See Swartz (2010) for a critique of operational definitions, emphasizing their perverse effects on science because they tend to fragment areas of research.

(1993), as it provides “the most fundamental individuating conditions” of trends.

h) Non-stationarity may refer to the mean and/or to the variance. The linear deterministic trend model, as well as many other, nonlinear deterministic models, are non-stationary in mean, whereas the driftless random walk type trend models are non-stationary in variance.

i) Visually, the non-oscillatory or non-cyclical feature is the main distinctive feature from the two other major components of the classical decomposition, the cycle and the seasonal components.

j) Although persistent long-term movements are mentioned, their origin is deliberately omitted. The forces that drive economic trends may be themselves persistent (as, e.g., technology or the stocks of human and physical capital in the case of output) or may be instead transitory but impart permanent or quasi-permanent effects (e.g., a large oil shock, a particularly important technological innovation, the outbreak of a pandemic virus or of a war, etc.). Economic theory has not yet reached a consensus about these forces and, for instance, in a survey of macroeconomic models with “scarring”-type effects, i.e., negative long-run effects, Cerras, Fatás and Saxena (2023) question the old popular wisdom according to which only supply shocks have long-term effects. Although these issues are still unsettled, the often assumed orthogonality in relation to cyclical and seasonal movements must be ruled out.

k) The terms long-run and low frequency are not easy to define. Their concrete meaning depends on the frequency of the observations, on the sample size and on the subject under study. Quantitatively, in the time domain, Müller and Watson (2024) establish that it corresponds to a time period that must represent “a non-negligible fraction of the sample period”. And they provide a few examples: “for example, when studying 70 years of post-WWII quarterly data, decadal variation is low-frequency, and when studying a decade of daily return data, yearly variation is low-frequency”. Moreover, as demonstrated in, e.g., Seater (1993) and Phillips (2010), the subject area is likely to perform a significant role: samples of “only a century” duration are insufficient to make correct inferences about climate change, i.e., about trends in temperature.

l) The inclusion of random walk type trends implies that features that are typical only of linear trends must be abandoned. Some examples are the following: characterized by “systematic upward or downward evolution with time”, "having a direction" and "monotonically throughout" (White and Granger, 2011), representing "regular or regularly changing" behaviour through long periods of time.

m) A special mention must be made to the smoothness condition, which is one of the most popularly assumed to hold but which I have omitted. For many, trend is even synonymous of smooth. The origin of this idea dates back to the beginning of the 20th century, to the Spencer graduation or moving average filter to smooth mortality rates and it was later also adopted in Leser's (1961) smoothing method that is known in economics as the Hodrick-Prescott filter. According to this approach, any time series can be decomposed into only two components, the trend (f_t) and the disturbance (c_t),

$$y_t = f_t + c_t,$$

and in this original decomposition, which was not meant to be applied to macroeconomics, there was no space to business cycles[9]. This means that either cycles are included in f_t — and possibly this is the reason why sometimes one find references to the "trend-cycle" — or the traditional decomposition with its irregular component does not hold and cycles are represented by the "perturbations", c_t. For the current purpose, it must be stressed that this trend component hardly has anything to do with what economists nowadays call trend. Its most important feature, if not the only one, was that it should be *smooth*, free from irregularities.

Hodrick and Prescott (1981, 1997) later adopted Leser's framework and opted for the second possibility, with c_t representing the cyclical component and f_t the isolated trend component, inevitably smooth, as in Spencer's mortality rates. And it is this smoothness condition that presides over the choice of the (smoothness penalty) parameter λ, not any optimality criterion based on some principle (see more on this in Lopes, 2025, section 4.3).

In conclusion then, besides the imagination of most economists, there is no palpable reason why trends must be smooth. As Harvey (2016) simply puts it:

> there is no fundamental reason, though, why a trend should be smooth, except that it is somewhat easier on the eye.

[9]In filtering or signal extraction theory, f_t is the *signal*, that needs to be extracted from the data, and c_t is the *noise*.

Therefore, procedures that are known to generate estimated trends that tend to be noisy, as the Nelson and Plosser (1982) one or the "Hamilton (2018) filter" are admissible and should not be downgraded due to some jaggedness.

At the empirical level, both the linear trend model and the driftless random walk type model usually generate smooth trends, particularly the first one, as random walks, though smoother than stationary processes, often display relatively long periods with some turbulence. Nevertheless, even observed linear trends very often display sharp, abrupt breaks in level that are usually considered as manifestations of the irregular component. Smooth trends are then obtained with curves joining a pre-break period with a post-break one, which means that global linearity cannot hold. An alternative has been to consider piecewise or segmented linear trends and hence piecewise smoothness as well.

6 Concluding remarks

The existence of several distinct models for the trend of economic time series, and particularly of two major disparate models, is one of the reasons explaining the absence of a precise, complete and widely accepted definition of trend. Taking advantage of the asymptotic properties of the standard textbook test for a (linear) trend, Rivas and Gonzalo (2020) have put forward an operational, "practical definition", enclosing those two major types of trends — the linear deterministic and the random walk type — into a single definition.

However, the tool they use to wrap the two cases is fragile and does not provide a sound basis for a definition, at least for one which purports to be operational. Statistical tests are tools for inductive inference and, no matter how powerful they are, they do not warrant uncertainty- and error-free statements. Some arbitrariness is also unavoidable as the outcome of the tests may depend on the chosen level of significance.

Despite David Hume's paradox on induction, science must often rely on it. However, when it comes to the conditions for a definition, the nature of statistical tests does not allow them to serve as a basis upon which solid research can be built. They are very useful econometric tools but their fragility does not permit them to perform the role of providing the conditions upon which deductive reasoning is based.

At the current state of knowledge it does not seem possible, however, to replace the definition in RG2020 with another of the same type. Therefore, the simple alternative that I propose is a necessary step back, to a less precise, vaguer definition, containing necessary (though not sufficient) conditions that are dispersed through several previous definitions. Hopefully it will serve for some time as a basis for further research on trends in economics which permits a firmer progression. Although the concept of trend certainly requires further clarification, for now measurement must proceed with an imprecise, somewhat vague definition. It appears to be preferable to theory based on apparently rigorous and precise but possibly wrong measurement.

Disclosure of interest

The author report there are no competing interests to declare

References

[1] Belnap, Nuel (1993). On rigorous definitions, *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 72 (2/3), pp. 115-146.

[2] Burge, Tyler (1993). Concepts, definitions, and meaning, *Metaphilosophy*, 24 (4), pp. 309-325.

[3] Canjels, Eugene and Mark W. Watson (1997). Estimating deterministic trends in the presence of serially correlated errors, *The Review of Economics and Statistics*, 79 (2), pp. 184-200.

[4] Gupta, Anil and Stephen Mackereth (2023). Definitions, *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta and Uri Nodelman.

[5] Hamilton, James D. (2018). Why you should never use the Hodrick-Prescott filer, *The Review of Economics and Statistics*, 100 (5), pp. 831-43.

[6] Harvey, Andrew C. (1989). Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press, Cambridge.

[7] Harvey, Andrew C. (2016). Trend analysis, *Wiley StatsRef: Statistics Reference Online*, John Wiley & Sons.

[8] Hodrick, Robert J. and Edward C. Prescott (1981). Post-war U.S. business cycles: an empirical investigation, *Discussion Paper* 451, Southwestern University.

[9] Hodrick, Robert J. and Edward C. Prescott (1997). Post-war U.S. business cycles: an empirical investigation, *Journal of Money, Credit and Banking*, 29 (1), pp. 1-16.

[10] Kiefer, Nicholas M. and Timothy J. Vogelsang (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests, *Econometric Theory*, 21, pp. 1130-1164.

[11] Leser, Conrad E. (1961). A Simple Method of Trend Construction, *Journal of the Royal Statistical Society, Series B*, Vol. 23 (1), pp. 91-107.

[12] Lopes, Artur S. (2025). *An Introduction to Univariate Analysis of Trends in Economic Time Series*, manuscript submitted for publication.

[13] Müller, Ulrich & Watson, Mark W. (2024). Low-Frequency Analysis of Economic Time Series, draft chapter for *Handbook of Econometrics*, Volume 7, edited by S. Durlauf, L.P. Hansen, J.J. Heckman, and R. Matzkin.

[14] Nelson, Charles R. and Charles I. Plosser (1982). Trends and random walks in macroeconomic time series: evidence and implications, *Journal of Monetary Economics*, 10, pp. 139-162.

[15] Persons, Warren M. (1919). Indices of business conditions, *The Review of Economics and Statistics*, vol. 1 (1), pp. 5-107.

[16] Phillips, P. C. B. (2010). The mysteries of trend, *Cowles Foundation Discussion Paper* no. 1771.

[17] Rivas, María Dolores G. and Jesús Gonzalo (2020). Trends in distributional characteristics: existence of global warming, *Journal of Econometrics*, 214, pp. 153–174.

[18] Seater, John J. (1993). World temperature-trend uncertainties and their implications for economic policy, *Journal of Business & Economic Statistics*, 11 (3), pp. 265-277.

[19] Swartz, Norman (2010). *Definitions, Dictionairies, and Meaning*, Simon Fraser University. `wwww.sfu.ca/~swartz/definitions.htm`, accessed 15 February 2025.