Non-Technical Summary for "All-Inside" Software from Casini and Perron (2021) "Change-Point Analysis of Time Series with Evolutionary Spectra"

ALESSANDRO CASINI[1] and PIERRE PERRON[2]

May 2021

Abstract

This note presents a non-technical summary of the methods in Casini and Perron (2021). Section 1 describes the main intuition behind the methods. Section 2 explains how the method can be used for the choice of sub-samples for regression analysis as well as for other type of analyses. Section 3 provides links to the current version and updates.

1 Non-Technical Overview

Casini and Perron (2021) contains new procedures to test the detection of various types of parameter instability in a time series (e.g., mean, variance and autocorrelation) and estimate the break dates.[3] The tests and the estimator are based on frequency domain statistics. This contrasts with most of the work in the structural breaks literature (e.g., change-points, regime-switching, time-varying parameters, etc.). The advantage of working with frequency domain statistics is that the spectral density function or spectrum contains information for the first and second moments of the time series of interest. Therefore, a test statistic or an estimator based on an estimate of the spectral density can simultaneously detect changes in the mean, volatility and autocovariance. This means that a single statistical procedure can be used to detect any change, abrupt or smooth, in the first and second moments. This has a considerable practical advantages over standard time-domain procedures. They require the user to use multiple tests in order to detect the changes of interest. For example, tests for breaks in the mean cannot detect breaks in the variance or autocovariance; and tests for breaks in the variance cannot detect breaks in the mean. In addition, some tests/estimators can detect abrupt breaks but have no power against smooth breaks whereas tests/estimators that work for smooth breaks do not for abrupt breaks. Since the user in general does not know a priori which feature of the time series has changed, one would need to run several tests and use several estimators to learn if any break is present at all. This is practically costly. Finally, sometimes these methods are coded in different software languages which impairs their applicability.

[1]Corresponding author at: Department of Economics and Finance, University of Rome Tor Vergata, Via Columbia 2, Rome, 00133, IT. Email: alessandro.casini@uniroma2.it.

[2]Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, US. Email: perron@bu.edu.

[3]The manuscript including the relevant theory for the methods for changes in the mean is in preparation, see Belotti, Casini, Catania, Grassi, and Perron (2021).

Our tests and estimators are based on an estimate of the spectral density over the frequency domain. We propose a single procedure that can detect any type of break in any feature of a time series.[4] Our procedure can detect any following types of parameter change:

- abrupt breaks;
- smooth changes, i.e., parameter changes that take sometime to take effect.
- changes in the smoothness of the parameter, e.g., the volatility can become rougher (i.e., less smooth; think about constant volatility versus slowly-varying volatility) from one regime to another without signifying a break.

Any of the above three types of change is relevant for economics and finance. They can be detected via changes in the mean, volatility and autocovariance. We now give some intuition for our procedure. An important summary measure of the features of a time series $\{X_t\}_{t=1}^{T}$ is the spectral density function, denoted by $f(t, \omega)$, where $\omega \in [0, \pi]$ is the frequency. In the frequency domain, a time series $\{X_t\}_{t=1}^{T}$ is seen as a combination of different components associated with different frequencies in the range $[0, \pi]$. For example, frequencies close to $\omega = 0$ are low frequencies. They are associated with components in X_t that take a long time to repeat themselves. On the other hand, frequencies close to π are high frequencies. They are associated with components in X_t that take a short time to repeat themselves. The dependence of $f(t, \omega)$ in t means that the properties of the series can change over time. To give an example, consider the spectral density of an AR(1) process

$$X_t = \rho_t X_{t-1} + e_t,$$

where ρ_t may depend on t and e_t is white noise $(0, \sigma^2)$. Then,

$$f(t, \omega) = \frac{\sigma^2}{2\pi} \left|1 - \rho_t \exp(-i\omega)\right|^2, \qquad \omega \in [0, \pi],$$

where $i = \sqrt{-1}$. Assuming smoothness of $f(t, \omega)$ in ω, one can detect changes in ρ_t by studying how the properties of $f(t, \omega)$ change over time. In the paper, we discuss how to consistently estimate $f(t, \omega)$. Intuitively, the estimate of $f(t, \omega)$ weights more observations that are close to t and less observations that are far from t. For example, if one is interested in knowing whether there is a break in ρ_t at time T^* with $1 < T^* < T$, one can estimate $f(t, \omega)$ using only observations prior to t and compare it with an estimate that uses observations after t (i.e., we exploit the fact that $f(t, \omega)$ is a nonlinear smooth function of ρ_t). Thus, a large change between the two estimates is evidence for a change in ρ_t. Note that the same idea can be exploited to detect changes in σ^2 since also σ^2 is summarized in $f(t, \omega)$.

The test statistics $\mathrm{S}_{\max,T}(\omega)$, $\mathrm{S}_{\mathrm{Dmax},T}$, $\mathrm{R}_{\max,T}(\omega)$ and $\mathrm{R}_{\mathrm{Dmax},T}$ in the paper exploit this idea. Suppose we split the sample in blocks each with n_T observations. The index of the block is $r = 1, \ldots, \lfloor T/n_T \rfloor$. Let $\widehat{f}_{r,T}(\omega)$ be the estimate of $f(rn_T, \omega)$. A large deviation between $\widehat{f}_{r,T}(\omega)$ and $\widehat{f}_{r+1,T}(\omega)$ suggests the presence of a break in the spectrum close to time rn_T and at frequency

[4]The name "All-Inside" reflects the feature that one needs to run only a single code to detect any break. That is, it is *"all inside"* one code.

ω. We can take the maximum over $r = 1, \ldots, \lfloor T/n_T \rfloor$ and $\omega \in [0, \pi]$ to ensure the maximal evidence of a parameter change. We refer to the paper for the full treatment. Here we only report a brief presentation of the test statistics. For some given frequency ω the test $\mathrm{S}_{\max,T}(\omega)$ is given by

$$\mathrm{S}_{\max,T}(\omega) = \max_{r=1,\ldots,M_T-2} \left| \frac{\widetilde{f}_{r,T}(\omega) - \widetilde{f}_{r+1,T}(\omega)}{\widehat{\sigma}_{f,r}(\omega)} \right|, \qquad \omega \in [-\pi, \pi], \tag{1.1}$$

where $\widehat{\sigma}_{f,L,r}(\omega)$ is an estimate of the standard deviation of $\widetilde{f}_{r,T}(\omega)$. When the frequency at which the break occurs is not known a priori, we propose the test $\mathrm{S}_{\mathrm{D}\max,T}$ which is based on

$$\max_{\omega_k \in \Pi'} \mathrm{S}_{\max,T}(\omega_k),$$

where $\Pi' \subset \Pi = \{\omega_1, \omega_2, \ldots, , \omega_{n_\omega - 1}, \omega_{n_\omega}\}$ is a set of frequencies over which the maximum is computed.

Another proposed test is $\mathrm{R}_{\max,T}(\omega)$ which is given by

$$\mathrm{R}_{\max,T}(\omega) = \max_{r=1,\ldots,M_T-2} \left| \frac{\widetilde{f}_{r,T}(\omega)}{\widetilde{f}_{r+1,T}(\omega)} - 1 \right|,$$

and is valid for a given frequency ω, while corresponding test which is agnostic about which frequency the break occurs is $\mathrm{R}_{\mathrm{D}\max,T}$ which is based on

$$\max_{\omega_k \in \Pi'} \mathrm{R}_{\max,T}(\omega_k).$$

The difference between $\mathrm{S}_{\max,T}(\omega)$ and $\mathrm{R}_{\max,T}(\omega)$ is that the latter is self-normalized such that one does not need to construct the estimate $\widehat{\sigma}_{f,r}(\omega)$.

The tests $\mathrm{S}_{\max,T}(\omega)$ and $\mathrm{R}_{\max,T}(\omega)$ depend on ω. The choice of ω is, of course, important as it involves different frequency components and hence different periodicities. If the user does not have a priori knowledge about the frequency at which the spectrum has a change-point, our recommendation is to run the tests for multiple values of $\omega \in [0, \pi]$. Even if the change-point occurs at some ω_0 and one selects a value of ω close but not equal to ω_0 the tests should still be able reject the null hypothesis given the differentiability of $f(u, \omega)$. Thus, one can select a few values of ω evenly spread on $[0, \pi]$.

In the paper we propose Algorithm 1 for the estimation of the break date, that is the date at which some change has occurred. The algorithm works as follows. If any of the tests $\mathrm{S}_{\max,T}(\omega)$, $\mathrm{S}_{\mathrm{D}\max,T}$, $\mathrm{R}_{\max,T}(\omega)$ and $\mathrm{R}_{\mathrm{D}\max,T}$ rejects the null hypothesis of no break, then the algorithm starts and searches for the date at which there is the largest deviation between $\widehat{f}_{r,T}(\omega)$ and $\widehat{f}_{r+1,T}(\omega)$. It searches over all time points in the set $\mathcal{T} = \{n_T, 2n_T, \ldots, T - n_T\}$. If the normalized deviation is greater than a certain threshold d_T^*, that time point is deemed as a break date, say $\widehat{T}_1$.

Then v_T observations in a neighborhood of $\widehat{T}_1$ are removed from $\mathcal{T}$, which is updated. The search for a possible second break repeats the above steps. The algorithm terminates when for any index $r \in \mathcal{T}$ the deviation between $\widehat{f}_{r,T}(\omega)$ and $\widehat{f}_{r+1,T}(\omega)$ is smaller than the threshold d_T^*.

2 Procedure for Sub-Sample Choice for Regression

In this section, we explain how to use our method for the choice of sub-samples for regression analysis. Suppose we want to estimate the following linear model by least-squares,

$$y_t = x_t'\beta + e_t, \qquad t = 1, \dots, T. \tag{2.1}$$

However, we suspect that β and/or the second moments of e_t may change over the sample. We can use the code `All-Inside.m` to split the sample in sub-samples such that we are confident that in each sub-sample β is constant and the errors are stationary. We proceed as follows:

1. Estimate (2.1) by least-squares over the full sample. Get the least-squares residuals $\{\widehat{e}_t\}$
2. Take $\{\widehat{e}_t\}$ as the input in Algorithm 1 and obtain the break dates. That is, test for breaks in the mean and second moments of $\{\widehat{e}_t\}$, and determine the relevant sub-samples using the estimated break dates such that the residual are zero-mean and stationary in each sub-sample.
3. Estimate (2.1) by least-squares separately for each sub-sample as determined in step 2.

The key idea is that if there are breaks or more general time variation in β this can be reflected in the properties of the residuals. A break in β results in a break in the mean of $\{\widehat{e}_t\}$ where the latter are the least-squares residuals obtained from a least-squares regression estimated over the full sample. The code `Example_OLS_Subsamples.m` presents a simple example to guide the user. Importantly, in this way our procedure can detect any type of time variation in β or in the variance and autocovariance of e_t (i.e., abrupt changes, smooth changes, etc).

3 Link to Implementation

For more details and updates, visit https://alessandro-casini.github.io/All-Inside/. Also accessible via https://alessandro-casini.com/software/.

References

BELOTTI, F., A. CASINI, L. CATANIA, S. GRASSI, AND P. PERRON (2021): "Testing and Estimating Abrupt and Smooth Breaks in the Mean of a Time Series," *Unpublished Manuscript, Department of Economics and Finance, University of Rome Tor Vergata.*

CASINI, A., AND P. PERRON (2021): "Change-Point Analysis of Time Series with Evolutionary Spectra," *Unpublished Manuscript, Department of Economics and Finance, University of Rome Tor Vergata.*