ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ

ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS

SCHOOL OF INFORMATION SCIENCES & TECHNOLOGY

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

ARDL Bounds test for Cointegration with Application on Income Inequality

By

Kleanthis K. Natsiopoulos

A THESIS
Submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfilment of the requirements for
the degree of Master of Science in Statistics

Athens, Greece
September 2018

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ & ΤΕΧΝΟΛΟΓΙΑΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ

ARDL Bounds test Συνολοκλήρωσης με Εφαρμογή στην Ανισοκατανομή Εισοδήματος

Κλεάνθης Κ. Νατσιόπουλος

ΔΙΑΤΡΙΒΗ

Που υποβλήθηκε στο Τμήμα Στατιστικής

του Οικονομικού Πανεπιστημίου Αθηνών

ως μέρος των απαιτήσεων για την απόκτηση

Μεταπτυχιακού Διπλώματος Ειδίκευσης στη Στατιστική

Αθήνα
Σεπτέμβριος 2018

DEDICATION

In dedication to my family and friends.

ACKNOWLEDGEMENTS

I would like to thank Assistant Professor A. Livada, my supervisor, for her guidance, enthusiasm and patience. I also wish to acknowledge the help provided by Professor S. Dimeli.

VITA

My career path started with my Bachelor of Science in Economics with major in Business Economics where I graduated first in class. Then I started my Master of Science in Statistics the requirements of which I complete with the present thesis while I am working as a data scientist for the last year.

ABSTRACT

Kleanthis Natsiopoulos

ARDL BOUNDS TEST FOR COINTEGRATION WITH APPLICATION ON INCOME INEQUALITY

September 2018

This thesis addresses the popular ARDL bounds test for cointegration and tries to demonstrate its underlying theoretical assumptions in a concise way so that someone interested in this area can easily apply and at the same time understand why and how it works. Often, misuse of a method or a test can lead to a bad situation or unexpectedly unwanted results. Thus, it is also crucial to understand under which circumstances the test doesn't work as expected. For this reason, four practical implementations of the test are presented showing some interesting behavior of the test in practice. This practical section explores the cointegrating relationships between the 1% top income share and the macroeconomic factors of credit, education, gdp, inflation, population growth and trade in order to reveal if there is a long-run relationship. This relationship is tested for four different countries, Greece, France, USA and UK trying to see if the income inequality is driven by the same factors and in the same way for such different economies. For the cases of Greece and France, although there were strong indications supporting the existence of such a relationship, due to a particular limitation in the test's methodology (endogeneity) we couldn't say for sure whether those results were valid or not. In the case of USA, the test concluded for the existence of a long-run relationship but a simple graphical inspection was enough to tell us that this was a false positive alarm (type I error) as this was a degenerate relationship. Finally, in the case of UK a not well defined model was supporting the long-run relationship hypothesis but a more carefully designed model was against this decision.

ΠΕΡΙΛΗΨΗ

Κλεάνθης Νατσιόπουλος

ARDL BOUNDS TEST ΣΥΝΟΛΟΚΛΗΡΩΣΗΣ ΜΕ ΕΦΑΡΜΟΓΉ ΣΤΗΝ ΑΝΙΣΟΚΑΤΑΝΟΜΗ ΕΙΣΟΔΗΜΑΤΟΣ

Σεπτέμβριο 2018

Αυτή η διπλωματική εργασία εξετάζει το δημοφιλές ARDL bounds test συνολοκλήρωσης και προσπαθεί να δείξει τις θεωρητικές παραδοχές πίσω από το τεστ με συνοπτικό τρόπο έτσι ώστε ο κάθε ενδιαφερόμενος να μπορεί εύκολα να εφαρμόσει και ταυτόχρονα να κατανοεί γιατί και πως λειτουργεί. Συχνά, η λανθασμένη χρήση μια μεθοδολογίας ή ενός τεστ μπορεί να οδηγήσει σε άσχημες καταστάσεις ή απρόσμενα και ανεπιθύμητα αποτελέσματα. Συνεπώς, είναι εξίσου σημαντικό να καταλάβουμε υπό ποιες συνθήκες το τεστ δεν λειτουργεί όπως θα περιμέναμε. Γι' αυτό το λόγο παρουσιάζονται τέσσερις πρακτικές εφαρμογές του τεστ δείχνοντας κάποιες ενδιαφέρουσες συμπεριφορές τους τεστ στην πράξη. Το πρακτικό αυτό κομμάτι ερευνά τις σχέσεις συνολοκλήρωσης μεταξύ του ανώτερου 1% μεριδίου εισοδήματος και των μακροοικονομικών παραγόντων πίστωσης, εκπαίδευσης, ΑΕΠ, πληθωρισμού, ρυθμού αύξησης του πληθυσμού και εμπορείου έτσι ώστε να αποκαλύψει αν υπάρχει μακροχρόνια σχέση. Αυτή η σχέση ελέγχεται για τέσσερις χώρες, την Ελλάδα, τη Γαλλία, τις ΗΠΑ και το Ηνωμένο Βασίλειο προσπαθώντας να δει αν η ανισοκατανομή του εισοδήματος οδηγείται από τους ίδιους παράγοντες και με τον ίδιο τρόπο σε τόσο διαφορετικές οικονομίες. Για την περίπτωση της Ελλάδας και της Γαλλίας, αν και υπήρχαν ισχυρές ενδείξεις που υποστήριζαν την ύπαρξη μιας τέτοιας σχέσης, λόγω ενός συγκεκριμένου περιορισμού στη μεθοδολογία του τεστ (ενδογένεια) δεν μπορούμε να είμαστε σίγουροι αν αυτά τα αποτελέσματα είναι έγκυρα ή όχι. Στην περίπτωση των ΗΠΑ, το τεστ υπέδειξε την ύπαρξη μακροχρόνιας σχέσης αλλά μια απλή γραφική απεικόνιση ήταν

αρκετή για να δείξει πως ήταν ψευδώς θετικό (σφάλμα τύπου I) καθώς η σχέση ήταν αποκλίνουσα. Τέλος, στην περίπτωση του Ηνωμένου Βασιλείου ένα κακώς προσδιορισμένο μοντέλο υποστήριζε την υπόθεση της μακροχρόνιας σχέσης αλλά ένα πιο προσεκτικά σχεδιασμένο μοντέλο ήταν αντίθετο με αυτή την απόφαση.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Since a lot of economic data are used in the academic and research field of economics but also on the business side, the need for us to have some valid and easy to apply theory for our analysis is crucial. The issue we are going to discuss probably emerges from economic data and the reason why this kind of data are somehow special is because most of these variables have a non-stationary behavior in terms of having a unit root or in other words being an integrated process of some order greater than zero. Its also true that OLS regression is probably the most frequently used technique because the estimation but also the inference are very easy to be done. The problem that arises when one tries to apply such a regression method using indegrated series is called *spurious regression*. Under this situation one gets misleading results that may lead to bad decisions. Although, there are cases where series that behave like this may end up having a stable and relationship, this phenomenon is called *cointegration*.

There have been almost 40 years from the time when this concept appeared for the first time and still remains an important topic in the econometrics literature. In the present thesis, we focus on a specific cointegration test called the *ARDL bounds test* proposed by M. Hashem Pesaran, Shin, & Smith (2001) and despite the fact that more than 15 years have been passed since then, this test in particular remains one of the most hot topics in the literature of econometrics. This is because of its ease of use, the model is estimated using the OLS method, the test is a classic F (or Wald) test on the joint significance of some parameters and its results are quite straightforward to interpret. Nevertheless, it is based on some extensive theoretical assumptions that need to be satisfied in order for the test results to be valid. The fact that the original paper of M. Hashem Pesaran et al. (2001) is highly technical, some of the underlying but important assumptions were not as clear unless one dives deep in the theoretical construction of the test which is in contrast with the advantages that this test offers for the practitioners.

For this reason, the first of the two targets that this thesis focuses on is explaining the core of the test and the underlying assumptions in a precise but concise way. This is done in the next three chapters discussing some basic time series concepts, the

ARDL model and its connection with the ECM, the multipliers and the dynamics, the multivariate analysis of the same concept, some of the assumptions that arise and the hypothesis test itself.

The second target of this thesis addresses the problem of income inequality and if and how there can be a long-run relationship between the 1% top income share and the macroeconomic factors of credit, education, gdp, inflation, population growth and trade. This relationship, if exists, is not clear if behaves with the same way for different economies like these of Greece, France, USA and UK. We apply the ARDL bounds test for cointegration in each of these cases separately and analyze and interpret the results. We don't stay only on a trivial rejection (or not) of the hypothesis but we focus on the reasons why a decision is made and how the test behaves even in situations where we normally wouldn't apply the test in the first place and we are comparing various results for a better understanding.

Chapter 2

Stationarity and integration

2.1 Describing stationarity

Starting with the description of stationarity, firstly we should separate it into two different definitions. The first one is called *strict stationarity* while the second one is called *weak stationarity*. A definition is:

Strict stationarity definition:

If the joint distributions $P(Y_{t_1} \leq \alpha_1, Y_{t_2} \leq \alpha_2, \ldots, Y_{t_k} \leq \alpha_k)$ are shift-invariant, meaning that they stay unchanged over time, then the stochastic processes of this time series is strictly stationary.

Technically this is equal to:

$$P(Y_{t_1} \leq \alpha_1, Y_{t_2} \leq \alpha_2, \ldots, Y_{t_k} \leq \alpha_k) = P(Y_{t_1+j} \leq \alpha_1, Y_{t_2+j} \leq \alpha_2, \ldots, Y_{t_k+s} \leq \alpha_k)$$
$$\forall t_1, t_2, \ldots, t_k, j \tag{2.1}$$

This means that the joint distribution of the processes depends only on the lags and leads (j) and not on time. Also, all the moments describing the stochastic processes are finite and don't depend on time either.

This definition of stationarity is very strict, as its name indicates, and it's not used in practice. For this reason, another more convenient and easier to be tested definition of stationarity is often used in practice. This is called weak stationarity:

Weak stationarity definition:

If the first two moments, the mean and the autocovariance, of a process exist and they don't depend on time (shift-invariant), then the process is called weak-stationary or covariance-stationary.

Technically this is equal to:

$$
\begin{aligned}
&\bullet \quad E(Y_t) \perp t \text{ (constant)} \\
&\bullet \quad Var(Y_t) = \sigma^2 < \infty \\
&\bullet \quad Cov(Y_t, Y_j) = Cox(Y_{t+r}, Y_{j+r}) \perp t \text{ (depends only on lags)}
\end{aligned}
\tag{2.2}
$$

If a process Y_t is strictly-stationary then the random variables Y_t are identically distributed $\forall t$ and $P(Y_{t_1} \leq \alpha_1, Y_{t_2} \leq \alpha_2) = P(Y_{t_1+j} \leq \alpha_1, Y_{t_2+j} \leq \alpha_2) \quad \forall j$

Additionally, if the process Y_t is strictly-stationary and at the same time $E(Y_t) < \infty$ and $Var(Y_t) < \infty$, then the process is also weakly-stationary. While the inverse does not hold, if a process is weakly-stationary but Gaussian[1] then it is also strictly-stationary.

Nowon, for convenience the term *stationary* will be referring to *covariance-stationary.*

2.2 Describing Integration

Many definitions of integration have been given along the previous years that are more or less similar to each other, depending on the scope of each researcher. The following statement is the formal definition proposed by (Engle & Granger, 1987):

> *A series with no deterministic component which has a stationary, invertible, ARMA representation after differencing d times, is said to be integrated of order d, denoted* $x_t \sim I(d)$

Now, connecting the above definition with the chapter 2 we can conclude that a variable $y_t \sim I(d)$, for $d = 0$ is a covariance-stationary[2] processes while for $d > 0$[3] the variable $(1 - L)^d y_t = \Delta^d y_t$ is a stationary processes.

At this stage, we note that the zero order of integration, namely $I(0)$, is a necessary but not sufficient condition for a variable to be stationary. This means that a variable may be $I(0)$ but at the same time it may not be stationary.

[1] The marginal distributions of the process to be Normal

[2] Note that in this whole sentence the term *stationary* refers to covariance-stationarity

[3] The values that d can take are not limited to integer numbers. For non-integer values of d, they are called fractional difference models.

Chapter 3

Cointegration

In this chapter, we will present the basic idea behind the concept of cointegration.

The concept of cointegration firstly appeared implicitly in the work of Davidson, Hendry, Srba, & Yeo (1978) through the Error Correction Model (ECM). Later, C. W. Granger (1981) firstly developed the theory of cointegration suggesting the term of cointegration and the relationship between the error correction models and cointegration. After that, the concept was formally further developed in detail with the works of C. Granger (1983) and Engle & Granger (1987), while the latest showed the integration of the short-run dynamics with long-run equilibrium (Maddala & Lahiri, 2009).

3.1 Describing and conceptualizing cointegration

The concept of cointegration is easy to understand through the example proposed by Davidson et al. (1978) about the consumption spendings model. They showed that although both consumption and income are non-stationary with a unit root, there is a long-run relationship between them which is a stationary process. This is the ratio between consumption and income which remains constant over time, so the linear stationary process is the log of consumption minus the log of income.

Technically, cointegration is a vector unit root process, say a $(k \times 1)$ vector of time series $\mathbf{y_t}$,[1] where its individual components are I(1) but there is some linear combination of $\mathbf{a'y_t}$ that is a stationary I(0) process, for some nonzero $(k \times 1)$ vector $\mathbf{a}$ which is called cointegrating vector (Hamilton, 1994). In this case, $\mathbf{y_t}$ is said to be cointegrated. In plain words, this means that even if some events may lead to permanent changes in the individual components of $\mathbf{y_t}$, there is some long-run equilibrium relationship

[1]Where each of the k time series $y_{i,t}$ is a $(T \times 1)$ vector

among them, described by $\mathbf{a'y_t}$, which hold them together[2].

$$\mathbf{y_t} = \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{k,t} \end{bmatrix} \tag{3.1}$$

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix} \tag{3.2}$$

$$\mathbf{a'y_t} = \begin{bmatrix} a_1 a_2 \cdots a_k \end{bmatrix} \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{k,t} \end{bmatrix} = a_1 y_{1,t} + a_2 y_{2,t} \cdots a_k y_{k,t} = \sum_{i=1}^{k} a_i y_{i,t} \tag{3.3}$$

The cointegrating vector $\mathbf{a}$ has to be normalized in order for its first element to be unity. Notice that for a stationary $\mathbf{a'y_t}$ process, the cointegrating vector $\mathbf{a}$ is not unique. If b is a nonzero scalar, then $b\mathbf{a}$ is also a cointegrating vector. This non-uniqueness of the cointegrating vector is also presented visually below in Figure 3.2 for a better understanding.

Generally, when there are more than two variables in $\mathbf{y_y}$, then there may be more than one (i.e. $h < k$) linearly independent $(k \times 1)$ cointegrating vectors $(\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_k})$ that result in more than one cointegrated relationships $\mathbf{a_i'y_t}$ that are stationary. This can be described by the $(h \times 1)$ stationary vector $\mathbf{A'y_t}$, where $\mathbf{A'}$ is a $(h \times k)$ matrix.

$$\mathbf{A'} = \begin{bmatrix} \mathbf{a_1'} \\ \mathbf{a_2'} \\ \vdots \\ \mathbf{a_h'} \end{bmatrix} \tag{3.4}$$

Hamilton (1994) shows in the triangular representation of a cointegrated system that in the case where $h = k$, in which the number of cointegrating vectors is equal to the number of variables, then $\mathbf{y_t}$ would be $I(0)$.

To sum up, when two variables are cointegrated we say that there is a long-run equilibrium relationship between them so that they don't drift too far apart over time. On the other hand, when the error term of their estimated long-run relationship is $I(1)$ the two series drift apart as time goes on and hence the estimated relationship doesn't really exist (Maddala & Lahiri, 2009).

Although it is very important for one to understand the mathematical representation of cointegration as long as the practical (e.g. the economic) meaning of cointegration

[2]Notice that despite the fact that the multiplications of these two vectors of dimensions $(1 \times k)$ and $(k \times 1)$ respectively results in a (1×1) vector, this vector itself contains the $(T \times 1)$ stationary time series as described in the Equation (3.3)

as both of theme described above, it would be very interested to see them in parallel. For this reasons, we will set a simple yet very informative example that is also used in many textbooks (Hamilton, 1994).

Consider the following bivariate system:

$$y_{1,t} = \gamma y_{2,t} + u_{1,t} \tag{3.5}$$

$$y_{2,t} = y_{2,t-1} + u_{2,t} \tag{3.6}$$

Where $u_{1,t}$ and $u_{2,t}$ are white noise (*WN*) processes, uncorrelated with each other, while $y_{2,t}$ is a random walk (*RW*) processes. Notice that $y_{1,t}$ and $y_{2,t}$ are both individually non-stationary $I(1)$ processes as $y_{2,t}$ is a *RW* and its first difference results in the $u_{2,t} \sim WN$ which is stationary. And the first difference of $y_{1,t}$ is a stationary $MA(1)$. The derivation of the above first differences are presented below.

$$\Delta y_{2,t} = u_{2,t} \tag{3.7}$$

$$\Delta y_{1,t} = \gamma \Delta y_{2,t} + \Delta y_{1,t} = \gamma u_{2,t} + u_{1,t} - u_{1,t-1} = \nu_t + \theta \nu_{t-1} \tag{3.8}$$

Where $\nu_t \sim WN$ and $\theta \neq -1$ given that γ is a nonzero scalar and $u_{2,t}$ is not a mass point at zero[3]. The proof along with a very clear instructive example of (3.8) is presented in Hamilton (1994) *pp.102-105*, where he shows how the sum of a $MA(1)$ and a *WN* uncorrelated processes produces another $MA(1)$ process.

Although $y_{1,t}$ and $y_{2,t}$ are both I(1), their linear combination $\mathbf{a}'\mathbf{y_t} = y_{1,t} - \gamma y_{2,t}$ is stationary as $\mathbf{y_t} = (y_{1,t}, y_{2,t})'$ is cointegrated. Here the cointegrating vector is $\mathbf{a}' = (1, -\gamma)$.

In the following example, we simulate the system described in the Equations (3.5) and (3.6). We generate the two time-series using a sample size of 5000 innovations with the first one to be equal to zero, where $u_{1,t}$ and $u_{2,t}$ are distributed as $N(0, 1)$ and setting the parameter $\gamma = 0.6$. In the Figure 3.1 we can see that the series $y_{1,t}$ and $y_{2,t}$ are obviously non-stationary. As we can notice the effect of the unit autoregressive roots lead to permanent changes (stochastic trends) in the variables as they diverge significantly from zero. Nevertheless, we can see that their linear combination using the real value of $\gamma = 0.6$ (which forms the cointegrating vector $\mathbf{a}' = (1, -0.6)$) is indeed stationary as it fluctuates steadily around zero.

[3] $E(u_{2,t}^2) > 0$

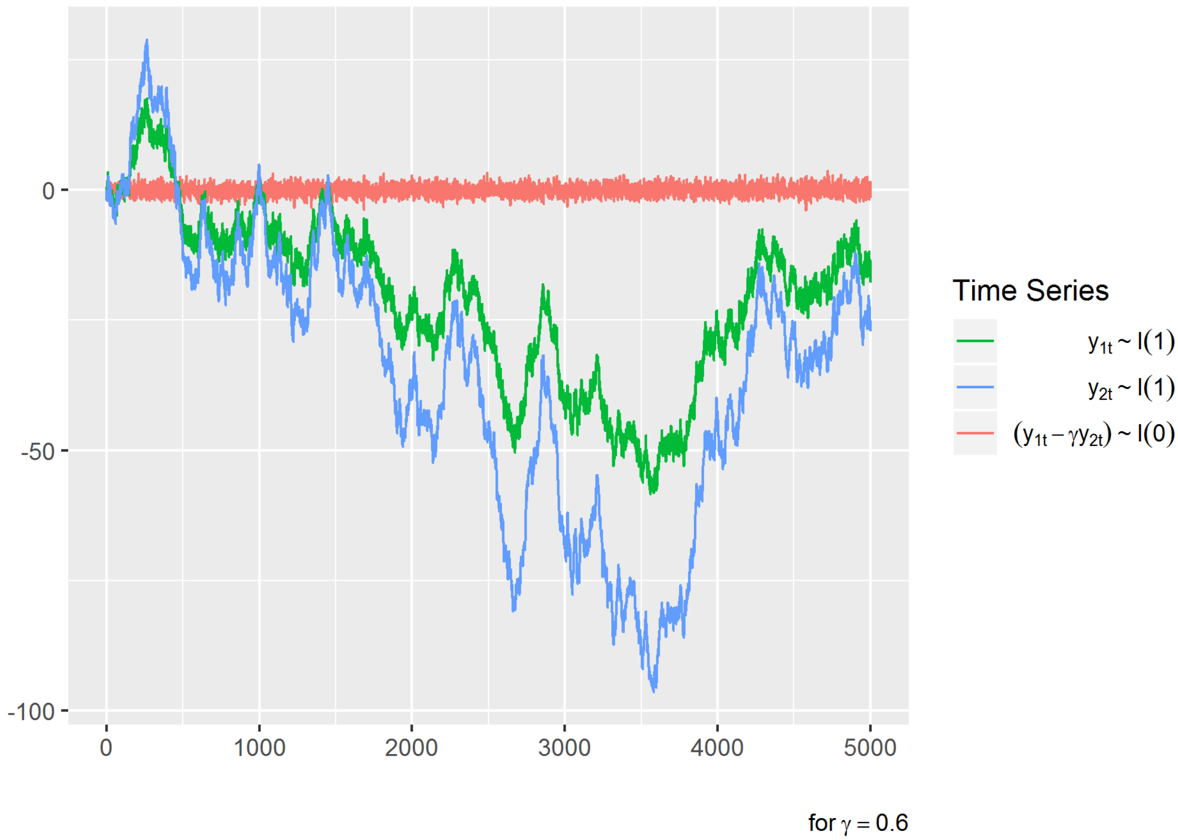

Figure 3.1: Cointegration simulation

In the Figure 3.2 we can see how there can not be more than one cointegrating vectors that are linearly independent of each other for a specific relationship and when it diverges from this unique value, the resulted series is not stationary. This is the case with $\gamma_1 = 0.4$ and $\gamma_2 = 0.8$ where they are minus and plus 0.2 respectively from the unique values of 0.6. Notice that γ is the second element of the cointegrating vector, while the first one is equal to unity (as the whole vector has to be normalized like that). Also, we show here that the cointegrating vector is not unique, in the sense that there can be a scalar b where $b\mathbf{a}$ is also cointegrating vector but they are obviously linearly dependent of each other. In our simulation example, this values is $b = 2.7$. Looking at the figure we can see that the resulted series is stationary as it fluctuates around zero but with a greater variation this time.

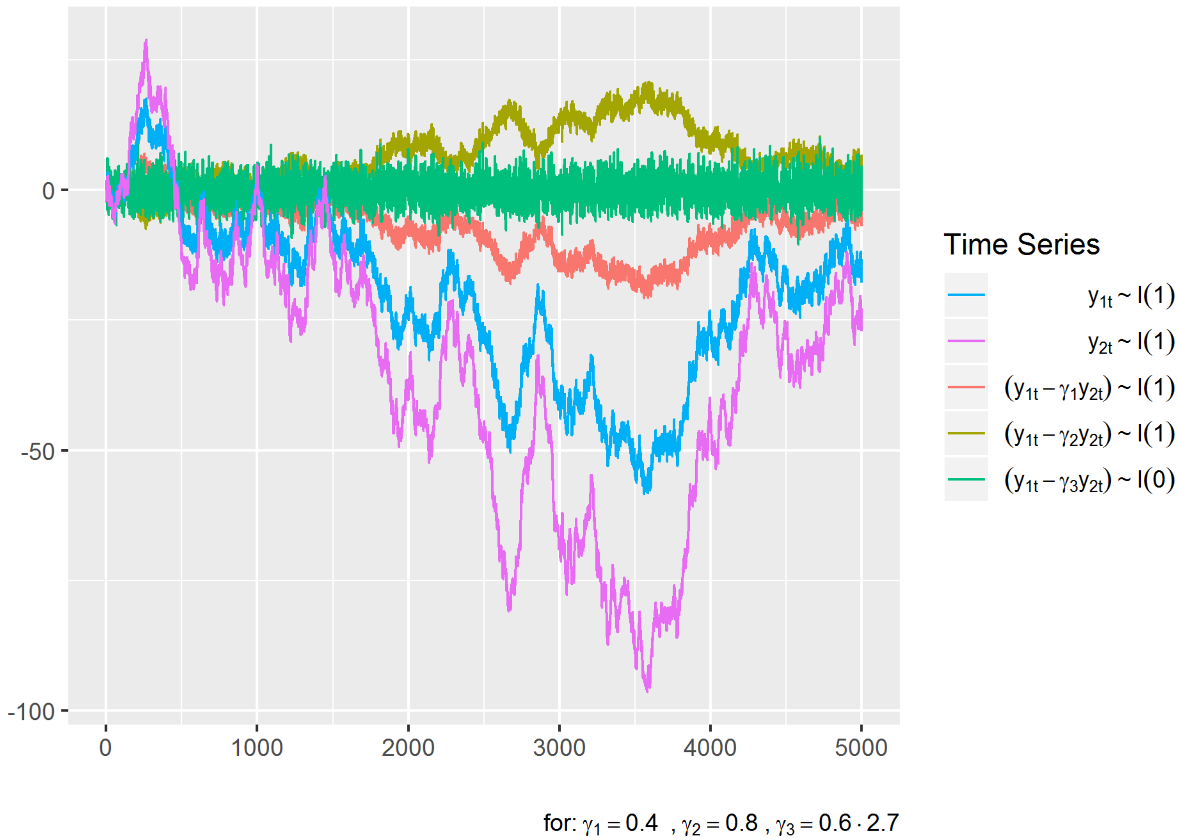

Figure 3.2: Non-cointegrated & non-unique cointegrating vectors simulation

Chapter 4

ARDL Bounds test

The ARDL bounds-test for cointegration was developed through the work of M. H. Pesaran & Shin (1998) and M. Hashem Pesaran et al. (2001). The test is practically performing a significance test on the parameters of the long-run variables included in the Conditional Error Correction Model (CECM) of the underlying VAR model. We explain in details the Conditional ECM latter in Chapter 4.1.3. In the cointegration analysis we use very often the ARDL platform and this is due to its ability to isolate and estimate the long-run relationship among the variables.

As the authors argue, one of the advantages of this model is that using the ARDL model framework, we can have robust estimations of cointegration under possible misspecification of the order of integration of the included variables.

To illustrate that, we consider the following possible cases:

- When all variables are $I(d)$ for $0 \leq d$ and are not cointegrated. Then for the case that $d = 0$ we can estimate Eq. (4.2) in levels using OLS. For the rest of the cases we can do the same after taking appropriate differences in order to end up with $I(0)$ variables. The initial variables may be differenced but we can consider that the estimated ARDL model is in the levels of the new variables namely $(1 - L)^d y_t \equiv \Delta_d y_t$.
- When all variables are $I(1)$ so they are cointegrated. We can estimate the long-run relationship using a simple OLS in levels and we can also estimate the short-run dynamics and the speed of adjustment to the cointegrating relationship constructing an Error Correction Model (ECM).
- When we have a mix of $I(1)$ and $I(0)$ variables and some of the $I(1)$ variables are cointegrated. And here is where the ARDL Bound-test takes over.

The traditional cointegration tests such as Engle & Granger (1987), Phillips & Ouliaris (1990), Johansen (1995) etc. are not able to handle the last case where the order of integration between the variables differs as these tests consider only the cases where all the variables are integrated of the same order (i.e. $I(1)$). Not only the ARDL bounds-test can handle the cases where there is a mix of $I(0)$ and $I(1)$ variables but it also eliminates the cases where a test that requires all variables to be $I(1)$ is applied while some variables where mistakenly estimated as $I(1)$ but their true nature is $I(0)$. In these cases, applying a tests like this would be invalid.

4.1 ARDL model, estimation and inference

In this section, we present the framework on the basis of which the ARDL bounds test is built on along with the corresponding practical models. We present both the theoretical Data Generating Process (DGP) and the regression model which we apply in practice.

4.1.1 General ARDL model

First, we describe the DGP of the general ARDL$(p, q_1, \dots, q_k)$ model:

$$\psi(L)y_t = \alpha_0 + \alpha_1 t + \sum_{j=1}^{k} \beta_j(L)x_{j,t} + \varepsilon_t \tag{4.1}$$

In this form, we use two lag polynomials, the AR operator $\psi(L)$ and the MA operator $\beta_j(L)$ that can be found in Eq. (B.1) and Eq. (B.9) respectively. L is the lag operator, α_0 is the constant term, α_1 is the coefficient of the linear trend, t is the vector representing the linear trend and ε_t is the innovations. This is the theoretical framework while in practice we estimate it using the following regression model[1]:

$$y_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^{p} \psi_i y_{t-i} + \sum_{j=1}^{k} \sum_{l_j=0}^{q_j} \beta_{j,l_j} x_{j,t-l_j} + \varepsilon_t \tag{4.2}$$

Where α_0 is the constant term, α_1 is the coefficient of the linear trend, ψ_i is the coefficient of the i^{th} lag of y_t, β_{j,l_j} is the coefficient of the l_j^{th} lag of the $x_{j,t}$ independent variable and ε_t is the innovations. The open form of Eq. (4.2) is presented in Eq. (B.11). Thus, the general ARDL$(p, q_1, \dots, q_k)$ model except for the possibly existing deterministic components (the constant and the trend), it also contains all the p lags of the variable y_t, all the q_j lags for each of the k variables $x_{j,t}$ and the each of the k variable $x_{j,t}$ in levels.

Another representation of an ARDL$(p, q_1, \dots, q_k)$ model Eq.(4.1) can also be written as a function of the intertemporal dynamics. We derive to this equation applying the Beveridge-Nelson decomposition for a MA process Eq. (B.14) on Eq. (4.1). Here is the theoretical DGP of this representation:

[1] solving for y_t

4.1.2 Intertemporal ARDL model

$$\begin{aligned} y_t &= \alpha_0 + \alpha_1 t + \sum_{i=1}^{p} \psi_i y_{t-i} + \sum_{j=1}^{k} \beta_j(L) x_{j,t} + \varepsilon_t \\ &= \alpha_0 + \alpha_1 t + \sum_{i=1}^{p} \psi_i y_{t-i} + \sum_{i=1}^{p} \Big(\beta_j(1) + (1-L)\tilde{\beta}_j(L)\Big) x_{j,t} + \varepsilon t \\ &= \alpha_0 + \alpha_1 t + \sum_{i=1}^{p} \psi_i y_{t-i} + \sum_{i=1}^{p} \beta_j(1) x_{j,t} + \sum_{j=1}^{k} \tilde{\beta}_j(L) \Delta x_{j,t} + \varepsilon_t \end{aligned} \tag{4.3}$$

In particular, here we use the Eq. (B.14) as $\beta_j(L) = \beta_j(1) + (1-L)\tilde{\beta}_j(L)$ to account for each of the k regressors. The $\beta_j(1)$ described in Eq. (B.10) also uses the subscript j to account for every variable and it it actually the sum of the level's and all the lags' coefficients.

The regression model that we use in practice in order to estimate the Eq. (4.3) is:

$$y_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^{p} b_{0,i} y_{t-i} + \sum_{j=1}^{k} b_j x_{j,t} + \sum_{j=1}^{k} \sum_{l_j=0}^{q_j-1} c_{j,l_j} \Delta x_{j,t-l_j} + \varepsilon_t \tag{4.4}$$

Where this representation of the ARDL model consists of the possibly existing deterministic components (the constant and the trend), all the p lags of the variable y_t, each of the k variables $x_{j,t}$ in levels, the first differences of each of the k regressors $(\Delta x_{j,t})$ and the $q_j - 1$ lags for each of the k variables $\Delta x_{j,t}$.

4.1.3 Conditional Error Correction Model (CECM)

An ARDL model is equivalent to the Conditional Error Correction Model (CECM) of the underlying VAR as there is an 1-1 correspondence between them (Banerjee, 1993). As they carry the same information (for example regarding the estimation of the long-run multipliers), in the empirical part, Chapter 5, we present the ARDL form of the models but we also provide their conditional ECM form in the Appendix C.

As the Conditional ECM is just another representation of the ARDL model in Eq. (4.1) we can show that it derives from the general ARDL model following the next steps starting from the know relationship $\Delta y_t = y_t - y_{t-1}$. First replace y_t according to the Eq. (4.1). Then according to Eq. (B.4) we can replace $\sum_{i=1}^{p} \psi_i y_{t-i}$ with $\psi^*(L) y_t$ and then apply the Eq. (B.13) which is a reparameterization of the Beveridge-Nelson decomposition for an AR process. Gather the y_{t-1} terms, express the $x_{j,t}$ using the relationship $x_t = x_{t-1} + \Delta x_t$ and apply the Beveridge-Nelson decomposition for a MA process to its operator. Finally, notice that the coefficient of y_{t-1} is actually the $\psi(1)$ according to the Eq. (B.2).

Now rearranging the right hand parts of this equation we derive the following Eq. (4.5) which is in fact the Conditional ECM. In this form it is an Unrestricted CECM

(UCECM) as we allow for the parameters of the once lagged variables in levels (y_{t-1} and $x_{j,t-1}$) to be estimated. The Eq. (4.5) shows the theoretical framework or the DGP of the UCECM.

Unrestricted CECM

$$\begin{aligned}\Delta y_t = {} & \alpha_0 + \alpha_1 t \\ & - \psi(1) y_{t-1} + \sum_{j=1}^{k} \beta_j(1) x_{j,t-1} \\ & + \widetilde{\psi}^*(L) \Delta y_{t-1} + \sum_{j=1}^{k} \widetilde{\beta}_j(L) \Delta x_{j,t-1} + \sum_{j=1}^{k} \beta_j(L) \Delta x_{j,t} + \varepsilon_t\end{aligned} \tag{4.5}$$

The model as we would estimate it in practice is shown in the Eq. (4.6).

$$\begin{aligned}\Delta y_t = {} & \alpha_0 + \alpha_1 t \\ & + b_0 y_{t-1} + \sum_{j=1}^{k} b_j x_{j,t-1} \\ & + \sum_{i=1}^{p-1} c_{0,i} \Delta y_{t-i} + \sum_{j=1}^{k} \sum_{l_j=1}^{q_j-1} c_{j,l_j} \Delta x_{j,t-l_j} + \sum_{j=1}^{k} d_j \Delta x_{j,t} + \varepsilon_t\end{aligned} \tag{4.6}$$

Now that we know how to form the Unrestricted form of the CECM it is very easy to transform in into a Restricted CECM (RCECM) which does not include the terms y_{t-1} and $x_{j,t-1}$ themselves but the (possibly cointegrating) relationship between them. This can be done by slightly changing the Eq. (4.5) and setting $\psi(1)$ to be the common multiplier for the relationship formed by y_{t-1} and $x_{j,t-1}$. The theoretical DGP of the RCECM is in Eq. (4.7).

Restricted CECM

$$\begin{aligned}\Delta y_t = {} & \alpha_0 + \alpha_1 t \\ & - \psi(1)(y_{t-1} - \sum_{j=1}^{k} \frac{\beta_j(1)}{\psi(1)} x_{j,t-1}) \\ & + \widetilde{\psi}^*(L) \Delta y_{t-1} + \sum_{j=1}^{k} \widetilde{\beta}_j(L) \Delta x_{j,t-1} + \sum_{j=1}^{k} \beta_j(L) \Delta x_{j,t} + \varepsilon_t\end{aligned} \tag{4.7}$$

The model as we would estimate it in practice is shown in the Eq. (4.8).

$$\begin{aligned}\Delta y_t = {} & \alpha_0 + \alpha_1 t \\ & + b_0 ECT_{t-1} \\ & + \sum_{i=1}^{p-1} c_{0,i} \Delta y_{t-i} + \sum_{j=1}^{k} \sum_{l_j=1}^{q_j-1} c_{j,l_j} \Delta x_{j,t-l_j} + \sum_{j=1}^{k} d_j \Delta x_{j,t} + \varepsilon_t\end{aligned} \tag{4.8}$$

It is clear that the Error Correction Term (ECT_t) is the relationship in levels between y_t and $x_{j,t}$ because if a relationship exists in time t, it also exists in time t-1. So, if this relationship indeed exists, it is called cointegrating relationship and it is actually the once lagged errors from the estimated relationship between y_t and $x_{j,t}$.

Looking at the Eq. (4.7) we notice that the coefficients that accompanies the variables $x_{j,t-1}$ are in fact the long-run multipliers[2] and that the coefficient of the ECT_{t-1} term is the $\psi(1)$.

The RCECM model is a very interesting one and it's often the first model we estimate since we have favorable results from a cointegration test. This is because the coefficient of the ECT_{t-1} term is another way to further support (or reject) our conclusions about the existence of cointegration. If the ECT_{t-1} term is statistically significant, this means that the cointegrating relationship (which it represents) exists. The coefficient, which appears as b_0 in the Eq. (4.8) and has to be negative in sign, has also a very interesting interpretation. It shows the speed of adjustment back to the long-run equilibrium. Its absolute value $|b_0|$ can be interpreted as the percentage of which the divergence from the equilibrium is reduced in each time unit (e.g. each year) and the ratio $\frac{1}{|b_0|}$ indicates the time (measured in time units) that the system takes to get back in equilibrium.

Another interesting relationship among all of the previously referred models is the one about $\psi(1)$. In the UCECM Eq. (4.6) and the RCECM (4.8) this value expresses itself through a single coefficient.

$$\psi(1) = -b_0 \tag{4.9}$$

In the general ARDL Eq. (4.2) and the Intertemporal ARDL Eq. (4.4) appears as the sum of the coefficients of the lagged dependent variables.

$$\psi(1) = \sum_{i=1}^{p} b_{0,t-i} \tag{4.10}$$

4.2 Long-run dynamics

Exploring the connection between the general form of the ARDL model (4.2) and the ARDL model including the intertemporal dynamics Eq. (4.4) we realize that they are just two different representations of the same thing. In fact, we can derive the coefficients of the general ARDL representation from the estimated coefficients of the intertemporal dynamics representations and vice versa. Opening the Eq. (4.4) replacing the $\Delta x_{j,t-l_j}$ with $L^{l_j}(x_{j,t} - x_{j,t-1})$ one can notice that the corresponding coefficients of the levels and all the lags for each of the k regressors $(x_{j,t}, x_{j,t-1}, \ldots, x_{j,t-(q_j-1)}, x_{j,t-qj})$

[2] In the RCECM form these are not estimated here but instead they are pre-estimated and they are used along with the y_t and $x_{j,t}$ variables to form the ECT_t term which is used here with one lag.

are accordingly:

$$\begin{aligned}
&\beta_{j,0} = b_j + c_{j,0} && \text{coefficient of } x_{j,t} \\
&\beta_{j,1} = c_{j,2} + c_{j,1} && \text{coefficient of } x_{j,t-1} \\
&\beta_{j,2} = c_{j,3} + c_{j,2} && \text{coefficient of } x_{j,t-2} \\
&\vdots \\
&\beta_{j,q_j-1} = c_{j,q_j-1} + c_{j,q_j-2} && \text{coefficient of } x_{j,t-(q_j-1)} \\
&\beta_{j,q_j} = -c_{j,q_j} && \text{coefficient of } x_{j,t-q_j}
\end{aligned} \tag{4.11}$$

The Eq. (4.3) is also very useful if we want to derive the long-run relationship between y_t and the k regressors $x_{j,t}$, without using any information of the lag of y_t. An *ARDL(p,q)* model can be represented using iterative substitution as an infinite distributed lag model, from which we can understand how a shock in a variable affects future periods. The following *ARDL(1,1)* (4.12) for example, can be written as a distributed lag (DL) model as presented in Eq. (4.13):

$$y_t = a_0 + a_1 t + \psi_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t \tag{4.12}$$

$$y_t = (1 + \psi_1 + \psi_1^2 + \dots)a_0 + (1 + \psi_1 L + \psi_1^2 L^2 + \dots)(a_1 t + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t) \tag{4.13}$$

An infinite Distributed Lag model can be written as:

$$y_t = a_0 + a_1 t + \sum_{j=1}^{k} \sum_{l=0}^{\infty} \beta_{j,l} x_{j,t-l} + \varepsilon_t \tag{4.14}$$

And under a finite structure this becomes a $DL(q_1, \dots, q_k)$:

$$y_t = a_0 + a_1 t + \sum_{j=1}^{k} \sum_{l_j=0}^{q_j} \beta_{j,l_j} x_{j,t-l_j} + \varepsilon_t \tag{4.15}$$

With another representation of this to be[3]:

$$y_t = a_0 + a_1 t + \sum_{j=1}^{k} \theta_j x_{j,t} + \sum_{j=1}^{k} \sum_{l_j=0}^{q_j-1} \gamma_{j,l_j} \Delta x_{j,t-l_j} + \xi_t \tag{4.16}$$

The coefficients α_1 and θ_j are the long-run multipliers of the trend and the independent variables x_j respectively. These multipliers measure the total effect on the dependent variable after a unit change in the dependent ones, and their estimation is one of our main goals if this cointegrating (long-run) relationship between the variables y_t and x_j (in levels) indeed exists.

[3]As we mentioned before, regressing on the levels of $x_{j,t}$ and all their q_j lags is equivalent to regressing on the levels of $x_{j,t}$, the levels of $\Delta x_{j,t}$ and their $q_j - 1$ lags.

4.3 Multipliers

Since we estimate the ARDL model of either form, the interpretation of the dynamic effects can be done using the so called multipliers. The long-run multiplier is often of big interest and thus we give special attention forming an example in which we replicate and validate the results from the Chapter 5.4.

4.3.1 Formulas for multipliers

short-run **or** ***impact multiplier***

- The effect on y_t from a unit change in x_t:

With respect to the DL Eq. (4.14) or (4.15):

$$\frac{\partial y_t}{\partial x_t} = \beta_0 \tag{4.17}$$

interim multiplier

- The effect on y_{t+s} (s steps ahead) from a unit change in x_t:

With respect to the general ARDL Eq. (4.1):

$$\frac{\partial y_{t+s}}{\partial x_t} \tag{4.18}$$

With respect to the DL Eq. (4.14) or (4.15):

$$\sum_{l=0}^{s} \beta_l \tag{4.19}$$

In particular, the effect on y_{t+1} from a unit change in x_t with respect to the general ARDL Eq. (4.1)[4] and the DL Eq. (4.15) respectively:

$$\frac{\partial y_{t+1}}{\partial x_t} = \beta_1 + \psi_1 \beta_0 \tag{4.20}$$

$$\sum_{l=0}^{1} \beta_l \tag{4.21}$$

[4] see the example in Eq. (4.13) to understand the structure in an open form

long-run **or** ***total multiplier***

The long-run multipliers appear in the Eq. (4.16) as θ_j.

- The total effect on y from a unit change in x_t:

With respect to the DL Eq. (4.14) and (4.15) respectively:

$$\theta = \sum_{l=0}^{\infty} \beta_l \tag{4.22}$$

$$\theta = \sum_{l=0}^{q} \beta_l \tag{4.23}$$

With respect to the general ARDL Eq. (4.2):

$$\theta = \frac{\sum_{l=0}^{q} \beta_l}{1 - \psi_1}$$

Generally, if the autoregressive order is p: (4.24)

$$\theta = \frac{\sum_{l=0}^{q} \beta_l}{1 - \sum_{i=1}^{p} \psi_i}$$

With respect to the Intertemporal ARDL model, Eq. (4.4):

$$\hat{a}_1 = \frac{\hat{\alpha}_1}{1 - \sum_{i=1}^{p} \hat{b}_{0,i}} \qquad \text{long-run multiplier of trend} \tag{4.25}$$

$$\hat{\theta}_j = \frac{\hat{b}_j}{1 - \sum_{i=1}^{p} \hat{b}_{0,i}} \qquad \text{long-run multiplier of } x_{j,t} \tag{4.26}$$

With respect to the UCECM, Eq. (4.5) and (4.6):

$$\hat{a}_1 = \frac{\hat{\alpha}_1}{\hat{b}_0} \qquad \text{long-run multiplier of trend} \tag{4.27}$$

$$\hat{\theta}_j = \frac{\hat{b}_j}{\hat{b}_0} \qquad \text{long-run multiplier of } x_{j,t} \tag{4.28}$$

Notice that the denominator of the above formulas is actually the $\psi(1)$. Now we can see that the long-run parameters are $\hat{\theta}_j = \psi^{-1}(1)\beta_j(1)$. In other words, it is equal to the sum of the coefficients of the levels and all their q_j lags for each of the k regressors x_j (i.e. on a DL model).

Concerning the Eq. (4.1), a notable point here is that $\psi(L)$ must be invertible in order for a stable relationship between y_t and x_t to exist, and hence the long-run multiplier to make sense.

4.3.2 Long-run multiplier example

We can give an example on how to calculate the long-run multipliers using the results of one of the models in the empirical part. For example we can use the estimated model for the case of France which is an $ARDL(1,0,1,1,0,1,0)$. The long-run multipliers can be found in the Table 5.15. In our example we will recalculate the long-run multiplier for the variable `edu` for which we can see from the Table 5.15 that its estimation is 1.294[5].

We can re-estimate this number using the Table 5.13 which is the implementation of the Eq. (4.2), that is the $ARDL(1,0,1,1,0,1,0)$ model in the general ARDL form. Using the formula in Eq. (4.24) and the estimated coefficients from the Table 5.13 we can see that:

$$\frac{(-1.134 + 1.598)}{1 - 0.641} = 1.292$$

We can also re-estimate it using the results from the Unrestricted Conditional ECM model in the Table C.8 which is the implementation of the UCECM model in Eq. (4.6) and plug it in the formula in Eq. (4.28). Thus we have:

$$\frac{0.464}{-(-0.358)} = 1.296$$

4.4 Multivariate analysis

What we have discussed so far concerns one particular model of the underlying VAR, the so called conditional, in which the variable we are primarily interested in is the variable y_t and the independent variables in this model are the so called marginal variables $\mathbf{x}_t = (x_{1,t}, \ldots, x_{k,t})'$.

The underlying VAR model can be described as:

$$\boldsymbol{\Phi}(L)(\mathbf{Z}_t - \boldsymbol{\mu} - \boldsymbol{\gamma}_t) = \boldsymbol{\epsilon}_t \tag{4.29}$$

Where $\mathbf{Z}_t = (y_t, x_{1,t}, \ldots, xk, t)'$ is the $(k+1)$ superset of y_t and $\mathbf{x}_t$, $\boldsymbol{\mu}$ is the $(k+1)$ intercept vector, $\boldsymbol{\gamma}_t$ is the $(k+1)$ trend vector, $\boldsymbol{\Phi}(L)$ is the square matrix lag polynomial Eq. (B.7) and $\boldsymbol{\epsilon}_t$ is the $(k+1)$ vector of innovations.

Using a model of those that was mentioned previously (e.g. the UCECM model), we treat it as a stand alone model while the true underlying VAR probably contains extra information about the interrelationships among the variables for which we don't account following this single model. Due to this fact, we describe here some of the previous concepts as a parallel system instead of a stand alone equation that helps us to dig a little further in technical details.

[5]small differences from the estimated long-run multipliers are due to rounding errors

4.4.1 Vector Error Correction Model (VECM)

Rewriting the Unrestricted Conditional ECM Eq. (4.6) under the multivariate representation we can write the Unrestricted VECM (UVECM) here.

Unrestricted VECM

$$\Delta \mathbf{z}_t = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 t - \boldsymbol{\Phi}(1)\mathbf{z}_{t-1} + \widetilde{\boldsymbol{\Phi}}^*(L)\Delta \mathbf{z}_t + \boldsymbol{\epsilon}_t \tag{4.30}$$

Where $\widetilde{\boldsymbol{\Phi}}^*(L)$ is the analogous matrix to Eq. (B.6), the $\boldsymbol{\Phi}(1)$ is the cointegration matrix, $\boldsymbol{\Phi}(1)\mathbf{z}_{t-1}$ is the matrix containing the cointegrating relationships analogous to $\psi(1)ECT_{t-1}$ and $\widetilde{\boldsymbol{\Phi}}^*(L)\Delta \mathbf{z}_t$ are the short-run dynamics.

or in another form:

$$\begin{aligned}\begin{bmatrix}\Delta y_t \\ \Delta \mathbf{x}_t\end{bmatrix} = \begin{bmatrix}\alpha_{y0} \\ \boldsymbol{\alpha}_{x0}\end{bmatrix} + \begin{bmatrix}\alpha_{y1} \\ \boldsymbol{\alpha}_{x1}\end{bmatrix} t \\ - \begin{bmatrix}\Phi_{yy}(1) & \boldsymbol{\Phi}_{yx}(1) \\ \boldsymbol{\Phi}_{xy}(1) & \boldsymbol{\Phi}_{xx}(1)\end{bmatrix}\begin{bmatrix}y_{t-1} \\ \mathbf{x}_{t-1}\end{bmatrix} \\ + \begin{bmatrix}\widetilde{\Phi}^*_{yy}(L) & \widetilde{\boldsymbol{\Phi}}^*_{yx}(L) \\ \widetilde{\boldsymbol{\Phi}}^*_{xy}(L) & \widetilde{\boldsymbol{\Phi}}^*_{xx}(L)\end{bmatrix}\begin{bmatrix}\Delta y_y \\ \Delta \mathbf{x}_t\end{bmatrix} + \begin{bmatrix}\epsilon_{yt} \\ \boldsymbol{\epsilon}_{xt}\end{bmatrix}\end{aligned} \tag{4.31}$$

4.4.2 Endogeneity

Under the context of the univariate analysis (e.g. the ECM Eq. (4.6)) all the marginal variables are treated as exogenous. Nonetheless, in the multivariate analysis of the VAR Eq. (4.29) there may be several endogenous variables and these endogenous variables are correlated to each other in the VECM Eq. (4.30) system too.

The error vector in the Unrestricted VECM Eq. (4.30) is $\boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Omega})$ where:

$$\boldsymbol{\Omega} = \begin{bmatrix}\omega_{yy} & \boldsymbol{\omega}_{yx} \\ \boldsymbol{\omega}_{xy} & \boldsymbol{\omega}_{xx}\end{bmatrix} \tag{4.32}$$

Where $\boldsymbol{\Omega}$ is the covariance matrix which carries and transfers the correlations among variables between each marginal equation. So, in order for the univariate analysis using the ECM to be valid, it requires that the effects on y_t emerging from this single model to be the same as those which would have passed to y_t from the whole VECM system.

This can happen only in the case when the only effect on y_t under the multivariate analysis is the direct effects from the marginal variables through the conditional equation. If the conditional variable y_t (the variable we are focusing on) also participates in other cointegrating relationship, the effect which naturally would have passed through the covariance matrix $\boldsymbol{\Omega}$ of the error $\boldsymbol{\epsilon}_t$, will be omitted in the univariate case. In this case any inference on the ECM and the tests based on this would be invalid.

Taking a closer look at Eq. (4.31) we can see that the error from the conditional equation (the first row) can be analyzed as follows.

$$\epsilon_{yt} = \boldsymbol{\omega}_{yx}\boldsymbol{\omega}_{xx}^{-1}\boldsymbol{\epsilon}_{xt} + u_{yt} \tag{4.33}$$

Where the new error $u_{yt} \sim N(0, \omega_{yy} - \boldsymbol{\omega}_{yx}\boldsymbol{\omega}_{xx}^{-1}\boldsymbol{\omega}_{xy})$ is independent from $\boldsymbol{\epsilon}_{xt}$.

At this point we rewrite the VECM Eq. (4.31) forming the conditional equation such that the error term to be the new independent error.

$$\begin{aligned}\begin{bmatrix}\Delta y_t \\ \Delta \mathbf{x}_t\end{bmatrix} &= \begin{bmatrix}\alpha_{y0} \\ \boldsymbol{\alpha}_{x0}\end{bmatrix} + \begin{bmatrix}\alpha_{y1} \\ \boldsymbol{\alpha}_{x1}\end{bmatrix} t \\ &- \begin{bmatrix}\Phi_{yy}(1) - \boldsymbol{\omega}_{yx}\boldsymbol{\omega}_{xx}^{-1}\boldsymbol{\Phi}_{xy}(1) & \boldsymbol{\Phi}_{yx}(1) - \boldsymbol{\omega}_{yx}\boldsymbol{\omega}_{xx}^{-1}\boldsymbol{\Phi}_{xx}(1) \\ \boldsymbol{\Phi}_{xy}(1) & \boldsymbol{\Phi}_{xx}(1)\end{bmatrix}\begin{bmatrix}y_{t-1} \\ \mathbf{x}_{t-1}\end{bmatrix} \\ &+ \Big((\mathbf{I}_{k+1} - \boldsymbol{\Psi})\widetilde{\boldsymbol{\Phi}}^*(L) + \boldsymbol{\Psi}\Big)\Delta \mathbf{z}_t + \begin{bmatrix}u_{yt} \\ \boldsymbol{\epsilon}_{xt}\end{bmatrix}\end{aligned} \tag{4.34}$$

The matrices at the left side of $\Delta \mathbf{z}_t$ intentionally appear in a linear form to save some space. But the matrix form of this part is trivial if we just notice that:

$$\boldsymbol{\Psi} = \begin{bmatrix}0 & \boldsymbol{\omega}_{yx}\boldsymbol{\omega}_{xx}^{-1} \\ \mathbf{0}_k & \mathbf{0}_{k\times k}\end{bmatrix} \tag{4.35}$$

Looking at the Eq. (4.34) we can clearly see that this information can only pass through $\boldsymbol{\Phi}_{xy}(1)$ that now revealed itself in the conditional equation. So if we ensure that $\boldsymbol{\Phi}_{xy}(1) = \mathbf{0}$, so that the effect that is driven from each of the marginal equations to the conditional doesn't exists, only then we can safely continue with our univariate cointegration analysis using the Conditional ECM models.

In conclusion, although there may or may not be other cointegrating relationships between the $\mathbf{x}_t$ themselves, there must be at most one cointegrating relationship between y_t and $\mathbf{x}_t$ in order for the univariate analysis to be valid. If this should be the case, the $\mathbf{x}_t$ are called weakly exogenous for all the parameters in the conditional equation.

Until now we have set up the cointegrating relationship under the univariate analysis as if all the regressors were exogenous and also the cointegrating relationship under the multivariate case. Although, none of these are very useful in practice but what we can do is to frame the cointegrating relationship for the univariate case assuming that the marginal variables are at most weakly exogenous and there is only one long-run relationship between these variables and y_t. As we have discussed this requires that $\boldsymbol{\Phi}_{xy}(1) = \mathbf{0}$, so assuming that this is true and replacing it in the matrix on the left side of $\Delta \mathbf{z}_t$ in Eq. (4.34) we can construct a new cointegrating relationship for the univariate conditional model under this assumption[6].

$$\Phi_{yy}(1)y_{t-1} - \Big(\boldsymbol{\Phi}_{yx}(1) - \boldsymbol{\omega}_{yx}\boldsymbol{\omega}_{xx}^{-1}\boldsymbol{\Phi}_{xx}(1)\Big)\mathbf{x}_{t-1} \tag{4.36}$$

[6]Note that the cointegrating matrix in Eq. (4.31) is $\boldsymbol{\Phi}(1)$ while in the Eq. (4.34) is $(\mathbf{I}_{k+1} - \boldsymbol{\Psi})\boldsymbol{\Phi}(1)$

And since the equilibrating relationship is supposed to be stationary, it is hence stable over time around zero. According to this, combining the Eq. (4.36) and Eq. (4.34) and writing the equation in terms of y_t we end up with the long-run relationship in levels.

$$y_t = \frac{\alpha_{y0}}{\Phi_{yy}(1)} + \frac{\alpha_{y1}}{\Phi_{yy}(1)} t - \left(\frac{\boldsymbol{\Phi}_{yx}(1) - \boldsymbol{\omega}_{yx}\boldsymbol{\omega}_{xx}^{-1}\boldsymbol{\Phi}_{xx}(1)}{\Phi_{yy}(1)} \right) \mathbf{x}_t + u_t \tag{4.37}$$

4.5 Hypothesis testing

What we are primarily interested about is to test whether there is a linear combination of the independent variables with the dependent ones that forms a cointegrating (long-run) relationship. In this Chapter we describe two test. The first one (the F-bounds test) is a test for the absence of cointegration and it is practically a joint significance F-test (or Wald-test) on some parameters of the UCECM Eq. (4.6). The second one (the t-bounds test) can be used if the results from the F-bounds test are statistically significant to reduce the possibility of a false positive result.

4.5.1 Deterministics in the Long-Run

One can say that the F-bounds test is non-standard test as we have a different specification depending on whether the deterministic components (constant (α_0) and linear trend (α_1) from Eq. (4.5) and (4.6)) enters the UCECM but also depending on whether they enter the long-run relationship too. In general, the common parts of the Null and the Alternative hypothesis of the F-bounds test are:

$$\begin{aligned} H_0 &= \psi(1) \cap \{\beta_j(1)\}_{j=1}^{k} = 0 \\ H_1 &= \psi(1) \cup \{\beta_j(1)\}_{j=1}^{k} = 0 \end{aligned} \tag{4.38}$$

As for the existence of the deterministics in the UCECM and the restrictions of whether they may enter the long-run relationship we can discriminate between 5 different cases:

- Case 1 (No constant, no trend)
- Case 2 (Restricted constant, no trend)
- Case 3 (Unrestricted constant, no trend)
- Case 4 (Unrestricted constant, Restricted trend)
- Case 5 (Unrestricted constant, Unrestricted trend)

Where in case that a deterministic component doesn't exists, it is absent from every equation implying that:

$$\begin{aligned} &\boldsymbol{\mu} = \mathbf{0} \ \text{ or } \ \boldsymbol{\gamma} = \mathbf{0} \\ \text{With respect to UCECM: } &\alpha_0 = 0 \ \text{ or } \ \alpha_1 = 0 \\ \text{With respect to UVECM: } &\boldsymbol{\alpha}_0 = 0 \ \text{ or } \ \boldsymbol{\alpha}_1 = 0 \end{aligned} \tag{4.39}$$

In case where a deterministic component is restricted it appears in the UCECM but is restricted to a specific linear combination of $\psi(1)$ and $\beta_j(1)$ so that it also enters the cointegrating relationship ECT_{t-1}.

$$\begin{aligned} &\boldsymbol{\mu} \neq \mathbf{0} \ \text{ or } \ \boldsymbol{\gamma} \neq \mathbf{0} \\ \text{With respect to UCECM: } &\alpha_0 = \psi(1)\mu_y + \sum_{j=1}^{k} \beta_j(1)\mu_{x_j} \\ &\alpha_1 = \psi(1)\gamma_y + \sum_{j=1}^{k} \beta_j(1)\gamma_{x_j} \\ \text{With respect to UVECM: } &\boldsymbol{\alpha}_0 = \boldsymbol{\Phi}(1)\boldsymbol{\mu} + (\sum_{i=1}^{p} i\boldsymbol{\Phi}_i)\boldsymbol{\gamma} \\ &\boldsymbol{\alpha}_1 = \boldsymbol{\Phi}(1)\boldsymbol{\gamma} \end{aligned} \tag{4.40}$$

In the case where a deterministic component is unrestricted it participates in the UCECM but without any restrictions about it (except of the restriction of $\neq 0$ that it forces it to be part of the UCECM). This way it is estimated as a regular scalar deterministic component and it doesn't enters the long-run relationship ECT_{t-1}.

$$\begin{aligned} &\boldsymbol{\mu} \neq \mathbf{0} \ \text{ or } \ \boldsymbol{\gamma} \neq \mathbf{0} \\ \text{With respect to UCECM: } &\alpha_0 \neq 0 \ \text{ or } \ \alpha_1 \neq 0 \\ \text{With respect to UVECM: } &\boldsymbol{\alpha}_0 \neq 0 \ \text{ or } \ \boldsymbol{\alpha}_1 \neq 0 \end{aligned} \tag{4.41}$$

Additionally, when a component is under this restriction, it is also added in the joint significance test under the Null (restricted also to be $= 0$) as it is part of the relationship now.

4.5.2 F-bounds test

The common parameters that always exist in every case are the following and the restricted ones should be added if needed.

$$\begin{aligned} \text{With respect to UCECM: } &H_0 : \psi(1) = \beta_j(1) = 0, \forall j \\ \text{In practice: } &H_0 : b_0 = b_j = 0, \forall j \end{aligned} \tag{4.42}$$

Replacing appropriately the elements from (4.39), (4.40) and (4.41) in the Eq. (4.6), (4.8) and (4.42) we can construct the models and the test for each of the 5 cases.

Notice that the Null hypothesis with respect to the UCECM under the assumption of $\mathbf{\Phi}_{xy}(1) = \mathbf{0}$ is:

$$\text{F-bounds test } H_0 : \Phi_{yy}(1) = 0 \quad \text{and} \quad \mathbf{\Phi}_{yx}(1) - \boldsymbol{\omega}_{yx}\boldsymbol{\omega}_{xx}^{-1}\mathbf{\Phi}_{xx}(1) = \mathbf{0}'_k \tag{4.43}$$

The Eq. (4.43) is very important because this is what the test is all about. Starting to describe how the F-bounds test works we should start talking about the rank of the matrices. It is easy to define that since there are k marginal variables the cointegrating matrix associated with the matrix of the marginals can have a maximum rank of $r_x = k$ and a minimum of $r_x = 0$, where $rk(\mathbf{\Phi}_{xx}(1)) = r_x$ is the rank of the matrix $\mathbf{\Phi}_{xx}(1)$. We should also define that r_z is the rank of the cointegrating matrix $(\mathbf{I}_{k+1} - \mathbf{\Psi})\mathbf{\Phi}(1)$ which accounts for the whole VAR.

Lets now discriminate between two polar cases. The first is the case where $r_x = k$, which is equivalent to $\mathbf{x}_t \sim I(0)$. Remember that this is the maximum rank that can be achieved for this matrix. The other polar case is where $r_x = 0$ which means that $\mathbf{x}_t \sim I(1)$ and this is the minimum possible rank for the matrix.

The whole system contains the k marginal variables plus the conditional one, hence the cointegrating matrix associated with the matrix $\mathbf{z}_t$ can only have a minimum rank of $r_z = r_x$ and a maximum of $r_z = r_x + 1$.

Finally, notice that the Null Hypothesis Eq. (4.43) is satisfied only when $r_z = r_x$. The only possible alternative for r_z is to be equal to $r_x + 1$ which as we easily understand, and we will describe in detail later, can have three possible outcomes meaning that the tested joint equalities can break either at the same time or one at a time.

Combining the above information we can see how one can test for the H_0 in Eq. (4.43). Suppose that we know that r_x is a known number between 0 and k. If we calculate the critical value that corresponds to this value of r_x from the non-standard limiting distribution and then compare this with the F-statistic from the joint Wald test Eq. (4.42) we are eventually testing for the H_0 in Eq. (4.43). As per usual, if the F-statistic is greater than the critical value then we reject the H_0.

One of the advantages of the F-bounds test is that we don't have to know for sure the exact rank r_x or in other words the order of integration of $\mathbf{x}_t$. Instead, we can use the two polar cases we spoke about before and calculate two critical values. One for the case where $\mathbf{x}_t \sim I(0)$ or $r_x = k$, say ξ_L, and one for the case where $\mathbf{x}_t \sim I(1)$ or $r_x = 0$, say ξ_U. These are the lower and the upper bounds respectively. So irrespective of whether $\mathbf{x}_t$ is $I(0)$, mutually cointegrated or $I(1)$ we can compare the F-statistic with the two polar critical values and we can have the following possible results.

- $F < \xi_L < \xi_U$: Whatever the order of integration of $\mathbf{x}_t$ is, we are unable to reject the H_0. Hence, no cointegration between y_t and $\mathbf{x}_t$ exists.
- $\xi_L < F < \xi_U$: If $r_x = k$, and so the critical value is ξ_L, we can reject the H_0. If $r_x = 0$, and so the critical value is ξ_U, we can't reject the H_0. Accordingly, knowing the precise order of integration we can conclude, but this requires further testing and it takes away the basic advantage of the F-bounds test.

- $\xi_L < \xi_U < F$: We can reject the H_0 even if the order of integration of $\mathbf{x}_t$ is $I(1)$ and so r_x takes the minimum possible value which is 0. This also means that $\boldsymbol{\Phi}_{xx}(1) = 0$ and so $\boldsymbol{\omega}_{yx}\boldsymbol{\omega}_{xx}^{-1}\boldsymbol{\Phi}_{xx}(1) = 0$ which reduces the H_0 in Eq. (4.43) to $H_0: \ \Phi_{yy}(1) = 0 \quad \text{and} \quad \boldsymbol{\Phi}_{yx}(1) = \mathbf{0}_k'$.

But what does rejecting the H_0 actually means? Remember that the H_0 in Eq. (4.43) consists of two equalities, so rejecting the possibility of these two happening together doesn't always mean that they both differ from zero. We end up with three possible scenarios.

- $\Phi_{yy}(1) = 0, \quad \boldsymbol{\Phi}_{yx}(1) \neq \mathbf{0}_k'$: The cointegrating relationship is nonsensical as $\Phi_{yy}(1)$ is the denominator in Eq. (4.37) and the cointegrating relationship is not defined. Still $y_t \sim I(1)$.

- $\Phi_{yy}(1) \neq 0, \quad \boldsymbol{\Phi}_{yx}(1) = \mathbf{0}_k'$: The cointegrating relationship exists but it is degenerate. This means that a relationship exists but it is through the short-run dynamics $\Delta\mathbf{x}_t$ and although it is seemingly stable it diverges in the long-run. Under this possible scenario $y_t \sim I(0)$[7].

- $\Phi_{yy}(1) \neq 0, \quad \boldsymbol{\Phi}_{yx}(1) \neq \mathbf{0}_k'$: A cointegrating relationship exists. Interestingly, while this may mean that $y_t \sim I(1)$ and hence a usual cointegrating relationship exists, this may be also the case where $y_t \sim I(0)$. This also means that all the marginal variables in $\mathbf{x}_t \sim I(0)$, implying that $\mathbf{z}_t \sim I(0)$, and so the system cointegrating matrix is full rank ($r_z = (k+1)$). To correct the above statement about a 'cointegrating relationship', a long-run relationship also exists in this case (an OLS estimation in levels would not lead to spurious regression) but this is not what we call a *usual cointegrating relationship*.

4.5.3 t-bounds test

We realize that just by rejecting the H_0 we can't be sure if a usual cointegration exists or the results are degenerate or just nonsensical. Trying to eliminate at least the first false positive case (the nonsensical), M. Hashem Pesaran et al. (2001) proposed an additional test for $\Phi_{yy}(1) = 0$ through a non-standard ADF type regression using the usual t-statistic. Unfortunately, the t-bounds test can only be used in three our of the five variants of the equation in regards to the deterministics. These are the cases where there is no restriction on the deterministic components (`Case I`, `Case III` and `Case V`).

One uses the t-bounds test in the same manner as the F-bounds test whilst in this case the lower bound, ζ_L, corresponds to $\mathbf{x}_t \sim I(1)$ and the upper bound, ζ_U, corresponds to $\mathbf{x}_t \sim I(0)$. As per usual, the two sided t-test rejects the Null if the absolute value of the t-statistic is greater than the absolute value of the critical value.

[7] This is the case in Section 5.5 where the dependent variable is trend stationary

- $|t| < |\zeta_L| < |\zeta_U|$: Failing to reject the H_0 we automatically concludes that $\Phi_{yy}(1) = 0$ and thus we are in the first scenario of the nonsensical relationship.

- $|\zeta_L| < |t| < |\zeta_U|$: As with the F-bounds test, we reject the H_0 if $\mathbf{x}_t \sim I(0)$ but not if $\mathbf{x}_t \sim I(1)$. In order to conclude for the case where $\mathbf{x}_t$ is mutually cointegrated among themselves we have to know the rank of the cointegrating matrix and compare with the relevant critical value.

- $|\zeta_L| < |\zeta_U| < |t|$: One can reject the Null Hypothesis that $\Phi_{yy}(1) = 0$ and that leaves us with two possible results. Either $\boldsymbol{\Phi}_{yx}(1) \neq \mathbf{0}'_k$ and there is a cointegrating relationship either $\boldsymbol{\Phi}_{yx}(1) = \mathbf{0}'_k$ and there is a degenerate relationship.

4.6 Test requirements

After all this technical details, here is a good place to sum up the whole practical process and especially the requirements under which the application of the test is valid. These restrictions that were mentioned in details previously are the followings.

Order of integration

One of the first things we usually check before any time series analysis is the order of integration of the series. This is crucial in this case also, but instead, this particular test does not distinguish between the $I(0)$ and the $I(1)$ series as long as they play the role of the independent variables. So, in this case, we care more about testing whether the independent variables are $I(2)$ (or greater) or not. Any $I(0)$ or $I(1)$ independent variable can participate in the model but not the $I(2)$ ones.

As for the dependent variable, this has to be exactly $I(1)$ for a cointegrating relationship to exist. In addition, two practical examples are presented in this paper, in the chapter 5.5 and the chapter 5.6, where the dependent variables were found to be $I(0)$ and in fact trend stationary. In these cases, we continue on purpose with the analysis to show how degenerate cases like these behave in practice.

Additionally, potential structural breaks have to be taken into account along the testing process of a Unit Root. ADF type tests are known to have low statistical power and a potential outlier may force the test to incorrect diagnosis. The same may happen with a structural break. A series can have a stationary behavior in the first regime and a different but still stationary behavior in the next regime. In this case, a conventional Unit Root test will be confused, as it has not taken into account the break in the series, and will conclude in favor of the existence of a Unit Root.

Serially independent and homogeneous error term

Theoretically, a correctly specified ARDL model, using the appropriate number of lags for the dependent as long as for the independent variables, is free of autocorrelation problems. Therefore, it is essential that an ARDL model that its residuals are serially correlated is not a well specified model and it doesn't represent the true DGP of the underlying series. In fact, in a case like this, the estimated coefficients will be biased and inconsistent and the standard errors will be invalid.

Another problem arises when the error term is heterogeneous. In contrast to the previous case of autocorrelation, here in the case where the error term is heterogeneous, the OLS estimators will be still consistent but the standard error will be again invalid. So, this is less of a problem when it comes to the estimation of the parameters but it is when we have to make inference.

The usual desirable residuals properties (normality etc.) should also hold.

Dynamically stable parameters

Reminding that the main goal of this test is to find whether there is a long run relationship between the variables. So, the estimated parameters have to be dynamically stable in order for our final estimated long run relationship to be sensible. Many test can be applied for this reason but in this paper we use the usual CUSUM and CUSUMSQ tests, despite the fact that they are known for low statistical power.

A single cointegrating relationship for y_t

As we discussed in details in the Chapter 4.4.2, there must be at most one cointegrating relationship between y_t and $\mathbf{x}_t$. Otherwise the results from our univariate analysis will be incorrect because we don't account for these simultaneous effects. In the Chapter 5 we also model separately every variable to test for other cointegrating relationships in the two cases (Greece and France) where a long-run relationship was found but the results (especially for the case of France) were not supporting our conditional model.

Chapter 5

Application: Income inequality

5.1 Data summary

The sample we used for the main modeling consists of yearly data that spans from 1971 to 2014, while could extend our sample to the past and to more recent data but this was not possible for all variables. Notice that each country is referred using the standard ISO code[1].

Our choice for income inequality measurement is the 1% Top Income Shares (`tis01`). The data for France, USA, and United Kingdom were obtained from the World Inequality Database[2] while those for Greece was computed using the same appropriate methodology of Piketty (2001) (Appendix B) by A. Livada and K.Chryssis who provided me with the data set for Greece. The computation of the 1% tis indicator was based on the pre-tax national income. More specifically, for the cases of Greece and France the estimation of the 1% tis was based entirely on the tax units while for United Kingdom, due to data availability reasons, we have used a mixed approach. Tax units for the period 1971-1989 and individual units for the period 1990-2014[3]. For the USA, the 1% tis was estimated based on the equal split as this was the only measure available.

The 1% tis variable for United Kingdom, which was finally estimated as described above, had a missing value (for the sample under investigation) at the year 1980 which was replaced using a cubic spline interpolation.

The Figure 5.1 shows the long history of the 1% tis variable for all the countries. We can see that there is a common U-shape for all the countries. One could say that there

[1] Greece (GRC), France (FRA), United States of America (USA), United Kingdom of Great Britain and Northern Ireland (GBR)

[2] `http://wid.world`

[3] The final series created has a smooth continuity as the methodology of estimation (based on tax units, individual or equal split) makes no big difference on the estimated tis

is a deterministic decline for all countries for a long period of time[4] and a deterministic rising thereafter. Under this point of view we should assume that there is a structural break[5] and continue the rest of the analysis under this spectrum. The other alternative is to consider this behavior as the result of a *RW* that fluctuates around a mean creating the sense of deterministic trends in some parts of the series.

In this thesis we treat each of the series as a special case. The `tis01` series of USA has reached its historical high and it doesn't seem to dramatically change its behavior. We consider that this case incorporates a deterministic trend. The `tis01` series for GBR has reached its historical high around 2009 and it seems like having a slightly declining way from 2007 since the last observation in 2014, similar to the decline from 1971 to 1978. For the rest of the series, covering the most of the time span in our sample, we suppose that it is driven from a deterministic linear trend[6]. A declining shape applies also for FRA after 2007 and in this case it seems like a repeated cycle mimicking the behavior of the series in the years from 1945 to 1985. Finally, GRC looks like it has also passed its historical high but although its behavior is quite the same as this for the other countries with some lag of one or two decades we believe that is safer not to force a deterministic upward trend in the cases of GRC and FRA.

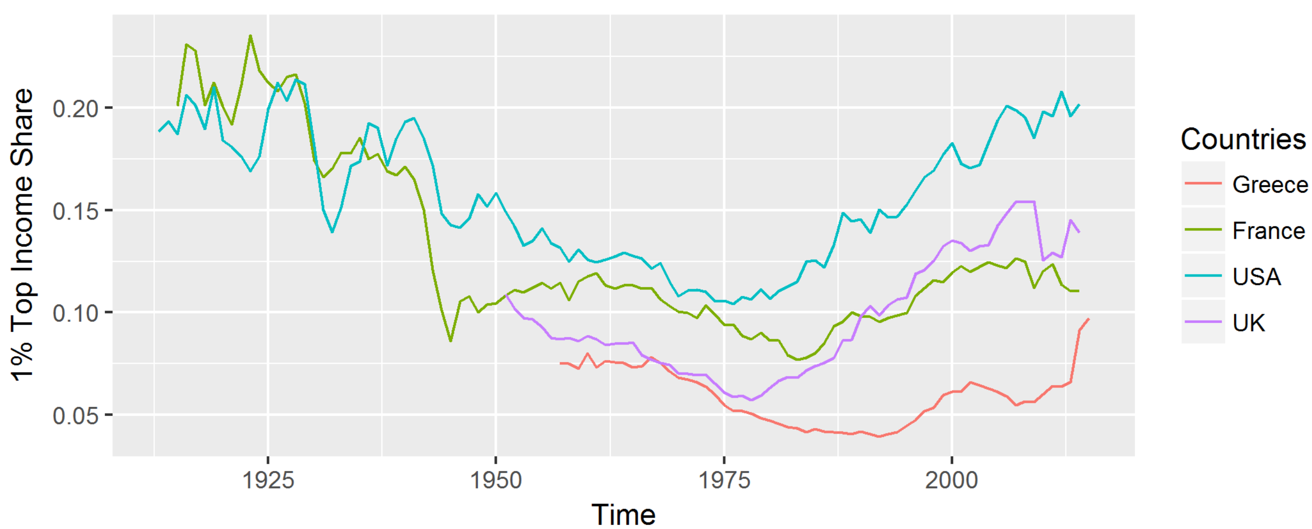

Figure 5.1: Extended 1% tis series for all countries

In Table 5.1 we present some summary statistics about `tis01` series for each country. The results in the table are calculated using the working subsample covering the years from 1971 to 2014.

[4]Until the late '70 for USA and United Kingdom, around the mid 80' for France and around the mid 90' for Greece

[5]as the deterministic structure of the series breaks/changes

[6]Unfortunately, the current version of the software we use for the analysis, EViews 10, does not support a discrimination like this in the long-run relationship. Our effort to use dummy variables to model the short-run relationship didn't have a considerable impact to the results.

Table 5.1: 1% Top Income Share summary statistics

	Min.	**Median**	**Mean**	**St.Dev.**	**Max.**
tis01_GRC	0.039	0.053	0.054	0.011	0.091
tis01_FRA	0.077	0.100	0.103	0.015	0.127
tis01_USA	0.104	0.148	0.151	0.035	0.208
tis01_GBR	0.057	0.103	0.102	0.033	0.154

The explanatory variables we are using throughout this thesis along with their secondary source (database) are the following:

- `credit` denotes the domestic credit to private sector (% of GDP)
  - It includes loans, nonequity securities, trade credits etc provided to the private sector. — World Bank
- `edu` represents the enrolment in tertiary education (% of total population)
  - It is constructed dividing the total number of students enrolled at public and private tertiary education institutions by the total population of the country. — World Bank
- `gdp` denotes the GDP per capita (constant 2010 US$)
  - It is the GDP divided by midyear population — World Bank
- `infl` denotes the inflation, consumer prices (annual %)
  - It is the consumer prices as the percentage change from the previous year — OECD
- `popg` denotes the population growth (annual %)
  - It is calculated as the exponential rate of growth of midyear population, expressed as a percentage — World Bank
- `trade` denotes the Trade (% of GDP)
  - It is the sum of exports and imports of goods and services measured as a share of gross domestic product — World Bank

Getting a good sense of the data is very important. Among others, it helps us make decisions in situations where the statistical tests are there for us as a tool but the appropriate specification or a final decision has to be made based on our knowledge and our experience. For this reason, we present here the Figure 5.2 that shows the variable `tis01` in levels and the Figure 5.3 that shows the variable `tis01` in first differences for each of the four countries.

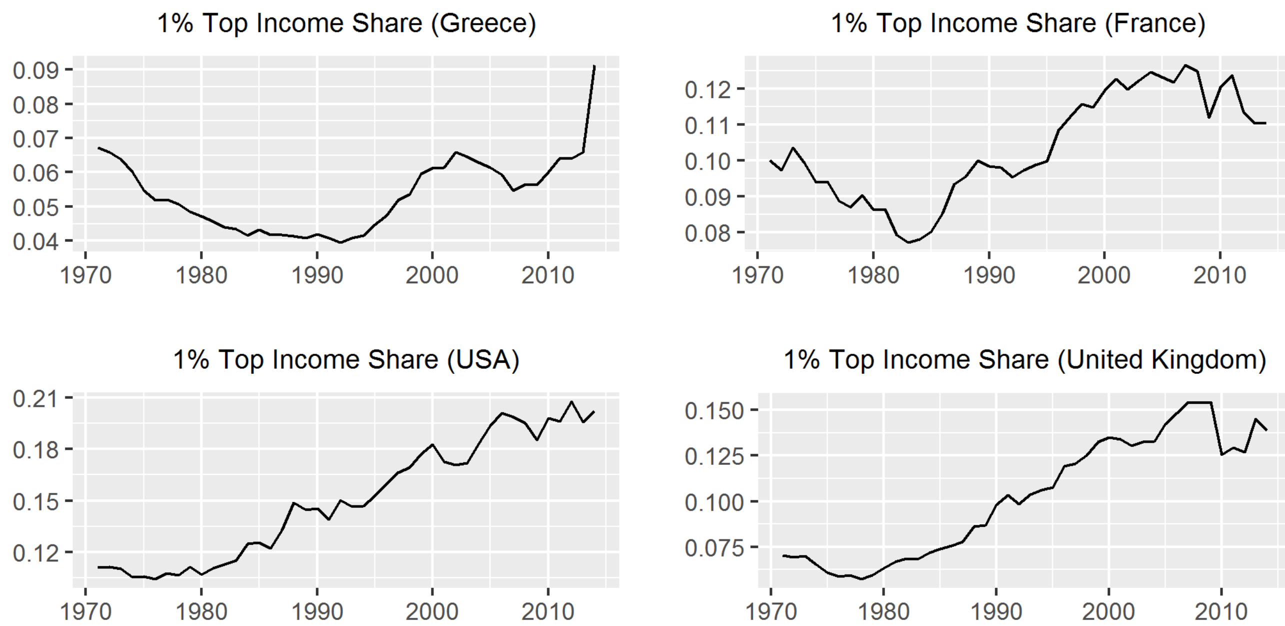

Figure 5.2: tis01, variables in levels

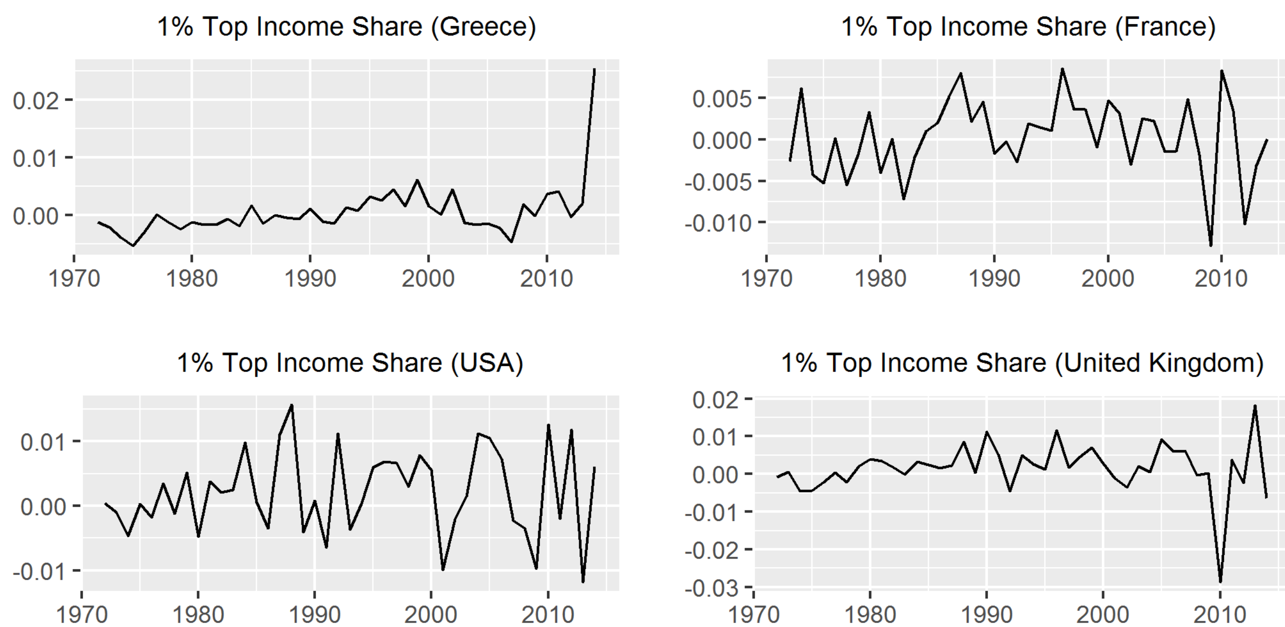

Figure 5.3: tis01, variables in first difference

Also, at the beginning of each of the following cases we present the figures of the explanatory variables in levels and in first differences. For each country, the figures with the variables in levels are the Figures 5.4, 5.11, 5.17 and 5.20 and the figures with the variables in first differences are the Figures 5.5, 5.12, 5.18 and 5.21.

5.2 Methodological strategy

In this section, we present the general methodological strategy we follow for the analysis.

First, we apply the appropriate stationarity and Unit Root tests to conclude for the order of integration of the variables. If the order of integration of the independent variables is not $I(2)$ or greater and the order of the dependent variable is $I(1)$ we are good to go for the rest of the analysis. Notice that when the results are not the optimal for applying the test, we still continue with this case to better understand the reasons that led us here.

Next, we select some candidate cases[7] that we think they may be appropriate for our data. This is a selection that is mainly based on the graphical behavior of the variable, our belief based on the economic theory, the results of the models (statistical significance of the deterministic trend components etc) and the final estimated results (whether they make sense or not).

When we decide on the appropriate case, we try to achieve the most parsimonious model that is free of some crucial problems as mentioned above in 4.6.

Once we decide on a suitable model that we think that describes the DGP of the underlying process reasonably well, we go ahead for the actual test of non-cointegration using the F-bounds test. Additionally, if the chosen case is `Case I`, `Case III` or `Case V` we can also make use of the t-bounds test in order to shed some light on the results and reduce the chances of a potential false positive.

If the results are favorable, we can continue forming the long-run relationship, calculating the long-run multipliers, the short-run dynamics, the speed of the adjustment to the long-run equilibrium etc., as described in the Chapters 4.1 and 4.3. If the results show no signs of cointegration among the variables, the formation of the long-run relationship and those that was described above would be meaningless because the results would be spurious.

5.3 The case of Greece

In the Table 5.2 we present some standard unit root test along with some breakpoint unit root tests which test for the existence of a UR under the hypothesis of a potential structural break on the deterministic trend components.

[7] The choice is among the `Case I`, `Case II`, `Case III`, `Case IV` and `Case V`

What we are actually interested in in these tables is to see if the dependent variable (`tis01`) is $I(0)$ because this would violate the requirements of the test as described in 4.6.

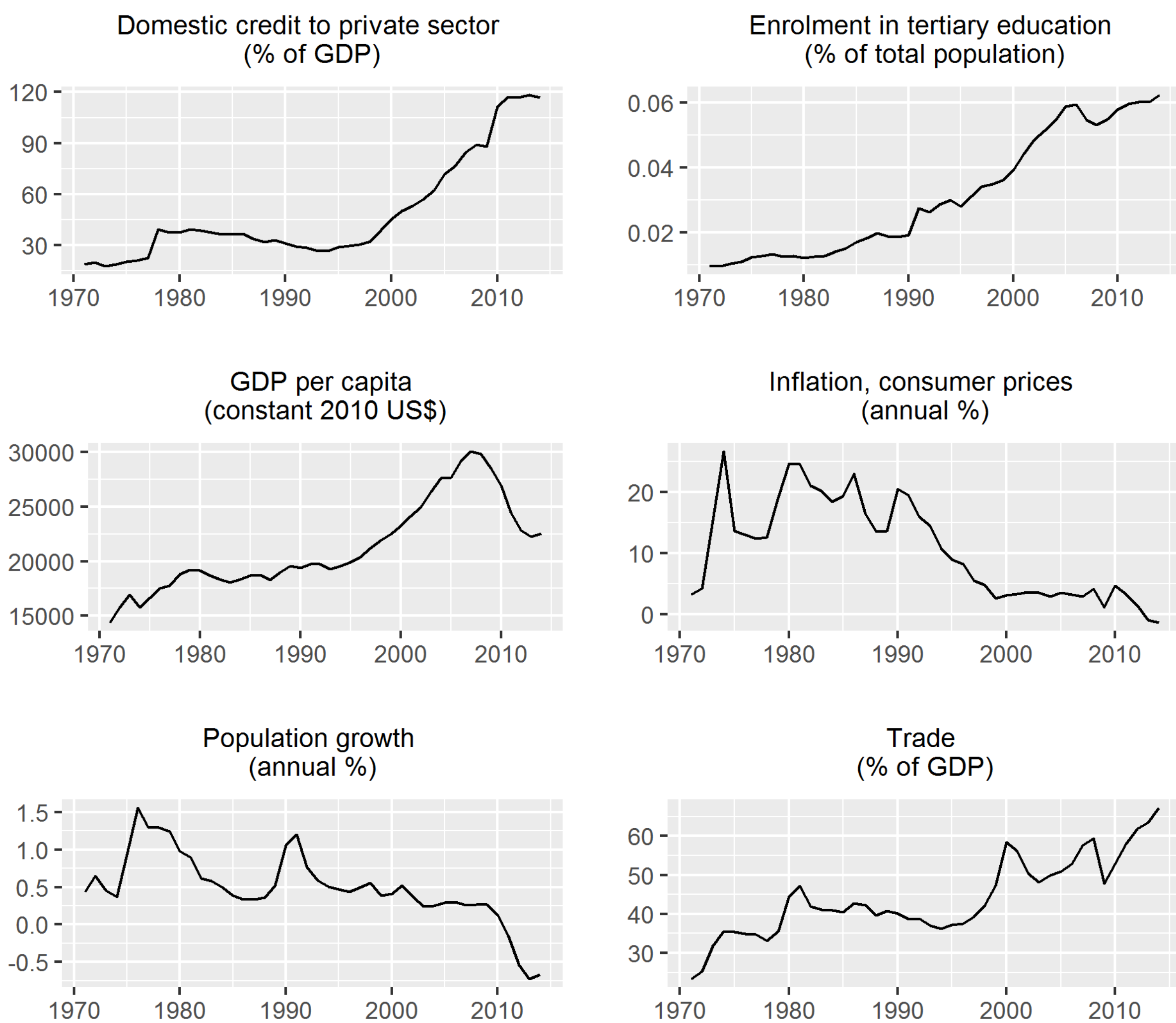

Figure 5.4: Greece, variables in levels

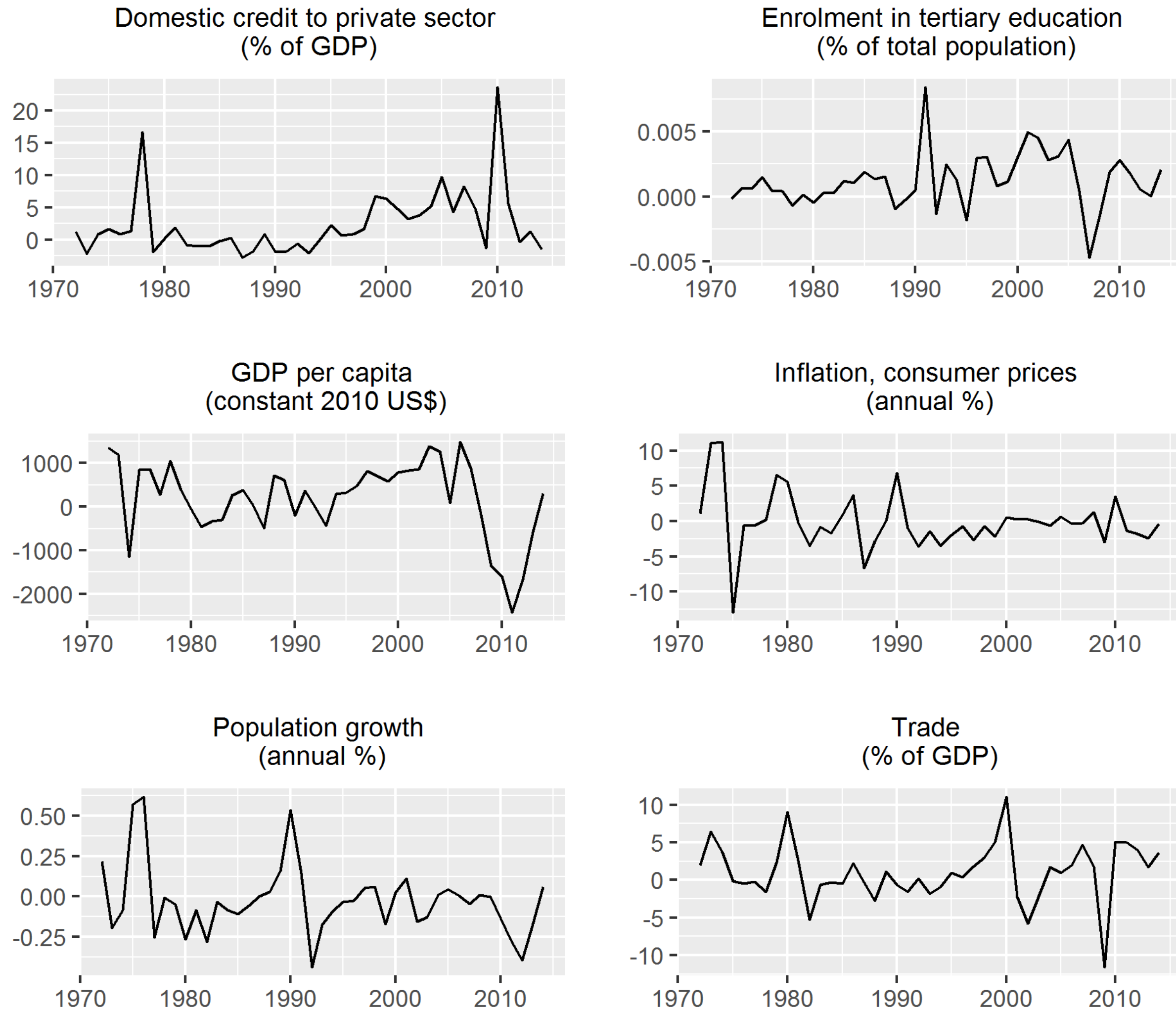

Figure 5.5: Greece, variables in first differences

We applied the same for the cases of the other countries too and they can be found in Tables 5.11, 5.17 and 5.22. Interpreting some interesting results from the Table 5.2 we can see that the first row, based on the ADF and PP tests, indicates that the variable `tis01` has a Unit Root. The KPSS test fails to reject the Null Hypothesis that the series is stationary but this test should not be interpreted in this cases as the H_0 of the test assumes that the series is trend stationary. Here it is clear that there isn't a deterministic trend in the data but the test gets tricked from the general upward trend in the data. In the rest of this paper, a decision may have been made even if some results seem to be in conflict. The reasoning behind this is not always stated in details as above but it should logically come out of the data and the situation. The results for some of the other variables are also mixed but for the purposes of this tests we don't have to make a conclusion on whether the independent variables are $I(0)$ or $I(1)$ as long as they are not $I(2)$.

Table 5.2: Unit Root tests for Greece in Levels

	Greece, variables in levels:											
	Standard Unit Root Tests											Breakpoint Unit Root Tests
	ADF	Exogenous	Lags	PP	Exogenous	KPSS	Exogenous	ADF_{bp}	Specification	Lags	Break Type	Break Date
tis01	0.911	n	0	0.561	n	0.347	c	-5.068^{*}	t-b	6	i	1990
credit	1.477	c	0	1.241	c	0.187^{**}	t	-3.371	t-c	0	i	1997
edu	0.622	c	0	0.517	c	0.171^{**}	t	-3.645	t-b	5	a	1990
gdp	-1.489	c	1	-1.717	c	0.081	t	-3.583	t-t	1	a	2003
infl	-1.480	c	0	-1.391	c	0.167^{**}	t	-4.431^{*}	c-c	0	i	1993
popg	-1.522	c	1	-0.489	c	0.100	t	-3.914	t-c	1	i	2010
trade	-1.089	c	0	-0.904	c	0.122^{*}	t	-4.387^{*}	t-t	1	i	1996
H_0:	Unit Root			Unit Root		No Unit Root		Unit Root				

Note: *p<0.1; **p<0.05; ***p<0.01

For the Exogenous columns, c indicates that a constant intercept participates in the test equation,
t indicates the participation of a linear trend along with the constant and n indicates than no exogenous variables participate.
The column Lags shows the number of lags that was used (where applicable).
The number of lags was chosen based on the BIC criterion. The AIC was also consistent as for the stability of the test results
but it overestimated the number of lags in some cases.
The first letter on the Specification column indicates the trend nature of the data (therefore a level less than the column Exogenous).
The second letter indicates in which of these instances the break occurs.
Here t indicates a break in the treand, c a break in the intercept and b indicates a break in both of them.
In the Break Type column, a stands for Additive Outlier Test and i stands for Innovational Outlier Test.
The first one is used when we assume that the break occurs immediately while the second one when the break occurs slowly.
The breakpoint selection is mainly based on Dickey-Fuller min t-statistic and when appropriate on the max absolute t-statistics of the intercept/trend.

Table 5.3: Unit Root tests for Greece in First Differences

	Greece, variables in first differences:											
	Standard Unit Root Tests											Breakpoint Unit Root Tests
	ADF	Exogenous	Lags	PP	Exogenous	KPSS	Exogenous	ADF_{bp}	Specification	Lags	Break Type	Break Date
Δtis01	-1.255	n	0	-1.115	n	0.527^{**}	c	-3.795	c-c	0	a	2014
Δcredit	-4.676^{***}	n	0	-4.915^{***}	n	0.385^{*}	c	-6.652^{***}	c-c	0	i	1998
Δedu	-5.277^{***}	c	0	-5.263^{***}	c	0.271	c	-5.897^{***}	c-c	0	a	2007
Δgdp	-3.366^{***}	n	0	-3.439^{***}	n	0.166	c	-4.709^{**}	c-c	0	i	2007
Δinfl	-6.004^{***}	n	0	-6.059^{***}	n	0.297	c	-7.414^{***}	c-c	0	i	1975
Δpopg	-4.718^{***}	n	0	-4.526^{***}	n	0.191	c	-5.962^{***}	c-c	1	i	1990
Δtrade	-5.109^{***}	n	0	-4.965^{***}	n	0.126	c	-6.691^{***}	c-c	1	i	2000
H_0:	Unit Root			Unit Root		No Unit Root		Unit Root				

Note: *p<0.1; **p<0.05; ***p<0.01

For the Exogenous columns, c indicates that a constant intercept participates in the test equation,
t indicates the participation of a linear trend along with the constant and n indicates than no exogenous variables participate.
The column Lags shows the number of lags that was used (where applicable).
The number of lags was chosen based on the BIC criterion. The AIC was also consistent as for the stability of the test results
but it overestimated the number of lags in some cases.
The first letter on the Specification column indicates the trend nature of the data (therefore a level less than the column Exogenous).
The second letter indicates in which of these instances the break occurs.
Here t indicates a break in the treand, c a break in the intercept and b indicates a break in both of them.
In the Break Type column, a stands for Additive Outlier Test and i stands for Innovational Outlier Test.
The first one is used when we assume that the break occurs immediately while the second one when the break occurs slowly.
The breakpoint selection is mainly based on Dickey-Fuller min t-statistic and when appropriate on the max absolute t-statistics of the intercept/trend.

In the Table 5.3 we present the same tests as in the previous Table 5.2 but here the variables under investigation are transformed into their first differences. Here we want to see whether the independent variables and the dependent one are $I(2)$ (actually greater than $I(1)$, in which case the initial variable can't participate in the model) or not.

We applied the same for the cases of the other countries too and they can be found in Tables 5.12, 5.18 and 5.23.

Table 5.3 shows that all the independent variables in first differences are stationary. In this particular case we have seen that they are not stationary in levels so we conclude that they are $I(1)$, but this conclusion is not necessary because we are fine as long as their order of integration is not greater than $I(1)$. On the other hand, the dependent variable `tis01` in first differences seems to be non-stationary containing a Unit Root.

But taking a closer look at the Figure 5.3 we notice that the last observation corresponding to the year 2014 is unexpectedly high and this drives the tests to conclude for the existence of a Unit Root. Naturally, a single observation in time should not significantly affect the internal process of the data. In order to test for this, in the next Table 5.4 we did the same tests but this time excluding the last observation for the year 2014. And as expected, the results turned over concluding that the dependent variable `tis01` is $I(1)$. As for the last observation, it's not known yet if it the start of a new regime, or a random shock. In case that we aim to forecast we should know if we should treat this as a new regime to set the rest of the series (from 2014 and on) with this new behavior or to control for this with a once-off dummy variable. However, we don't focus on this kind of forecasting analysis rather than we try to explain the relationship between `tis01` and the rest of the independent macroeconomic variables and so the use of a once-off dummy should be just fine for our case if needed.

Table 5.4: Unit Root tests for Greece in First Differences (excluding 2014)

	Greece, variables in first differences: Sample: 1971-2013 (excluding 2014)												
	Standard Unit Root Tests							Breakpoint Unit Root Tests					
	ADF	Exogenous	Lags	PP	Exogenous	KPSS	Exogenous	ADF_{bp}	Specification	Lags	Break Type	Break Date
Δtis01	-2.359^{**}	n	1	-3.917^{***}	n	0.429^{*}	c	-4.777^{**}	c-c	0	a	1984
H_0:	Unit Root			Unit Root		No Unit Root		Unit Root				

Note: *p<0.1; **p<0.05; ***p<0.01

For the Exogenous columns, c indicates that a constant intercept participates in the test equation,
t indicates the participation of a linear trend along with the constant and n indicates than no exogenous variables participate.
The column Lags shows the number of lags that was used (where applicable).
The number of lags was chosen based on the BIC criterion. The AIC was also consistent as for the stability of the test results
but it overestimated the number of lags in some cases.
The first letter on the Specification column indicates the trend nature of the data (therefore a level less than the column Exogenous).
The second letter indicates in which of these instances the break occurs.
Here t indicates a break in the treand, c a break in the intercept and b indicates a break in both of them.
In the Break Type column, a stands for Additive Outlier Test and i stands for Innovational Outlier Test.
The first one is used when we assume that the break occurs immediately while the second one when the break occurs slowly.
The breakpoint selection is mainly based on Dickey-Fuller min t-statistic and when appropriate on the max absolute t-statistics of the intercept/trend.

To sum up, as concerns the order of integration of our variables, we are good to go for the rest of the analysis for the case of Greece.

For the case of Greece many modeling attempts were made and the best found to be the scenario of `Case II` with a restricted constant. Other choices like `Case III` or `Case IV` either turned our nonsensical based on the t-bounds cointegration test or they were obviously not converging to equilibrium even graphically.

Initially, the best model we found was the $ARDL(1, 0, 0, 3, 0, 0, 0)$ using the `Case II` testing restrictions and this is the model that Tables 5.5, C.2, 5.6, C.1, C.4 and C.5 are referring to.

Testing this model for autocorrelation and heteroskedasticity issues, we found out that it suffers from heteroskedasticity. The results from the tests can be found in the second half of the Table C.2. For this reason, we estimated robust standard errors using heteroskedasticity consistent covariance matrix[8] and the results of the final model are listed in the Table 5.5. Our selected model $ARDL(1, 0, 0, 3, 0, 0, 0)$ is free of autocorrelation problems regarding the results in Table C.2 and what we are interested in now is the Table 5.6 where we can see that there are strong evidence for a long-run levels relationship. The Figure 5.6 shows how the estimated Long-Run cointegrating relationship fits to the real data.

The graphical representation is a very important tool along with the cointegration test. And this graph doesn't totally satisfy our conclusion for the existence of cointegration in levels. For this reason, we made a simple modification to this model. We constructed a once-off dummy variable controlling only for the last observation in the year 2014 (as we did for the Unit Root tests) and the results are summarised in the Tables 5.7, C.3, 5.8, C.6, 5.9 and 5.10 and they concern the selected model ARDL(1,0,1,1,1,0,0). In contrast with the rest of the independent dynamic regressors, the dummy variable `D2014` was forced to participate only in the short run part of the model as a fixed regressor and not in the long run relationship in order to avoid any potential manipulation of the long run dynamics. This way, we treat this observed point as a short run shock for which we control for. We also see how the best selected model changed from $ARDL(1, 0, 0, 3, 0, 0, 0)$ to $ARDL(1, 0, 1, 1, 1, 0, 0)$ with the use of a dummy variable affecting just one observation and so how influential this observation was that may was misleading the results. We notice that this model doesn't have heteroskedasticity problems anymore. This model also passes the autocorrelation and the F-bounds cointegration tests.

What's even more interesting is the Figure 5.9 where we clearly see how the second model fits the original data much more better. It probably fluctuates a little more in the first years from 1971 to the mid 90's but it captures much better the trend after that.

A very important check is the one about the dynamic stability of the parameters. The CUSUM test in Figure 5.7 shows that the parameters are dynamically stable and

[8] HAC (Newey-West)

while the CUSUM Square test in Figure 5.8 indicates a slight break during 1996, it doesn't go far from the 5% limits and it gets back to stability in 1999 indicating that the residuals variance is stable too.

The graph doesn't totally satisfy our decision but we could say that it is a quite good modeling approach. So, assuming that our conclusion for long-run relationship is true, we can proceed interpreting the results of our interest. In the Tables 5.10 and 5.9 where we can find information about the speed of adjustment back to the long-run equilibrium after an instant shock, the short-run multipliers and the long-run multipliers which form the final long-run relationship. The statistically significant $\mathtt{ECT}_{t-1}$ term and its negative sign is also a sign that supports our conclusion for the existenceof cointegration. Its coefficient is equal to -0.225 which indicates that 22.5% of the disequilibrium is corrected each year or equivalently that the disequilibrium is fully corrected in about 4 years. Also from the Table 5.10 we see that the variables `edu`, `gdp` and `infl` are all statistically significant showing that they have a short-run impact. From the other table we can see that except from the intercept the only variable which has a long-run impact is the variable `infl`. Both the short and the long run components of `infl` are statistically significant and so we can suggest that inflation has a strong causal effect. On the other hand, the variables `edu` and `gdp` have a weak causal effect.

To sum up, the long-run effect of `infl` is rather due to the currency change from drachmas to euro and its effect is negative both in the short and the long run which means that as the inflation rises, the top 1% income share decreases. The variable `gdp` has also a negative short-run effect and the only variable with a positive short-run effect is `edu`.

Although our prime interest is in exploring and estimating the long-run relationship, we also present the graphical representation of the performance of the ARDL model. The Figure 5.10 shows how the fitted ARDL model performs against the observed dependent series and also shows the corresponding residuals. We observe that the fit is much better than the fit of the long-run relationship as this is the estimated series from the ARDL model which is a platform that incorporates both the short and the long-run dynamics.

But as we have explained in previous chapters, what we have done so far is investigating this model as if it was separate from other influences. In fact, the underlying DGP is a VAR and if the variable `tis01` takes part in any cointegrating relationship with any of the other variable, this effect should have passed in our estimated model. Although, we estimated the model separately from any other influences. At this point we have to apply the same modeling procedure using as dependent variable every variables we previously treated as independent.

The results for the variables `credit`, `edu`, `gdp`, `infl`, `popg` were the same, that no cointegrating relationship was found. Interestingly, the model treating the variable `trade` as dependent revealed that there is a cointegrating relationship under the `Case I` where both the F and the t test were significant. The selected model was the

$ARDL(2, 2, 0, 0, 1, 0, 0)$ with the corresponding variables `trade`, `tis01`, `credit`, `edu`, `gdp`, `infl` and `popg` respectively. The graphical long-run relationship also looks very good fitting[9].

After this last finding, any inference on the previous model for `tis01` (as we did) would be incorrect as the variable `trade` in the model for `tis01` is probably not weakly exogenous. The test for every variable should normally have been done beforehand but we wanted to make this example on how we should interpret the results if this was appropriate.

Table 5.5: ARDL model, Greece

	ARDL(1,0,0,3,0,0,0)			
	Dependent variable: tis01			
	Coefficient	Std.Error	t-Statistic	P-Value
Intercept	0.059	0.011	5.573	0.000
tis01_{t-1}	0.563	0.044	12.718	0.000
credit_t	2.530e−04	2.550e−05	9.936	0.000
edu_t	0.292	0.034	8.524	0.000
gdp_t	−1.480e−06	1.220e−07	−12.185	0.000
gdp_{t-1}	2.290e−06	3.090e−07	7.405	0.000
gdp_{t-2}	−5.410e−07	4.090e−07	−1.323	0.196
gdp_{t-3}	−3.190e−06	5.010e−07	−6.361	0.000
infl_t	−3.950e−04	1.350e−04	−2.932	0.006
popg_t	5.510e−04	0.001	0.516	0.609
trade_t	1.680e−04	7.370e−05	2.276	0.030
Observations	41		Residual Std. Error	0.003
R^2	0.936		Log Likelihood	184.743
Adjusted R^2	0.915		AIC	−8.475
F-Statistic	44.055		BIC	−8.016
Prob(F-Statistic)	0.000			

Note: HAC standard errors & covariance

[9] These model are not presented here but they are available upon request

Table 5.6: F Bounds test for cointegration, Greece

	Value	Significance	I(0)	I(1)
		Asymptotic: n=10000		
F−statistic	6.415	10%	1.99	2.94
k	6	5%	2.27	3.28
Actual Sample Size	41	2.5%	2.55	3.61
		1%	2.88	3.99
		Finite Sample: n=45		
		10%	2.188	3.254
		5%	2.591	3.766
		1%	3.540	4.931
		Finite Sample: n=40		
		10%	2.218	3.314
		5%	2.618	3.863
		1%	3.505	5.121

F-Bounds test H_0: No levels relationship

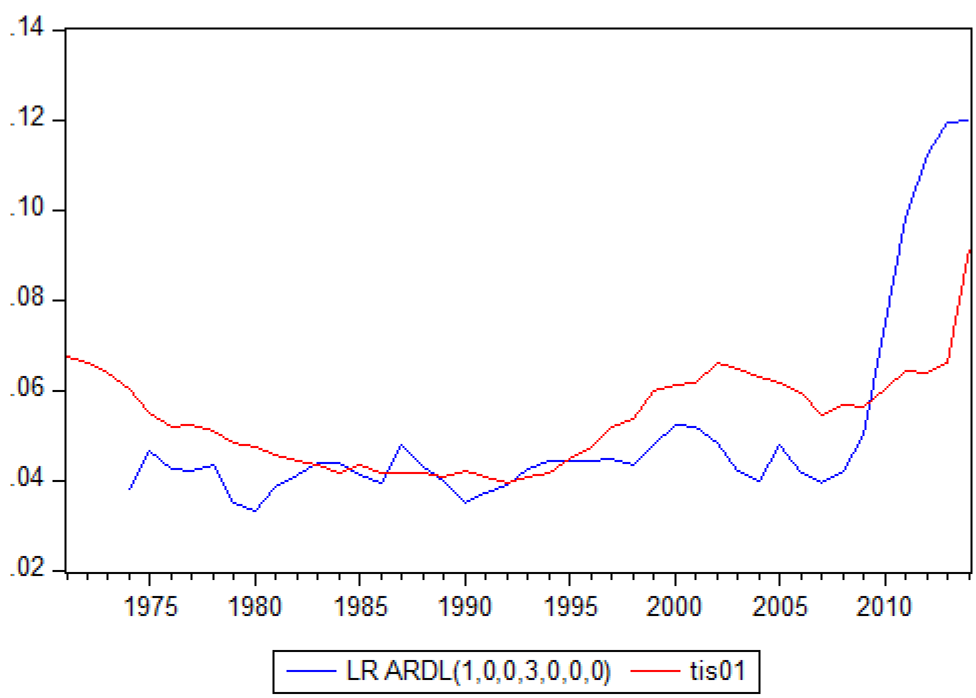

Figure 5.6: Greece, LR relationship, ARDL(1,0,0,3,0,0,0) Case II

Table 5.7: ARDL model with dummy, Greece

		ARDL(1,0,1,1,1,0,0)		
		Dependent variable: tis01		
	Coefficient	Std.Error	t-Statistic	P-Value
Intercept	0.018	0.004	4.751	0.000
tis01_{t-1}	0.775	0.040	19.376	0.000
credit_t	-2.18e-05	2.66e-05	-0.818	0.419
edu_t	0.267	0.120	2.226	0.033
edu_{t-1}	-0.355	0.147	-2.413	0.022
gdp_t	-1.19e-06	4.53e-07	-2.640	0.013
gdp_{t-1}	1.11e-06	4.91e-07	2.258	0.031
infl_t	-2.050e-04	8.58e-05	-2.393	0.023
infl_{t-1}	-2.750e-04	7.29e-05	-3.770	0.001
popg_t	-7.130e-04	0.001	-0.711	0.483
trade_t	1.070e-04	6.08e-05	1.757	0.089
D2014_t	0.024	0.002	11.249	0.000
Observations	43		Residual Std. Error	1.563e-03
R^2	0.984		Log Likelihood	223.853
Adjusted R^2	0.979		AIC	-9.854
F-Statistic	178.759		BIC	-9.362
Prob(F-Statistic)	0.000			

Table 5.8: F Bounds test for cointegration with dummy, Greece

	Value	Significance	I(0)	I(1)
		Asymptotic: n=10000		
F−statistic	7.831	10%	1.99	2.94
k	6	5%	2.27	3.28
Actual Sample Size	43	2.5%	2.55	3.61
		1%	2.88	3.99
		Finite Sample: n=45		
		10%	2.188	3.254
		5%	2.591	3.766
		1%	3.540	4.931
		Finite Sample: n=40		
		10%	2.218	3.314
		5%	2.618	3.863
		1%	3.505	5.121

F-Bounds test H_0: No levels relationship

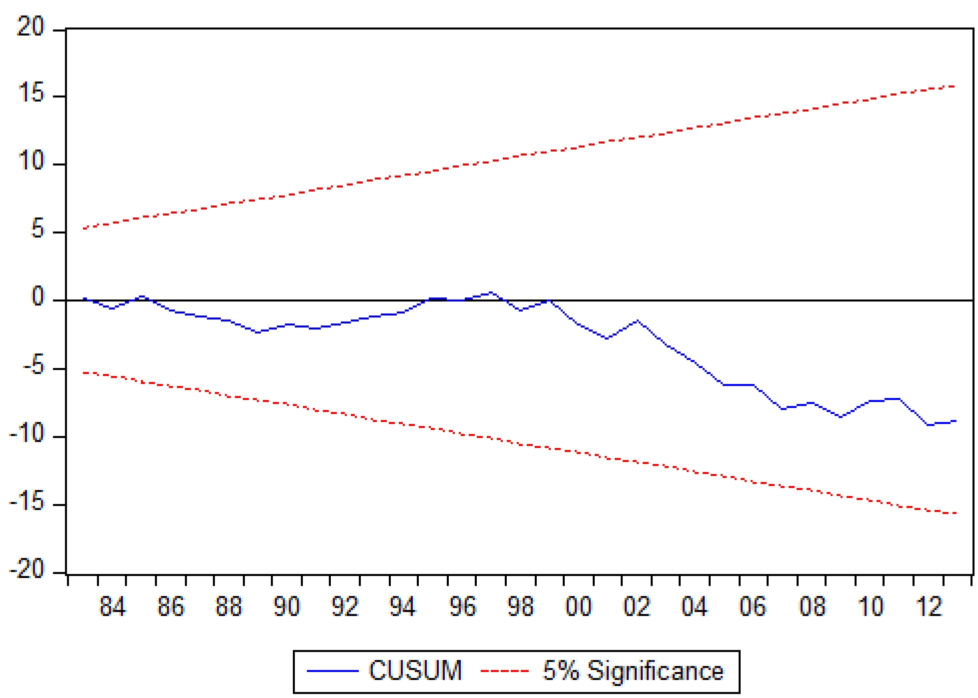

Figure 5.7: Greece, CUSUM test, ARDL(1,0,1,1,1,0,0) Case II

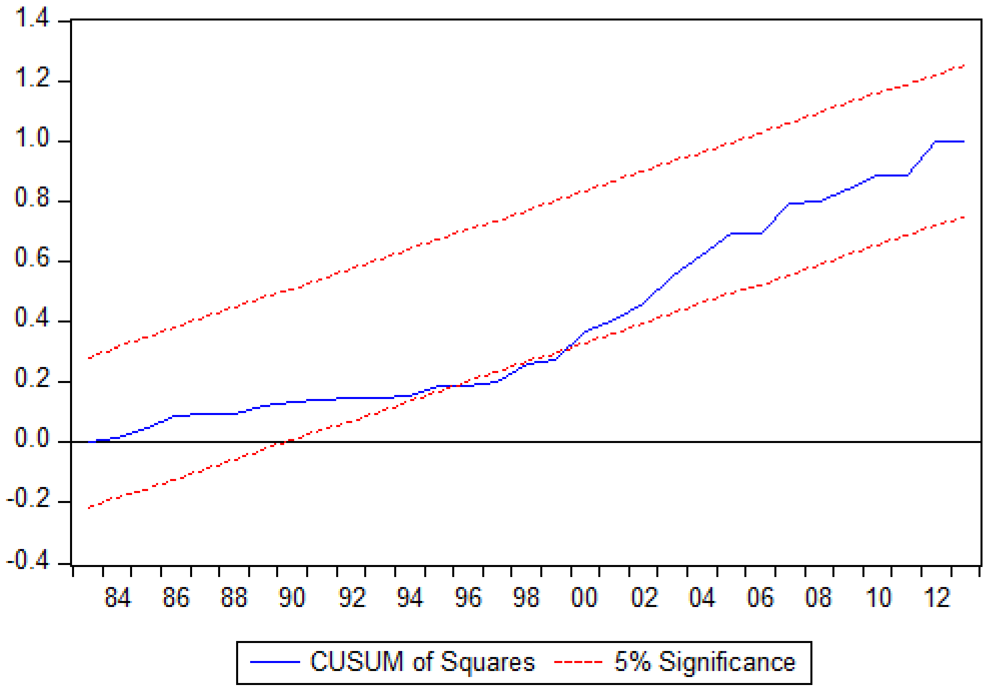

Figure 5.8: Greece, CUSUMSQ test, ARDL(1,0,1,1,1,0,0) Case II

Table 5.9: Levels Equation with dummy, Greece

	Dependent variable: tis01			
	Case 2: Restricted Constant and No Trend			
	Coefficient	Std.Error	t-Statistic	P-Value
Intercept	0.081	0.013	6.265	0.000
$credit_t$	-9.69e-05	1.250e-04	-0.775	0.444
edu_t	-0.391	0.335	-1.167	0.252
gdp_t	-3.81e-07	9.32e-07	-0.409	0.685
$infl_t$	-0.002	4.100e-04	-5.207	0.000
$popg_t$	-0.003	0.005	-0.705	0.486
$trade_t$	4.750e-04	2.790e-04	1.703	0.099

Table 5.10: ECM with dummy, Greece

	Coefficient	Std.Error	t-Statistic	P-Value
	Dependent variable: Δtis01			
	Case 2: Restricted Constant and No Trend			
Δedu_t	0.267	0.090	2.964	0.006
Δgdp_t	-1.19e-06	2.52e-07	-4.738	0.000
Δinfl_t	-2.050e-04	5.69e-05	-3.611	0.001
D2014_t	0.024	1.437e-03	16.368	0.000
ECT_{t-1}	-0.225	0.026	-8.763	0.000

Observations	43	Residual Std. Error	1.412e-03
R^2	0.915	Log Likelihood	223.853
Adjusted R^2	0.906	AIC	-10.179
BIC	-9.974		

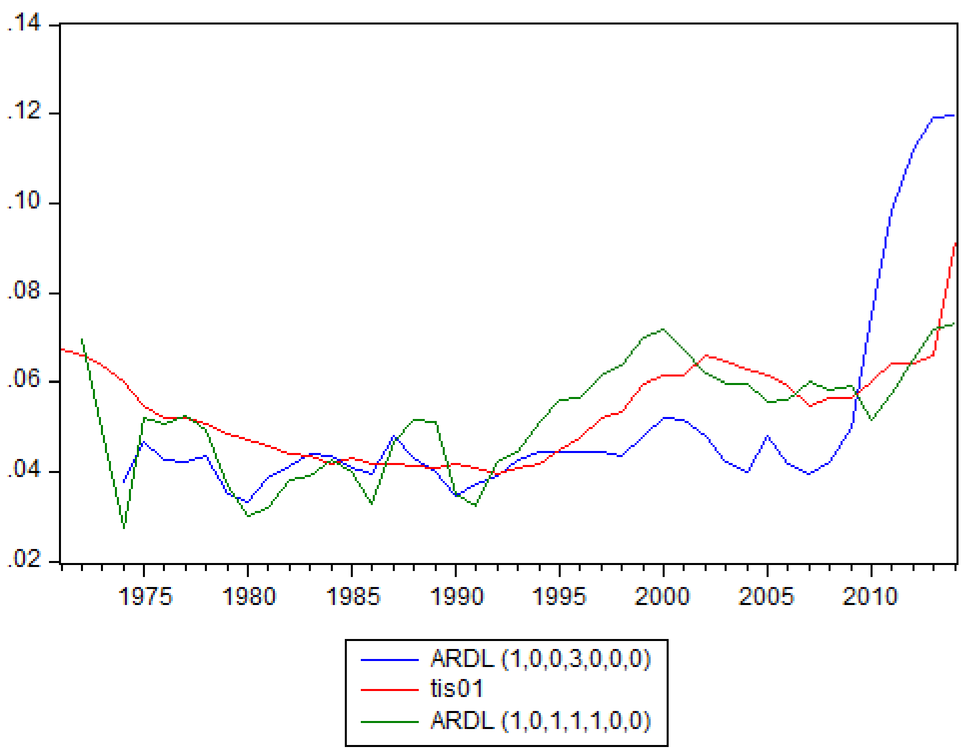

Figure 5.9: Greece, LR relationship, ARDL(1,0,1,1,1,0,0) Case II

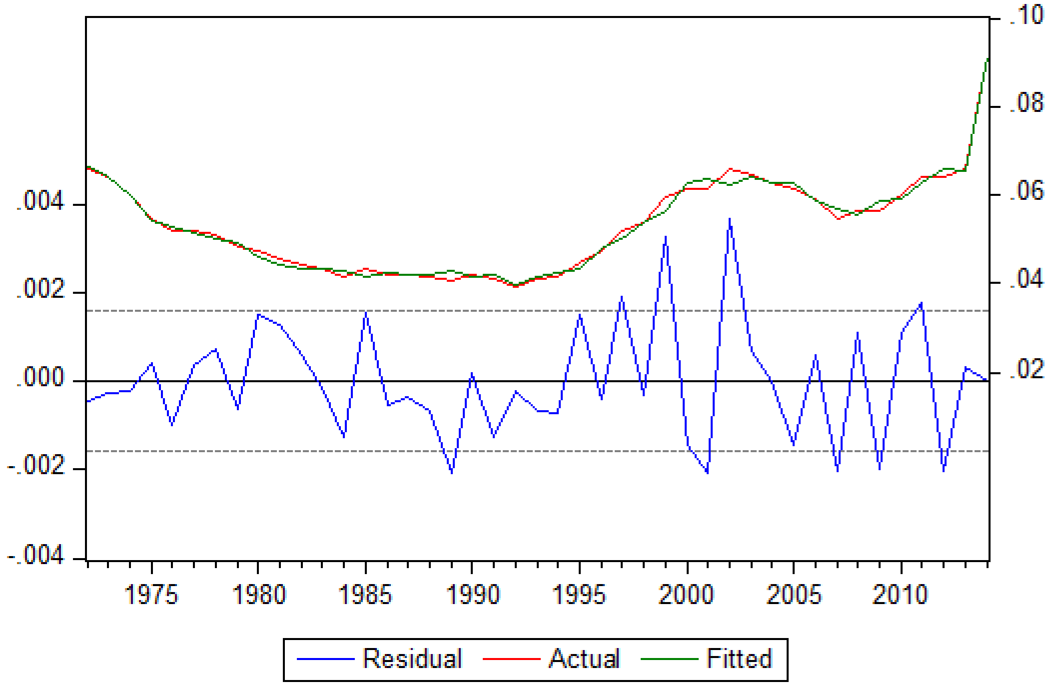

Figure 5.10: Greece, fit and error, ARDL(1,0,1,1,1,0,0) Case II

5.4 The case of France

From the Table 5.11 we see that the `credit` is stationary in levels and so it's $I(0)$. The standard ADF test gets confused because of the sudden drop. But assuming a deterministic trend (KPSS) or a structural break at the intercept of the slope at the year 1977 the conclusion is the former. The same happens also with `trade`. The variable `edu` is under the same situation. The linear deterministic trend here is more visible and the type of the break occurs more slowly. `infl` appears to be $I(0)$ under the hypothesis of a break at the intercept, `popg` seems to have Unit Root while the results for `gdp` are mixed. About the variable `tis01`, it clearly has a UR based on the tests.

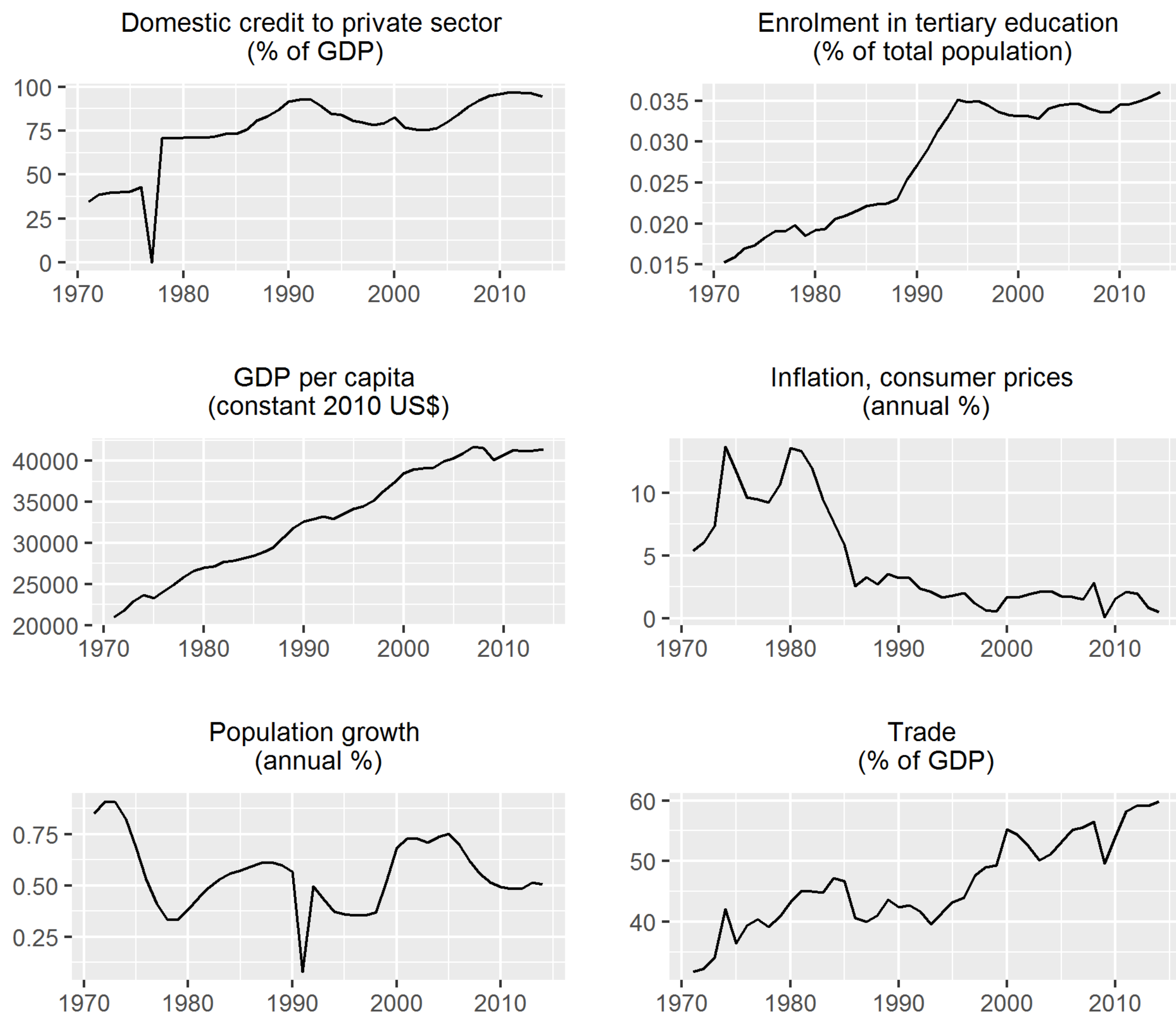

Figure 5.11: France, variables in levels

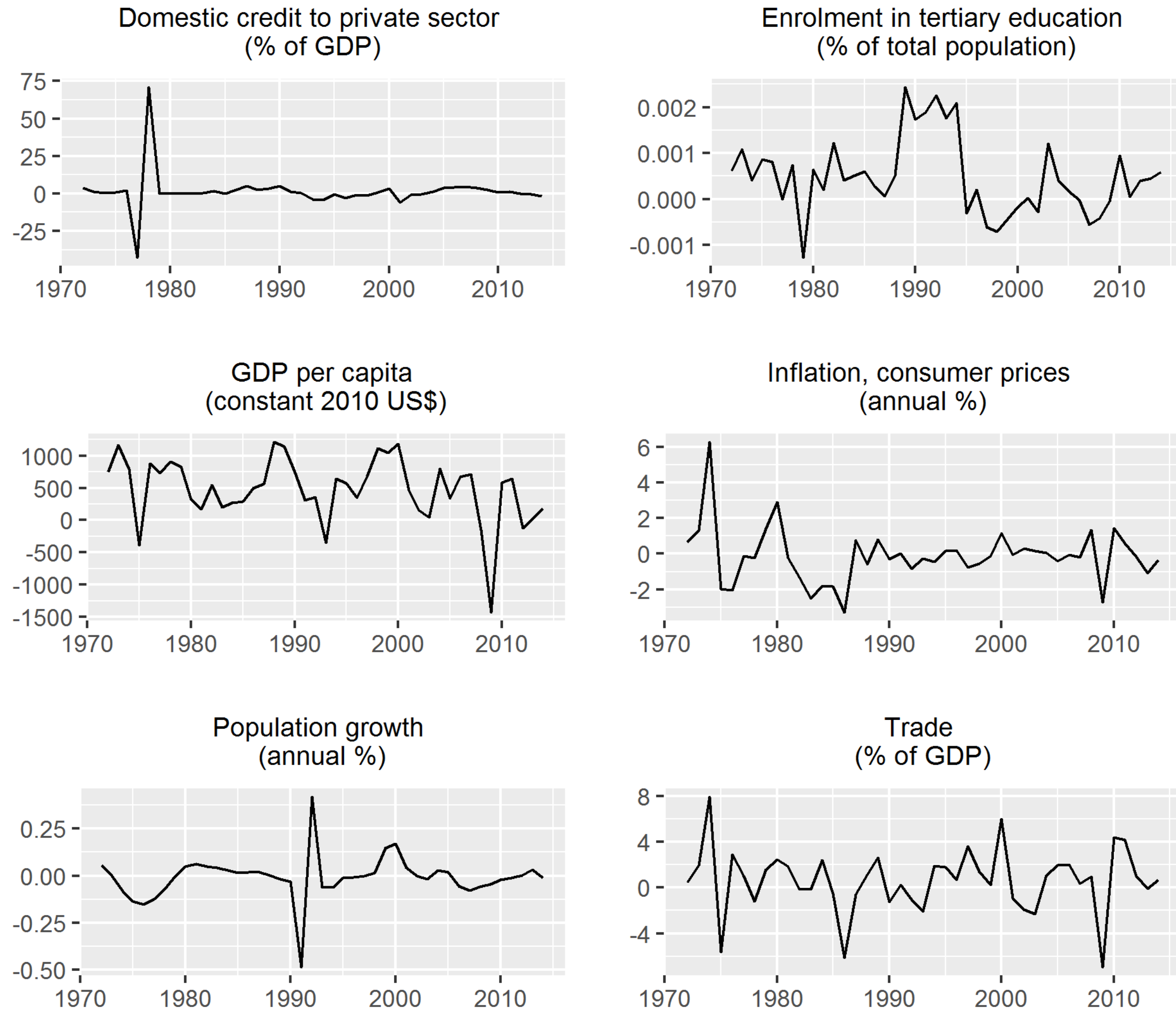

Figure 5.12: France, variables in first differences

The Table 5.12 provides very uniform results showing that all the variables are stationary at first differences and no variable is $I(2)$ or more. Once again, it's not necessary to conclude about whether a variable is $I(0)$ or $I(1)$ as long as it's not greater than $I(1)$.

Now we are set to go for the cointegration test. From the graphical representation of the `tis01` series for France in the Figure 5.2 and after experimenting with several models and cases, finally, the best model was found using the `Case II` scenario on a $ARDL(1,0,1,1,0,1,0)$ model specification. The model is presented in Table 5.13 and we can see from the Table C.7 that it passes all the diagnostic tests and so the model is well defined. Also, the Figures 5.13 and 5.14 are the graphical representations of the CUSUM and CUSUM Square tests and clearly show that the model parameters and variance are dynamically stable.

Now that all the requirements for the test are fulfilled we can take a look at the

Table 5.11: Unit Root tests for France in Levels

	France, variables in levels:											
	Standard Unit Root Tests											Breakpoint Unit Root Tests
	ADF	Exogenous	Lags	PP	Exogenous	KPSS	Exogenous	ADF_{bp}	Specification	Lags	Break Type	Break Date
tis01	0.213	n	0	0.195	n	0.601^{**}	c	−1.482	t-t	0	a	1982
credit	−1.952	c	1	−2.546	c	0.153	t	-9.578^{***}	t-c	0	a	1977
edu	−1.256	c	1	−1.302	c	0.144^{*}	t	-5.858^{***}	t-c	2	i	1988
gdp	−2.051	c	0	−1.913	c	$0.136*$	t	−4.226	t-b	1	i	2005
infl	−1.057	n	0	−1.106	n	0.618^{**}	c	-4.832^{**}	c-c	0	i	1984
popg	−1.199	n	0	−1.197	n	0.109	c	−3.458	c-c	1	i	1991
trade	−1.393	c	0	−1.233	c	0.105	t	-5.076^{**}	t-c	0	a	1985
H_0:	Unit Root			Unit Root		No Unit Root		Unit Root				

Note: *p<0.1; **p<0.05; ***p<0.01

For the Exogenous columns, c indicates that a constant intercept participates in the test equation,
t indicates the participation of a linear trend along with the constant and n indicates than no exogenous variables participate.
The column Lags shows the number of lags that was used (where applicable).
The number of lags was chosen based on the BIC criterion. The AIC was also consistent as for the stability of the test results
but it overestimated the number of lags in some cases.
The first letter on the Specification column indicates the trend nature of the data (therefore a level less than the column Exogenous).
The second letter indicates in which of these instances the break occurs.
Here t indicates a break in the treand, c a break in the intercept and b indicates a break in both of them.
In the Break Type column, a stands for Additive Outlier Test and i stands for Innovational Outlier Test.
The first one is used when we assume that the break occurs immediately while the second one when the break occurs slowly.
The breakpoint selection is mainly based on Dickey-Fuller min t-statistic and when appropriate on the max absolute t-statistics of the intercept/trend.

Table 5.12: Unit Root tests for France in First Differences

	France, variables in first differences:											
	Standard Unit Root Tests											Breakpoint Unit Root Tests
	ADF	Exogenous	Lags	PP	Exogenous	KPSS	Exogenous	ADF_{bp}	Specification	Lags	Break Type	Break Date
Δtis01	-6.221^{***}	n	0	-6.226^{***}	n	0.168	c	-6.860^{***}	c-c	0	i	2009
Δcredit	-10.044^{***}	n	0	-10.739^{***}	n	0.185	c	-12.804^{***}	c-c	0	i	1979
Δedu	-3.108^{***}	n	0	-3.040^{***}	n	0.159	c	-4.881^{**}	c-c	0	i	1994
Δgdp	-3.213^{***}	c	0	-3.034^{***}	n	0.406^{*}	c	-6.335^{***}	c-c	0	i	2009
Δinfl	-5.701^{***}	n	0	-5.672^{***}	n	0.104	c	-6.525^{***}	c-c	0	i	1976
Δpopg	-7.556^{***}	n	0	-7.518^{***}	n	0.109	c	-10.176^{***}	c-c	0	i	1991
Δtrade	-6.989^{***}	n	0	-7.049^{***}	n	0.093	c	-7.992^{***}	c-c	0	i	1986
H_0:	Unit Root			Unit Root		No Unit Root		Unit Root				

Note: *p<0.1; **p<0.05; ***p<0.01

For the Exogenous columns, c indicates that a constant intercept participates in the test equation,
t indicates the participation of a linear trend along with the constant and n indicates than no exogenous variables participate.
The column Lags shows the number of lags that was used (where applicable).
The number of lags was chosen based on the BIC criterion. The AIC was also consistent as for the stability of the test results
but it overestimated the number of lags in some cases.
The first letter on the Specification column indicates the trend nature of the data (therefore a level less than the column Exogenous).
The second letter indicates in which of these instances the break occurs.
Here t indicates a break in the treand, c a break in the intercept and b indicates a break in both of them.
In the Break Type column, a stands for Additive Outlier Test and i stands for Innovational Outlier Test.
The first one is used when we assume that the break occurs immediately while the second one when the break occurs slowly.
The breakpoint selection is mainly based on Dickey-Fuller min t-statistic and when appropriate on the max absolute t-statistics of the intercept/trend.

Table 5.14 and see that the F test statistic value is equal to 4.539 which is much greater even from the upper critical value for n=40[10] from the non asymptotic critical values provided by Narayan (2005), which indicates that there may be a cointegrating relationship between the independent variables and the dependent one. This hypothesis is also supported by the negative and statistically significant $\texttt{ECT}_{t-1}$ term in the Table 5.16 which also tells us that each year, the disequilibrium is reduced by 35.8% on average and so it takes less than three years to return back in the long-run equilibrium. The same table indicates that the variables `edu` and `gdp` have a short-run effect on `tis01` while the Table 5.15 shows that only the variable `popg` has a positive long-run effect on `tis01`.

Interestingly, we notice that `edu` and `gdp` have respectively a negative and a positive short-run effect in the case of France in contrast with the case of Greece where these variables had the opposite short-run effects.

We can also get a good feeling of the performance of our long-run relationship model from the Figure 5.15 where we see that the long-run trend is well captured except of the first few years until 1981 where there is a big divergence. The model also falsely predicts a weird drop for the year 1991 which is caused by the strange drop in the variable `popg` for the same year.

The whole performance of the ARDL model can be seen in the Figure 5.16 where the ARDL model seems to fit very well, much better than the long-run cointegrating relationship, but this is expected as the whole ARDL model includes also the short-run dynamics.

A quick try to apply the cointegration test using the rest of the variables as dependent revealed the following results[11]. The case with `credit` as dependent variable found to be cointegrating under the `Case I` using an $ARDL(1,2,1,0,3,0,3)$ model. The model for `edu` was also rejected the bounds test with the model $ARDL(1,3,0,1,3,1,3)$ under the `Case II`. The model for `gdp` had the same results using an $ARDL(1,0,0,0,2,0,0)$ under the `Case I` and also the model for `trade` under the same case with an $ARDL(1,0,0,0,1,0,0)$ model[12].

Clearly, the previous results for the model about `tis01` should not be interpreted in order to make inference about it because there are many (even one would also be a problem) endogenous variables that we should have accounted for their cointegrating relationship with `tis01`.

[10]in our case our sample size is n=43

[11]The order of the variables in the ARDL models has the following pattern. The first variable is always the dependent one, the second one is always `tis01`, the rest of the variables has the following sequence (excluding the one that is the dependent variable each time) `credit`, `edu`, `gdp`, `infl`, `popg`, `trade`.

[12]These models are not presented here but they are available upon request.

Table 5.13: ARDL model, France

	ARDL(1,0,1,1,0,1,0)			
	Dependent variable: tis01			
	Coefficient	Std.Error	t-Statistic	P-Value
Intercept	3.868e-03	7.107e-03	0.544	0.590
tis01_{t-1}	0.641	0.094	6.817	0.000
credit_t	1.480e-05	5.610e-05	0.264	0.793
edu_t	-1.134	0.726	-1.560	0.128
edu_{t-1}	1.598	0.744	2.147	0.039
gdp_t	6.200e-06	1.330e-06	4.673	0.000
gdp_{t-1}	-5.320e-06	1.090e-06	-4.895	0.000
infl_t	2.790e-04	3.360e-04	0.830	0.412
popg_t	2.958e-03	4.967e-03	0.595	0.555
popg_{t-1}	0.011	5.182e-03	2.201	0.035
trade_t	-4.510e-04	2.540e-04	-1.779	0.084
Observations	43		Residual Std. Error	0.003
R^2	0.965		Log Likelihood	191.949
Adjusted R^2	0.954		AIC	-8.416
F-Statistic	87.139		BIC	-7.966
Prob(F-Statistic)	0.000			

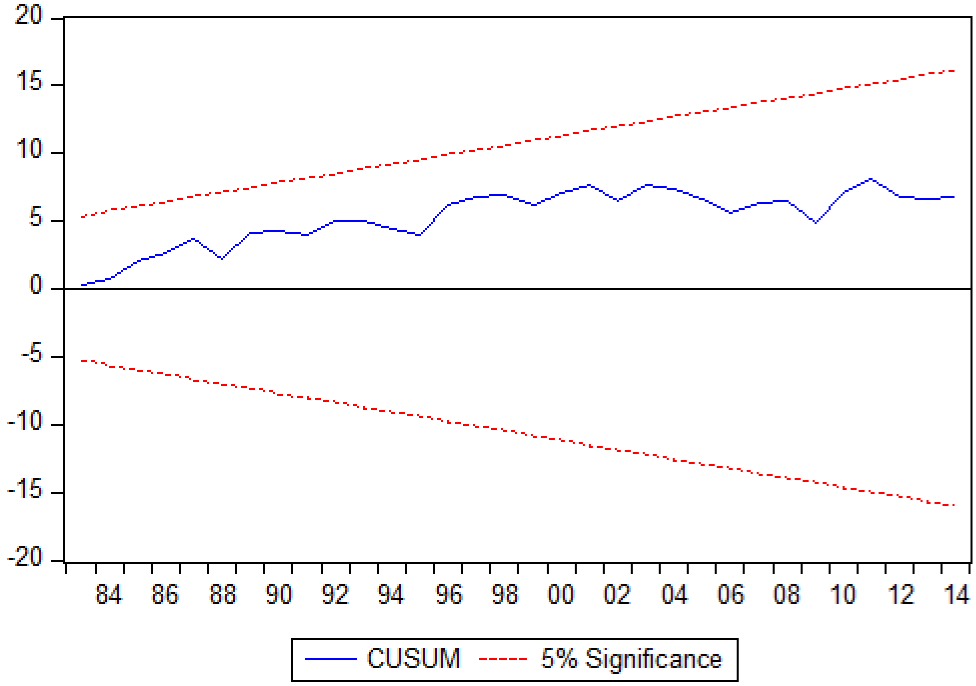

Figure 5.13: France, CUSUM test, ARDL(1,0,1,1,0,1,0) Case II

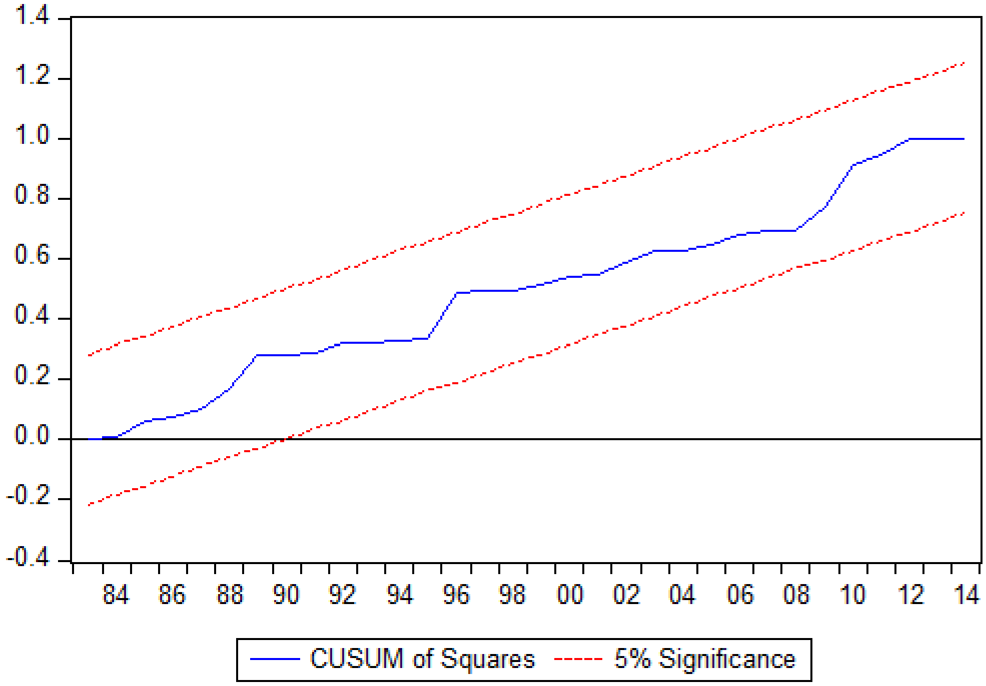

Figure 5.14: France, CUSUMSQ test, ARDL(1,0,1,1,0,1,0) Case II

Table 5.14: F Bounds test for cointegration, France

	Value	Significance	I(0)	I(1)
		Asymptotic: n=10000		
F-statistic	4.539	10%	1.99	2.94
k	6	5%	2.27	3.28
Actual Sample Size	43	2.5%	2.55	3.61
		1%	2.88	3.99
		Finite Sample: n=45		
		10%	2.188	3.254
		5%	2.591	3.766
		1%	3.540	4.931
		Finite Sample: n=40		
		10%	2.218	3.314
		5%	2.618	3.863
		1%	3.505	5.121

F-Bounds test H_0: No levels relationship

Table 5.15: Levels Equation, France

	Dependent variable: tis01 Case 2: Restricted Constant and No Trend			
	Coefficient	Std.Error	t-Statistic	P-Value
Intercept	0.010	0.019	0.551	0.585
credit_t	4.130e-05	1.620e-04	0.255	0.800
edu_t	1.294	0.895	1.445	0.158
gdp_t	2.470e-06	1.520e-06	1.619	0.115
infl_t	7.780e-04	8.730e-04	0.891	0.379
popg_t	0.040	0.012	3.111	0.000
trade_t	-1.258e-03	6.420e-04	-1.959	0.058

Table 5.16: ECM, France

	Dependent variable: Δtis01			
	Case 2: Restricted Constant and No Trend			
	Coefficient	Std.Error	t-Statistic	P-Value
Δedu_t	-1.134	0.505	-2.243	0.031
Δgdp_t	6.200e-06	8.280e-07	7.492	0.000
$\Delta popg_t$	2.958e-03	3.913e-03	0.755	0.455
ECT_{t-1}	-0.358	0.053	-6.652	0.000
Observations	43	Residual Std. Error	2.920e-03	
R^2	0.626	Log Likelihood	191.948	
Adjusted R^2	0.597	AIC	-8.742	
BIC	-8.578			

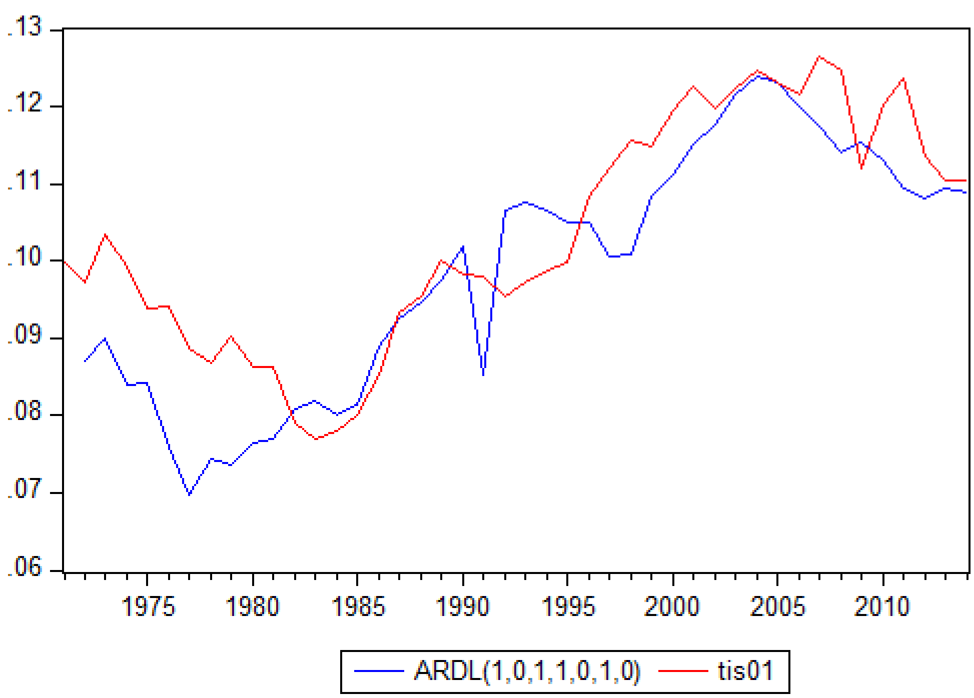

Figure 5.15: France, LR relationship, ARDL(1,0,1,1,0,1,0) Case II

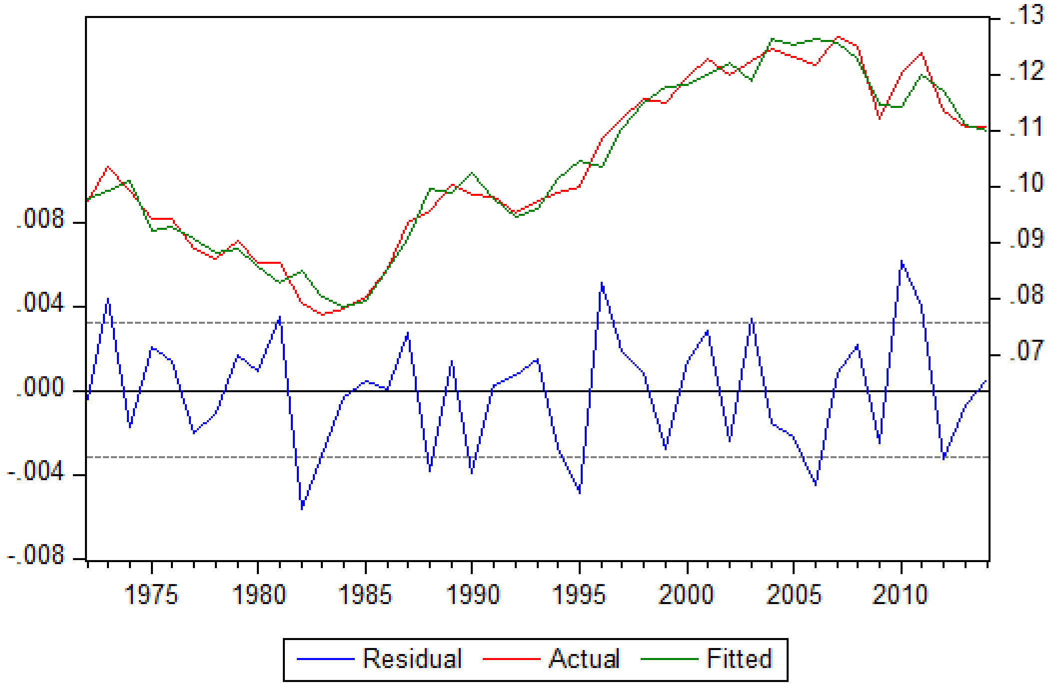

Figure 5.16: France, fit and error, ARDL(1,0,1,1,0,1,0) Case II

5.5 The case of USA

The results from Table 5.17 are mostly mixed among the tests. The only variable that all the tests agree in favor of the existenceof a Unit Root is `credit`. About the variable `tis01`, the AFD and the PP tests support the hypothesis of a Unit Root while the KPSS and the ADF breakpoint tests agree that there is not Unit Root. KPSS assumes a trend stationary process and the ADF breakpoint test also incorporates a deterministic linear trend including a break at the trend at the year 2011 where the `tis01` series seems to get flat. We continue believing the results of the KPSS test, that `tis01` is trend stationary and thus $I(0)$. A very interesting case would be the one that the ADF breakpoint test indicates but this implementation is not supported by the software we use as footnote 6 explains.

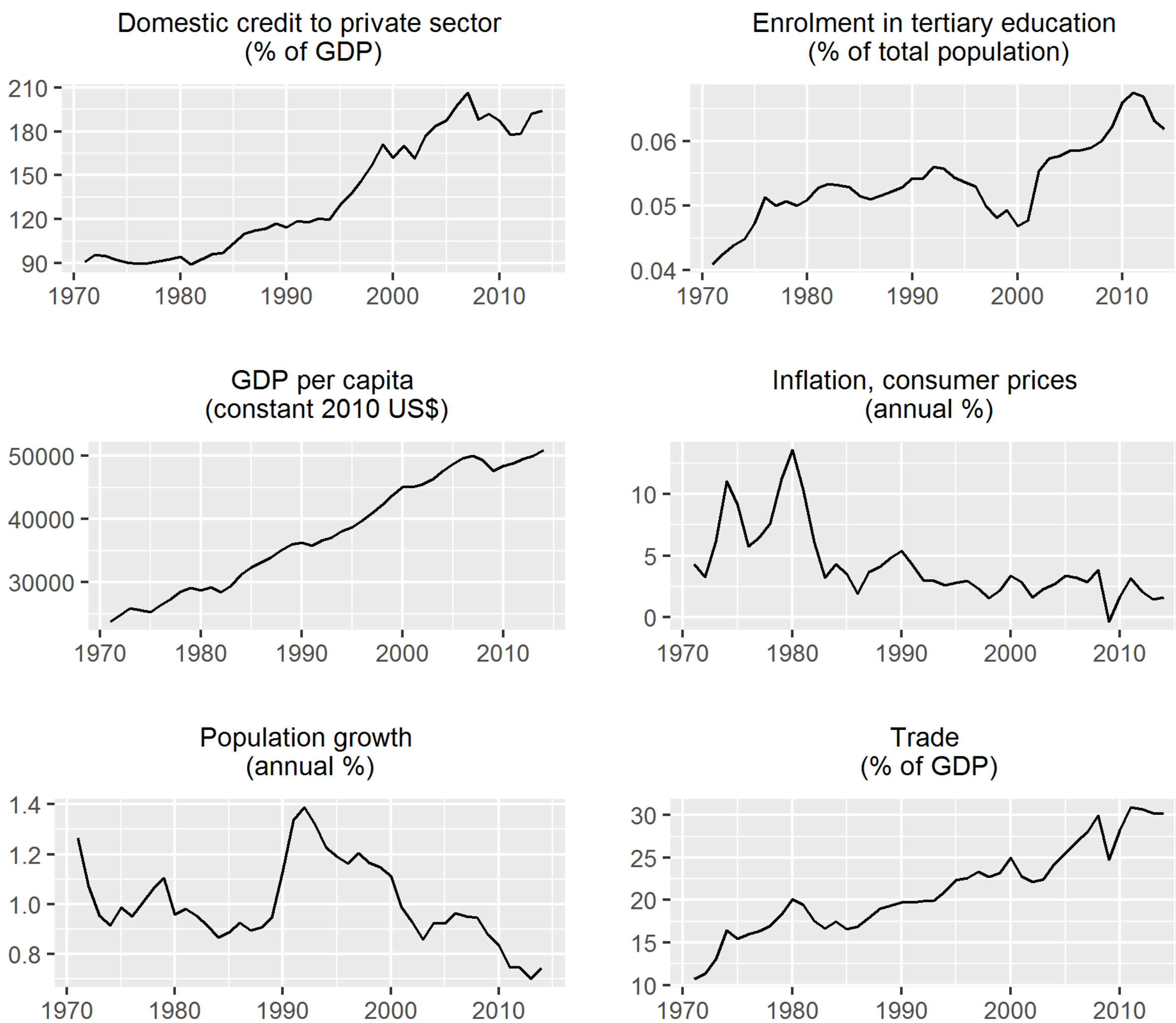

Figure 5.17: USA, variables in levels

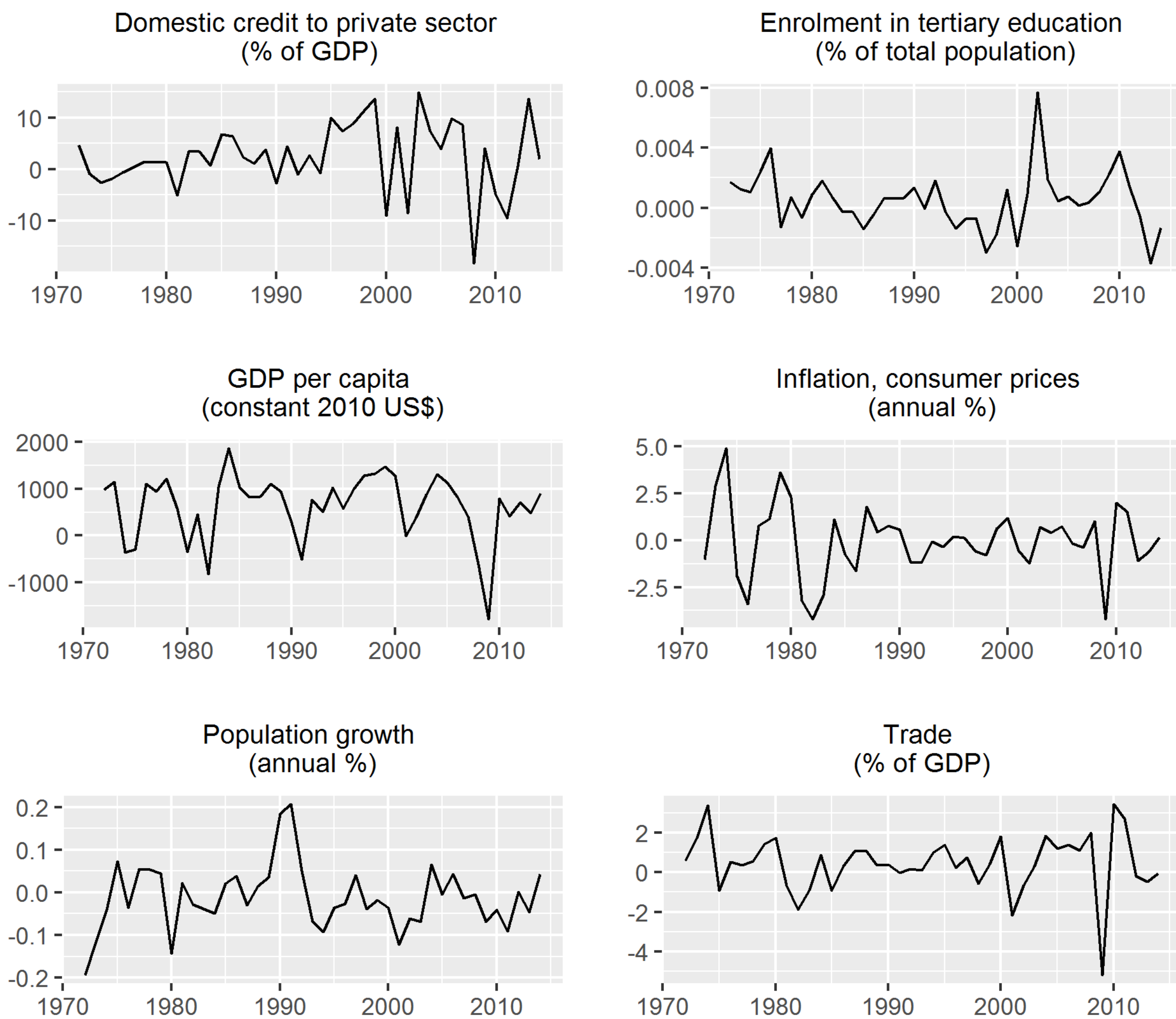

Figure 5.18: USA, variables in first differences

The Table 5.18 shows that all the independent variables in first differences are stationary. For the variable `tis01` in first differences, this time all the tests except the KPSS agree that it is stationary. According to the underlying assumptions that compose the requirements of the F-bounds test in order to be valid we should not continue applying the test as the dependent variable is trend stationary and so $I(0)$. Though, this case is a very interesting counterexample that in the end further supports our previous conclusion about the order of integration and although it doesn't give us any additional information about whether the series are mutually cointegrated, it gives us useful information about the behavior of the test under a case like this.

Along the modeling process many cases and order specifications were tested. Here we present the most representative ones for each case. Starting with `Case II` (using a restricted constant) and `Case III` (using an unrestricted constant) the best model for both cases found to be the $ARDL(1, 0, 0, 1, 0, 0, 0)$ but in both cases the F-bounds

Table 5.17: Unit Root tests for USA in Levels

	USA, variables in levels:											
	Standard Unit Root Tests											Breakpoint Unit Root Tests
	ADF	Exogenous	Lags	PP	Exogenous	KPSS	Exogenous	ADF_{bp}	Specification	Lags	Break Type	Break Date
tis01	−0.243	c	0	0.464	c	0.117	t	−4.402*	t-t	2	i	2011
credit	2.225	n	0	2.233	n	0.785***	t	−0.720	t-t	5	i	2004
edu	−1.879	c	1	−1.797	c	0.104	t	−4.343	t-b	1	i	1996
gdp	−0.463	c	1	−0.611	c	0.096	t	−3.792	t-c	1	i	2007
infl	−2.046	c	0	−2.017	c	0.093	t	−6.095***	c-c	1	a	1984
popg	−2.018	c	1	−1.855	c	0.141*	t	−7.322***	t-b	2	i	1989
trade	−1.264	c	0	−1.143	c	0.117	t	−4.918*	t-b	0	i	1981
H_0:	Unit Root			Unit Root		No Unit Root		Unit Root				

Note: *p<0.1; **p<0.05; ***p<0.01

For the Exogenous columns, c indicates that a constant intercept participates in the test equation,
t indicates the participation of a linear trend along with the constant and n indicates than no exogenous variables participate.
The column Lags shows the number of lags that was used (where applicable).
The number of lags was chosen based on the BIC criterion. The AIC was also consistent as for the stability of the test results
but it overestimated the number of lags in some cases.
The first letter on the Specification column indicates the trend nature of the data (therefore a level less than the column Exogenous).
The second letter indicates in which of these instances the break occurs.
Here t indicates a break in the treand, c a break in the intercept and b indicates a break in both of them.
In the Break Type column, a stands for Additive Outlier Test and i stands for Innovational Outlier Test.
The first one is used when we assume that the break occurs immediately while the second one when the break occurs slowly.
The breakpoint selection is mainly based on Dickey-Fuller min t-statistic and when appropriate on the max absolute t-statistics of the intercept/trend.

Table 5.18: Unit Root tests for USA in First Differences

	USA, variables in first differences:											
	Standard Unit Root Tests											Breakpoint Unit Root Tests
	ADF	Exogenous	Lags	PP	Exogenous	KPSS	Exogenous	ADF_{bp}	Specification	Lags	Break Type	Break Date
Δtis01	−6.613***	n	0	−6.612***	n	0.500**	c	−7.445***	c-c	0	i	1976
Δcredit	−6.091***	n	0	−6.175***	n	0.141	c	−7.305***	c-c	0	i	2007
Δedu	−4.394***	n	0	−4.419***	n	0.097	c	−5.473***	c-c	0	i	2002
Δgdp	−3.144***	n	0	−3.144***	n	0.085	c	−5.534***	c-c	0	i	2009
Δinfl	−5.482***	n	0	−6.728***	n	0.398*	c	−7.723***	c-c	1	a	1982
Δpopg	−4.697***	n	0	−4.640***	n	0.106	c	−5.805***	c-c	2	i	1991
Δtrade	−6.597***	n	0	−6.615***	n	0.149	c	−7.765***	c-c	0	i	2009
H_0:	Unit Root			Unit Root		No Unit Root		Unit Root				

Note: *p<0.1; **p<0.05; ***p<0.01

For the Exogenous columns, c indicates that a constant intercept participates in the test equation,
t indicates the participation of a linear trend along with the constant and n indicates than no exogenous variables participate.
The column Lags shows the number of lags that was used (where applicable).
The number of lags was chosen based on the BIC criterion. The AIC was also consistent as for the stability of the test results
but it overestimated the number of lags in some cases.
The first letter on the Specification column indicates the trend nature of the data (therefore a level less than the column Exogenous).
The second letter indicates in which of these instances the break occurs.
Here t indicates a break in the treand, c a break in the intercept and b indicates a break in both of them.
In the Break Type column, a stands for Additive Outlier Test and i stands for Innovational Outlier Test.
The first one is used when we assume that the break occurs immediately while the second one when the break occurs slowly.
The breakpoint selection is mainly based on Dickey-Fuller min t-statistic and when appropriate on the max absolute t-statistics of the intercept/trend.

tests, as shown in Tables 5.19 and 5.20, the F-statistic falls between the lower and the upper critical values and so the results are inconclusive[13].

But none of these cases seems a sound choice anyway as they don't incorporate the linear trend that we previously concluded that exists. So a better modeling approach would be the `Case IV` (using a restricted trend) which is presented in the Table 5.21. This is actually a very interesting case where the best order specification found to be the $ARDL(1,0,0,0,0,0,0)$. This means that no variable has a short-run effect on `tis01` and the whole series behavior is driven only through a single lag on the dependent variable which leaves the ARDL model with just an AR component and makes it a simple $ARX(1)$ which seems suspicious. If we were about to interpret the F-bounds test results in the Table 5.21 we would see that the F-statistic is greater than the upper bound and it would have rejected the Null hypothesis of no levels relationship concluding in favor of the existence of a cointegrating relationship.

Nonetheless, this would have been a crucial mistake. The reason is the one that was previously explained in the Chapter 4.5 at the requirements about the order of integration. Remember that even if we end up with statistically significant test results we are unable, using the test alone, to determine whether this is due to usual cointegration or we are under a degenerate case. In this case, we know in advance that the dependent variable is trend stationary and has a $I(0)$ process and we would normally expect the results of the F-bounds test, if properly specified, to be statistically significant but for the wrong reasons. In particular, as we have already explained in Chapter 4.5, we expect the first difference of the dependent variable in the conditional ECM model, to depend only on its own lagged levels and not on those of the independent (forcing) variables. This is exactly what our chosen model says through its $ARDL(1,0,0,0,0,0,0)$ specification. Moreover, a single look at the Figure 5.19 clearly suggests that this is a degenerate case.

[13]We have to note once again that since we assumed that the variable `tis01` $I(0)$ any results of the F-bounds tests are invalid and normally should not be interpreted. What we do here, is just an observation of how the test works under certain violated assumptions.

Table 5.19: F Bounds test for cointegration (Case II), USA

	Value	Significance	I(0)	I(1)
		Asymptotic: n=10000		
F-statistic	3.090	10%	1.99	2.94
k	6	5%	2.27	3.28
Actual Sample Size	43	2.5%	2.55	3.61
		1%	2.88	3.99
		Finite Sample: n=45		
		10%	2.188	3.254
		5%	2.591	3.766
		1%	3.540	4.931
		Finite Sample: n=40		
		10%	2.218	3.314
		5%	2.618	3.863
		1%	3.505	5.121

F-Bounds test H_0: No levels relationship
Corresponding model: ARDL(1,0,0,1,0,0,0)

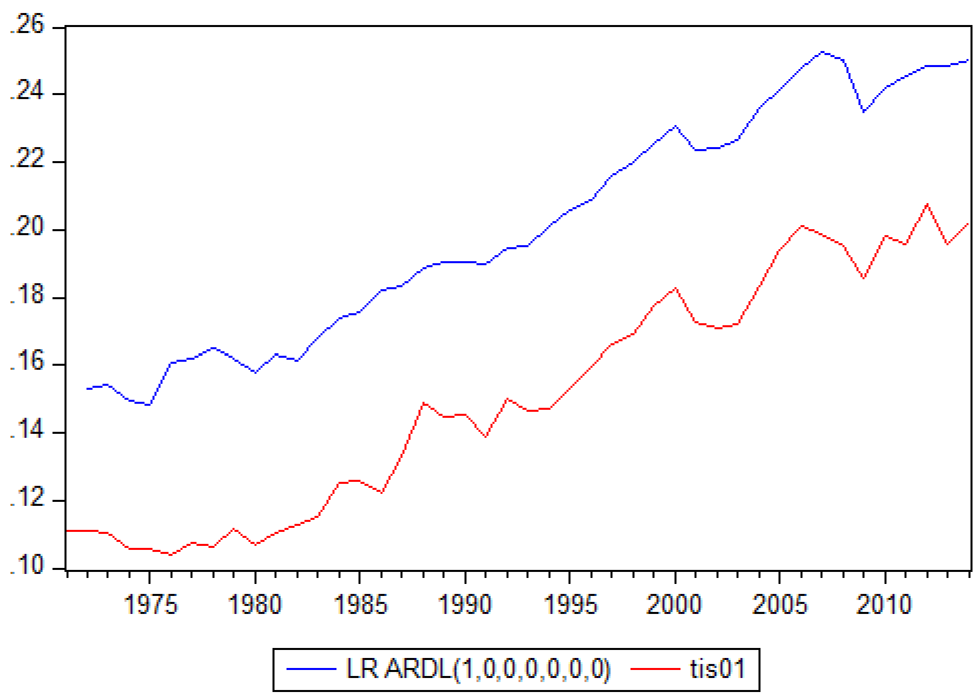

Figure 5.19: USA, LR relationship, ARDL(1,0,0,0,0,0,0) Case IV

Table 5.20: F Bounds and t Bounds test for cointegration (Case III), USA

	Value	Significance	I(0)	I(1)
		Asymptotic: n=10000		
F-statistic	3.307	10%	2.12	3.23
k	6	5%	2.45	3.61
Actual Sample Size	43	2.5%	2.75	3.99
		1%	3.15	4.43
		Finite Sample: n=45		
		10%	2.327	3.541
		5%	2.764	4.123
		1%	3.79	5.411
		Finite Sample: n=40		
		10%	2.353	3.599
		5%	2.797	4.211
		1%	3.8	5.643

	Value	Significance	I(0)	I(1)
t-statistic	-4.103	10%	-2.57	-4.04
		5%	-2.86	-4.38
		2.5%	-3.13	-4.66
		1%	-3.43	-4.99

F-Bounds test H_0: No levels relationship
t-Bounds test H_0: No levels relationship
Corresponding model: ARDL(1,0,0,1,0,0,0)

Table 5.21: F Bounds test for cointegration (Case IV), USA

	Value	Significance	I(0)	I(1)
		Asymptotic: n=10000		
F-statistic	5.459	10%	2.33	3.25
k	6	5%	2.63	3.62
Actual Sample Size	43	2.5%	2.9	3.94
		1%	3.27	4.39
		Finite Sample: n=45		
		10%	2.606	3.644
		5%	3.025	4.198
		1%	3.998	5.463
		Finite Sample: n=40		
		10%	2.634	3.719
		5%	3.07	4.309
		1%	4.154	5.699

F-Bounds test H_0: No levels relationship
Corresponding model: ARDL(1,0,0,0,0,0,0)

5.6 The case of United Kingdom

The results from the Table 5.22 show that the variables `infl`, `popg` and `trade` have a Unit Root, while the results are mixed for the rest. About the variable `tis01` for which we care the most in this table, it is hard to decide if it has a Unit Root or not.

The Table 5.23 show that all the variables are stationary in first differences. The independent variables are either $I(0)$ or $I(1)$ and the variable `tis01` is clearly stationary in first differences.

But we haven't take a decision yet about the variable `tis01`. The first option we have here is to trust the KPSS test and conclude that `tis01` is trend stationary, in other words $I(0)$. In this case, as we showed in the case of France, the test can't be applied. The second option is to assume that it has a Unit Root as the ADF and PP tests indicate.

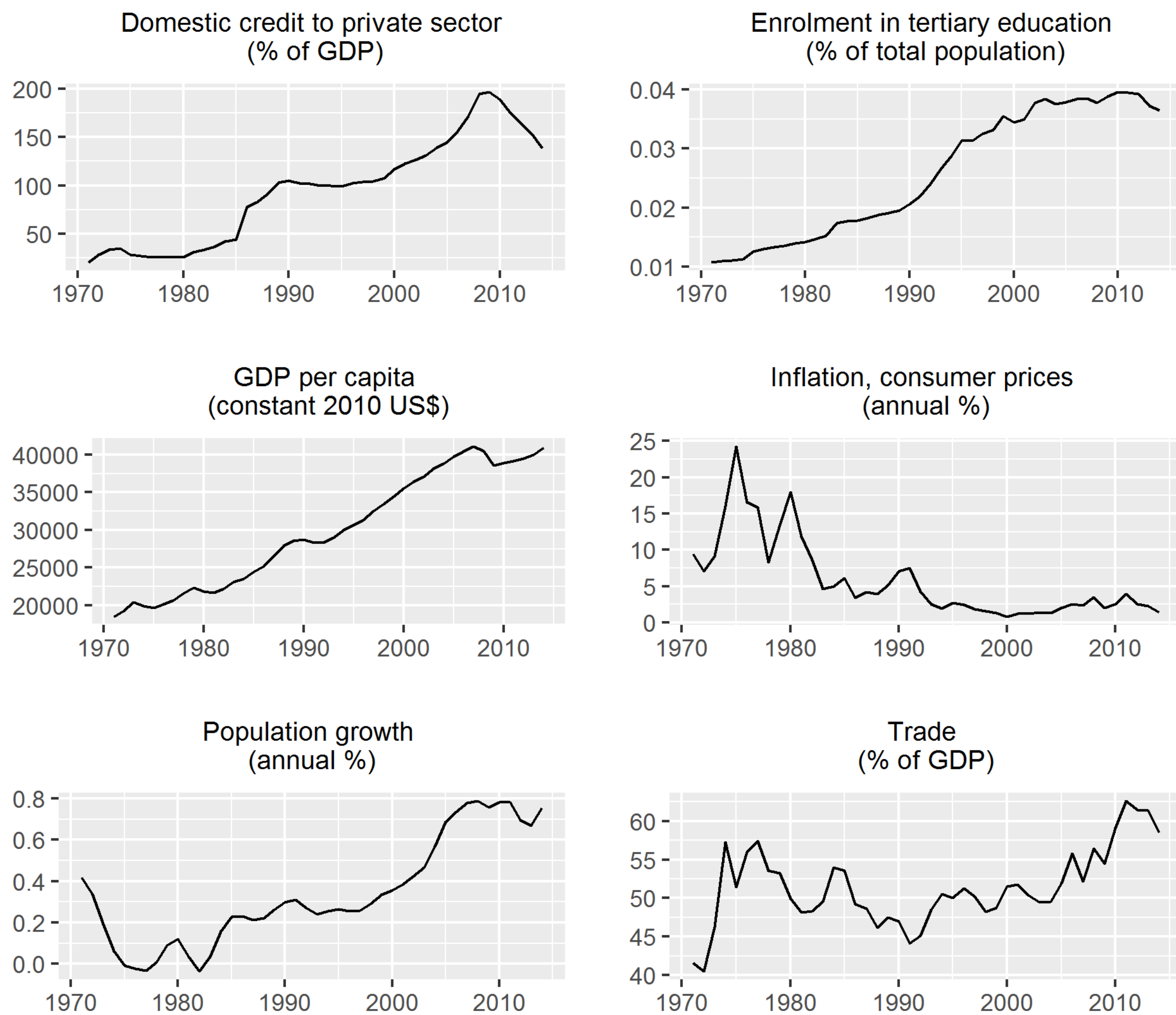

Figure 5.20: United Kingdom, variables in levels

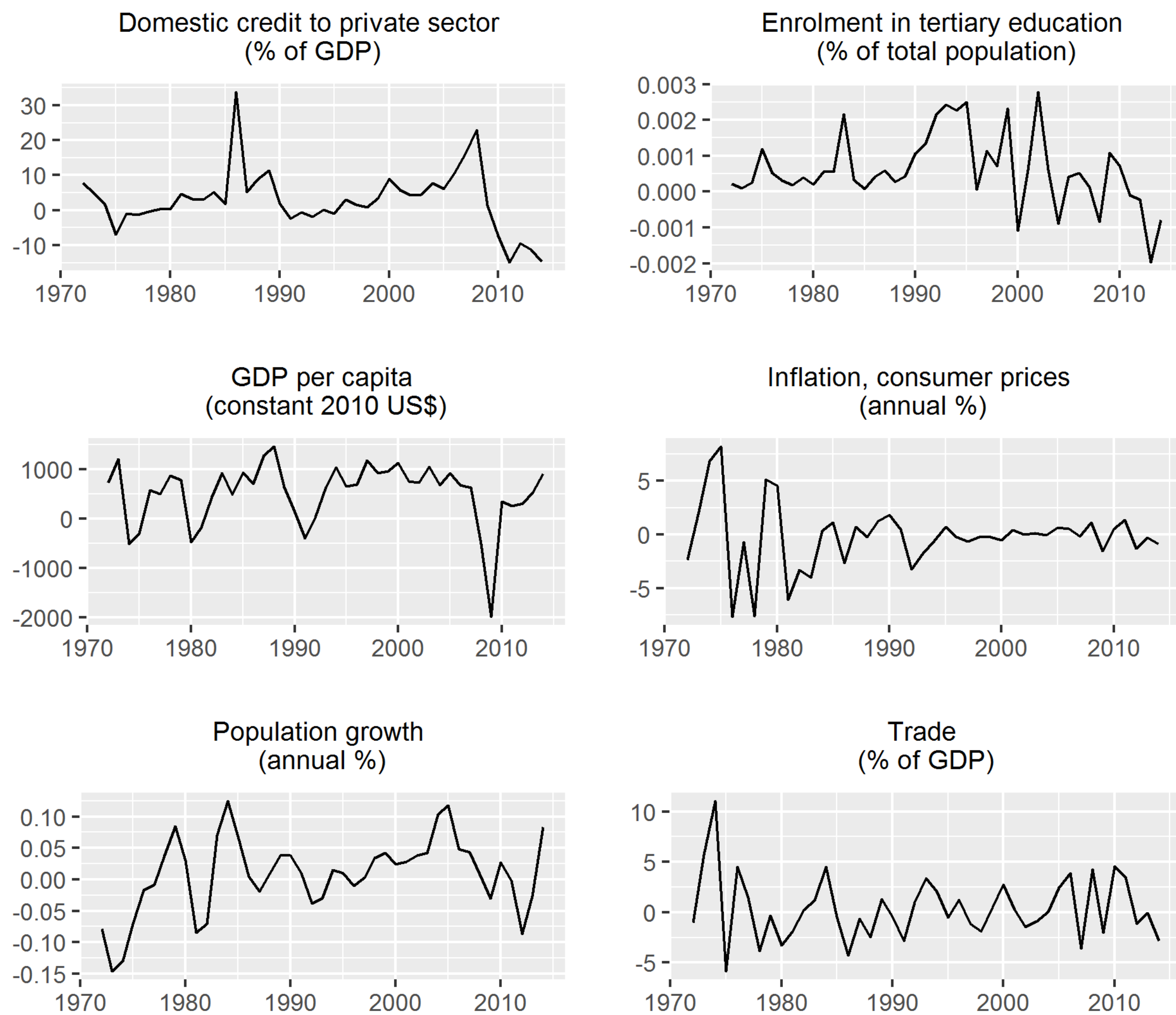

Figure 5.21: United Kingdom, variables in first differences

If this is true, we can continue doing the test using and appropriate case for the deterministic parts. In the opposite case, applying and interpreting the test would probably lead to wrong decisions as the test results would be invalid. Another try would be to model the deterministic parts accordingly as the series doesn't seem to have a linear trend throughout, rather than some declining parts in the very beginning and at the recent years but as explained this is not an option now.

Here we present three different cases, `Case I`, `Case II` and `Case IV` and for each one of those we include two models. The first one is the final model selected and the second one is best selected model only based on the BIC criterion without testing for potential autocorrelation and other problems. These models are all overparameterized and all have serial correlation problems. To understand the importance of this problem, we present the F-bounds and the t-bounds test results of these cases too. These results indicate in every case that there is a cointegrating relationship as the F-statistics

Table 5.22: Unit Root tests for United Kingdom in Levels

	United Kingdom, variables in levels:											
	Standard Unit Root Tests											Breakpoint Unit Root Tests
	ADF	Exogenous	Lags	PP	Exogenous	KPSS	Exogenous	ADF_{bp}	Specification	Lags	Break Type	Break Date
tis01	−0.559	c	0	−0.497	c	0.118	t	−3.239	t-t	0	a	2008
credit	−1.259	c	1	−1.171	c	0.055	t	−3.664	t-t	2	a	2006
edu	−1.136	c	0	−1.042	c	0.114	t	−2.361	t-c	0	i	1991
gdp	−0.405	c	1	−0.493	c	0.098	t	−3.406	t-t	1	i	2006
infl	−1.816	c	0	−1.583	c	0.162^{**}	t	-0.078^{*}	t-c	0	a	1980
popg	−0.867	c	5	−0.420	c	0.162^{**}	t	−2.764	c-c	7	i	2000
trade	−2.466	c	0	−2.536	c	0.154^{**}	t	−3.983	t-t	0	i	1992
H_0:	Unit Root			Unit Root		No Unit Root		Unit Root				

Note: *p<0.1; **p<0.05; ***p<0.01

For the Exogenous columns, c indicates that a constant intercept participates in the test equation,
t indicates the participation of a linear trend along with the constant and n indicates than no exogenous variables participate.
The column Lags shows the number of lags that was used (where applicable).
The number of lags was chosen based on the BIC criterion. The AIC was also consistent as for the stability of the test results
but it overestimated the number of lags in some cases.
The first letter on the Specification column indicates the trend nature of the data (therefore a level less than the column Exogenous).
The second letter indicates in which of these instances the break occurs.
Here t indicates a break in the treand, c a break in the intercept and b indicates a break in both of them.
In the Break Type column, a stands for Additive Outlier Test and i stands for Innovational Outlier Test.
The first one is used when we assume that the break occurs immediately while the second one when the break occurs slowly.
The breakpoint selection is mainly based on Dickey-Fuller min t-statistic and when appropriate on the max absolute t-statistics of the intercept/trend.

Table 5.23: Unit Root tests for United Kingdom in First Differences

	United Kingdom, variables in first differences:											
	Standard Unit Root Tests											Breakpoint Unit Root Tests
	ADF	Exogenous	Lags	PP	Exogenous	KPSS	Exogenous	ADF_{bp}	Specification	Lags	Break Type	Break Date
Δtis01	-6.892^{***}	n	0	-6.890^{***}	n	0.065	c	-7.834^{***}	c-c	0	i	1978
Δcredit	-3.246^{***}	n	0	-3.301^{***}	n	0.112	c	-5.318^{***}	c-c	0	i	2008
Δedu	-4.407^{***}	c	0	-4.367^{***}	c	0.268	c	-5.494^{***}	c-c	0	i	2002
Δgdp	-2.891^{***}	n	0	-2.891^{***}	n	0.108	c	-5.176^{***}	c-c	0	i	2009
Δinfl	-6.405^{***}	n	0	-7.761^{***}	n	0.488^{**}	c	-9.527^{***}	c-c	3	i	1986
Δpopg	-2.664^{***}	n	5	-2.732^{***}	n	0.338	c	-5.796^{***}	c-c	1	a	1998
Δtrade	-6.892^{***}	n	0	-6.890^{***}	n	0.065	c	-7.834^{***}	c-c	0	i	1978
H_0:	Unit Root			Unit Root		No Unit Root		Unit Root				

Note: *p<0.1; **p<0.05; ***p<0.01

For the Exogenous columns, c indicates that a constant intercept participates in the test equation,
t indicates the participation of a linear trend along with the constant and n indicates than no exogenous variables participate.
The column Lags shows the number of lags that was used (where applicable).
The number of lags was chosen based on the BIC criterion. The AIC was also consistent as for the stability of the test results
but it overestimated the number of lags in some cases.
The first letter on the Specification column indicates the trend nature of the data (therefore a level less than the column Exogenous).
The second letter indicates in which of these instances the break occurs.
Here t indicates a break in the treand, c a break in the intercept and b indicates a break in both of them.
In the Break Type column, a stands for Additive Outlier Test and i stands for Innovational Outlier Test.
The first one is used when we assume that the break occurs immediately while the second one when the break occurs slowly.
The breakpoint selection is mainly based on Dickey-Fuller min t-statistic and when appropriate on the max absolute t-statistics of the intercept/trend.

and the t-statistic in `Case I` are by far greater than any critical value regardless of the selected case which already seems strange. On the other hand, the well specified models doesn't have serial correlation problems.

When we are using the `Case I`, it's hard to check where the F-statistic lies, even using the critical values provided by Narayan (2005) because these values are reported for small sample sizes starting with 30 observations in increments of 5 and now we are in between as we see in the Table 5.24. For cases like this, we have constructed and carefully tested an algorithm written in R which simulates exact sample critical values under any scenario[14]. The lower and upper critical values for the exact sample size of n=43, k=6 and 5% significance were calculated as 2.30 and 3.66 respectively. The F-statistic is equal to 3.649 and so the test is inconclusive but for the `Case I` we can also apply the t-bounds test which is also inconclusive as the t-statistic falls between the lower and the upper critical values. Both the F and the t bounds tests are unable to reject the Null hypothesis for 5% level of significance but in both cases the statistic is very close to the critical value and so the test results are not very clear. We can see in the Figure 5.22 how the well specified model captures the fluctuations like the one in the 2000s better than the overparameterized model that is closer along the upward trend but it seems more like a straight line that becomes nearly horizontal at the 2000s and captures the mean rather than the fluctuating behavior.

The same happens with the `Case II`, where the F statistic in the Table 5.25 falls clearly between the two critical values and the Figure 5.23 has almost the same shape as the one in the previous figure. These two models are rejected for one more reason. The Figure 5.24 represents the CUSUM Square test that shows that the model parameters are not stable over time as there is a break starting in 1991 and stays off for a long time until 2010. This figure is almost identical for both the `Case I` and `Case II`.

Table 5.26 refers to the `Case IV` where this is a good example for us to show why the asymptotic critical values may be a simple solution but not always a good one, especially when our sample size is small. In this case, based on the asymptotic critical values we should reject the Null hypothesis of no levels relationship. But looking at the actual sample critical values by Narayan (2005) doesn't help much because our sample size is between the reported ones. Using again our simulation algorithm, we calculated the appropriate critical values for this case as 3.053 and 4.269 for the lower and upper limits respectively. According to this, we can't reject the Null hypothesis and this is another inconclusive case. The Figure 5.25 also supports this statement as both of the models diverge from `tis01`.

[14]The R codes are currently part of an R package under development, which also provides some useful functionalities that are missing from other existing software. The package will be published in the near future but the R codes that provide the critical values are available upon request.

Table 5.24: F Bounds and t Bounds test for cointegration (Case I), United Kingdom

	Value	Significance	I(0)	I(1)
		Asymptotic: n=10000		
Model	ARDL(1,0,0,1,0,0,0)	10%	1.75	2.87
F-statistic	3.649	5%	2.04	3.24
k	6	2.5%	2.32	3.59
Actual Sample Size	43	1%	2.66	4.05
		Finite Sample: n=45		
Model	ARDL(3,0,1,1,0,0,2)	10%	1.89	3.10
F-statistic	8.800	5%	2.29	3.64
k	6	1%	3.14	4.81
Actual Sample Size	41			
		Finite Sample: n=40		
		10%	1.92	3.17
		5%	2.32	3.70
		1%	3.32	5.01

	Value	Significance	I(0)	I(1)
Model	ARDL(1,0,0,1,0,0,0)	10%	-1.62	-3.70
t-statistic	-3.966	5%	-1.95	-4.04
		2.5%	-2.24	-4.34
Model	ARDL(3,0,1,1,0,0,2)	1%	-2.58	-4.67
t-statistic	-6.147			

F-Bounds test H_0: No levels relationship
t-Bounds test H_0: No levels relationship

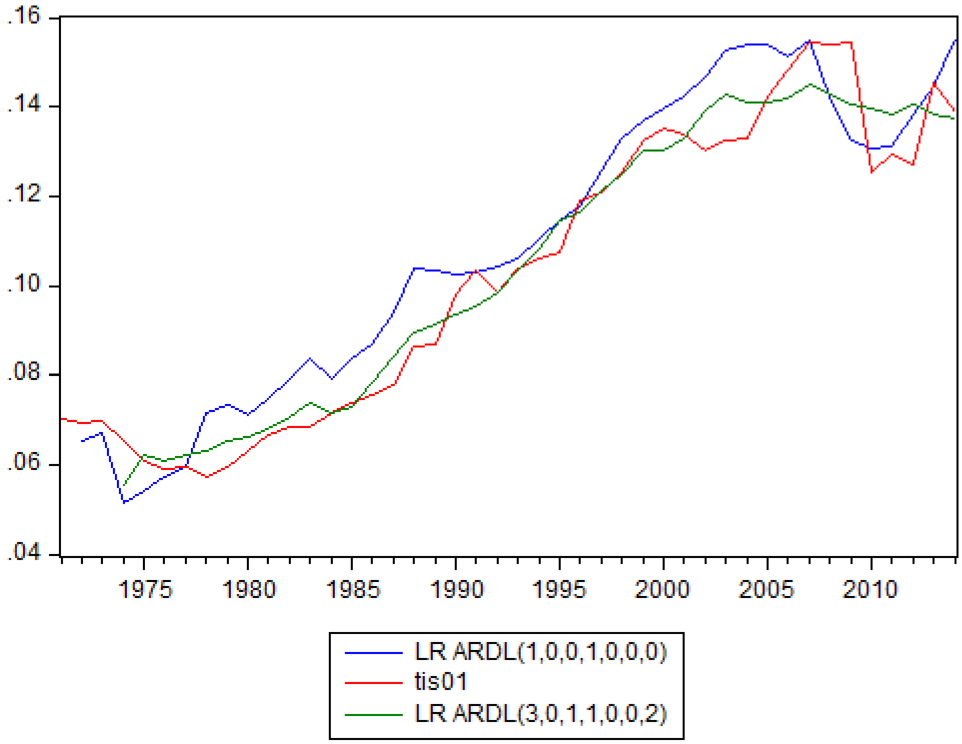

Figure 5.22: United Kingdom, LR relationships, Case I

Table 5.25: F Bounds test for cointegration (Case II), United Kingdom

	Value	Significance	I(0)	I(1)
		Asymptotic: n=10000		
Model	ARDL(1,0,0,1,0,0,0)	10%	1.99	2.94
F-statistic	3.166	5%	2.27	3.28
k	6	2.5%	2.55	3.61
Actual Sample Size	43	1%	2.88	3.99
		Finite Sample: n=45		
Model	ARDL(3,0,0,1,0,0,2)	10%	2.19	3.25
F-statistic	7.859	5%	2.59	3.77
k	6	1%	3.54	4.93
Actual Sample Size	41			
		Finite Sample: n=40		
		10%	2.22	3.31
		5%	2.62	3.86
		1%	3.51	5.12

F-Bounds test H_0: No levels relationship

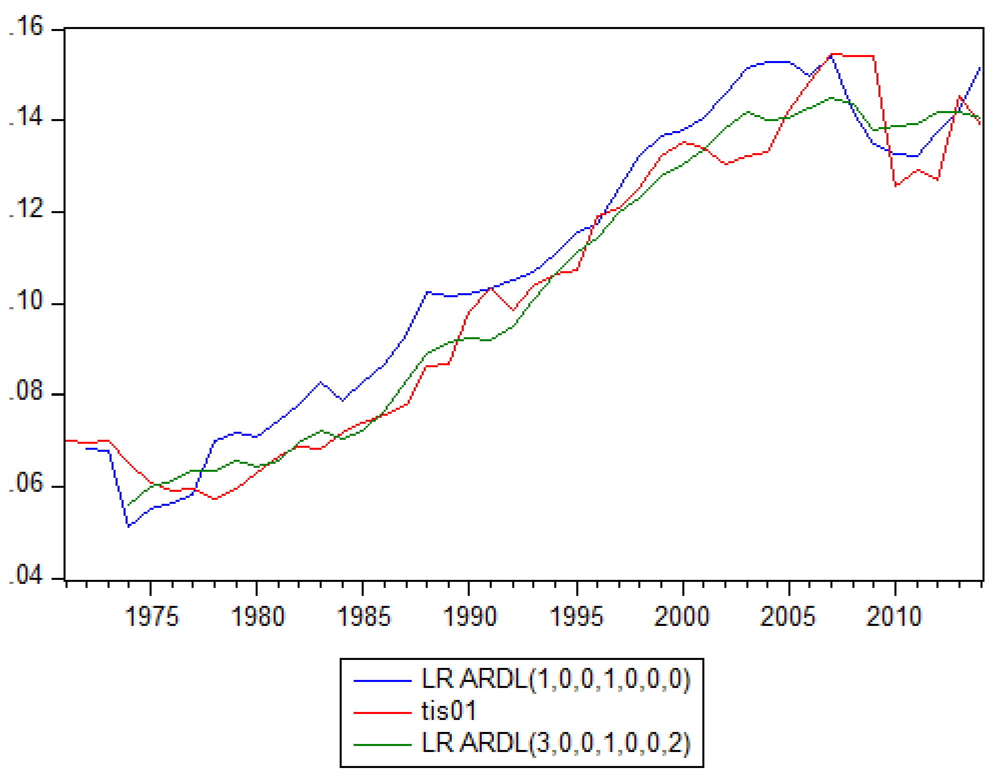

Figure 5.23: United Kingdom, LR relationships, Case II

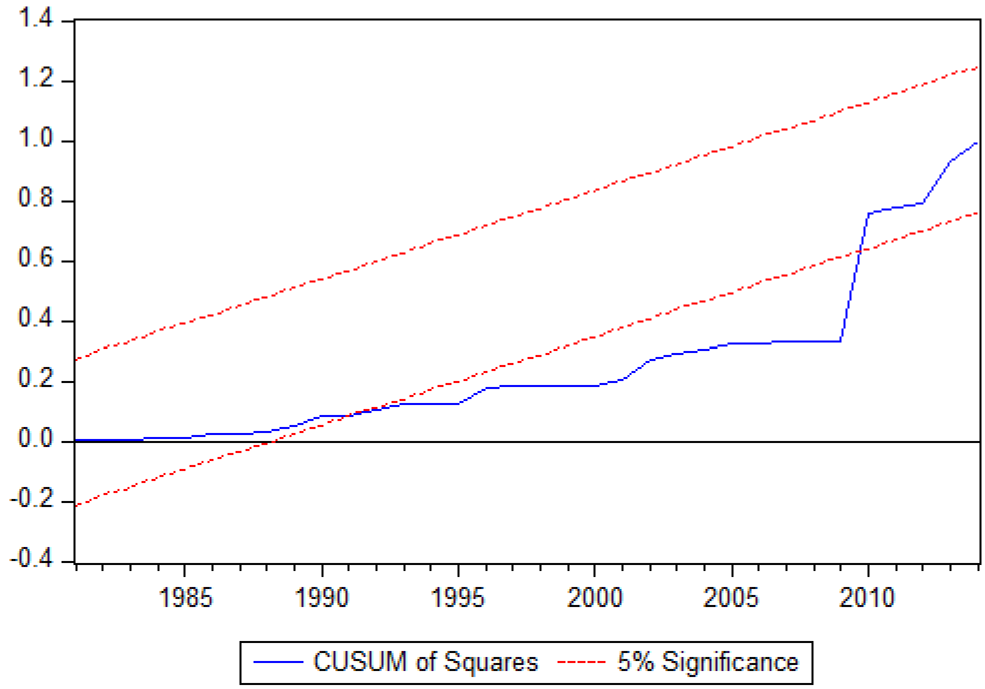

Figure 5.24: United Kingdom, CUSUMSQ test, Case I and Case II

Table 5.26: F Bounds test for cointegration (Case IV), United Kingdom

	Value	Significance	I(0)	I(1)
		Asymptotic: n=10000		
Model	ARDL(1,0,0,1,0,1,0)	10%	2.33	3.25
F-statistic	4.221	5%	2.63	3.62
k	6	2.5%	2.90	3.94
Actual Sample Size	43	1%	3.27	4.39
		Finite Sample: n=45		
Model	ARDL(3,0,0,1,0,0,0)	10%	2.61	3.64
F-statistic	6.807	5%	3.03	4.20
k	6	1%	4.00	5.46
Actual Sample Size	41			
		Finite Sample: n=40		
		10%	2.63	3.72
		5%	3.07	4.31
		1%	4.15	5.70

F-Bounds test H_0: No levels relationship

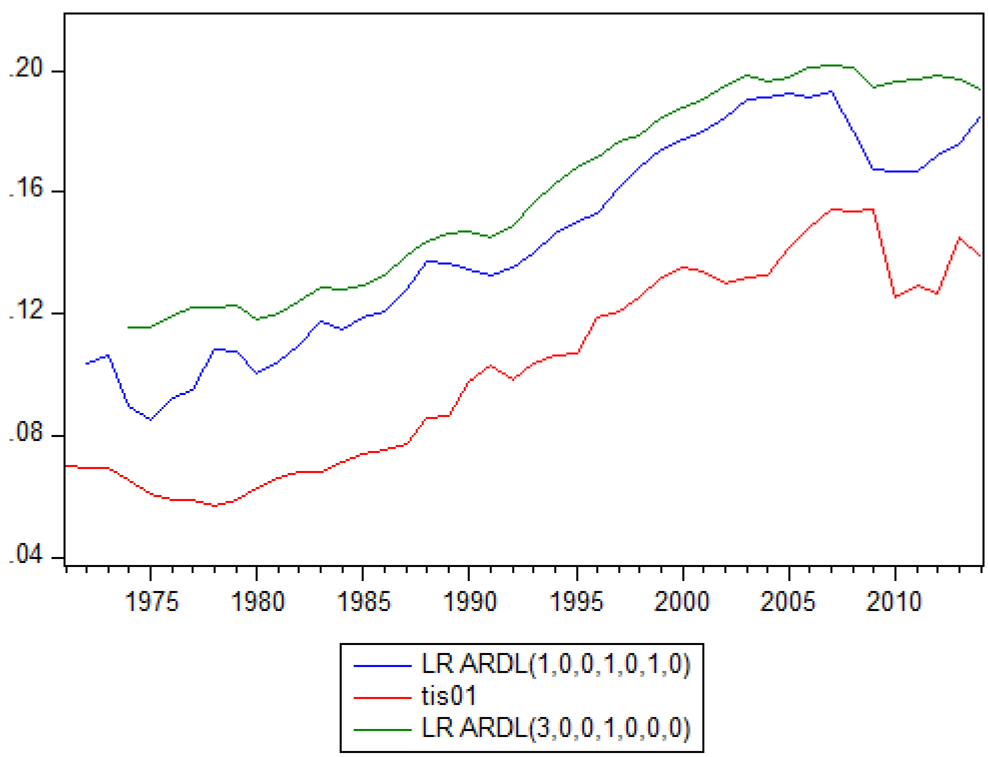

Figure 5.25: United Kingdom, LR relationships, Case IV

Summary and Conclusion

This thesis addresses the popular ARDL bounds test for cointegration and tries to demonstrate its underlying theoretical assumptions in a concise way so that someone interested in this area can easily apply and at the same time understand why and how it works. Often, misuse of a method or a test can lead to a bad situation or unexpectedly unwanted results. Thus, it is also crucial to understand under which circumstances the test doesn't work as expected. For this reason, four practical implementations of the test are presented showing some interesting behavior of the test in practice.

So, the present thesis consists of two parts. The first part is the theoretical one where we set out some basic time series theory and notation, then we explain the origin of the ARDL model, its connection with the Conditional ECM and how one can calculate and interpret the multipliers of the model. Next, we explain the connection with the underlying VAR and the corresponding VECM model and how endogeneity may be a problem under which the ARDL bounds test would be inappropriate to be used. The final section of the theoretical part explains the motivation behind the test and the analyses the Null hypothesis based on the VAR model, the Conditional ECM and also based on the CECM under the assumption that the problem of endogeneity is not present. The F-bounds and the complementary t-bounds tests are then presented under this last assumption.

The second part of this thesis is the practical application where the goal is to examine whether there is a long-run relationship between the 1% top income share and the macroeconomic factors of credit, education, gdp, inflation, population growth and trade. This application is splitted in four sections analysing separately the case of Greece, France, USA and UK. The case of Greece is similar to the one of France where the test concluded in favor of the existence of a long-run relationship but testing further for other possible cointegrating relationships in the whole system of equations, other cointegrating relationships between the `tis01` and the other explanatory variables were also found and so our first conclusion was unreliable. In the case of USA the unit root tests shows that `tis01` is trend stationary and thus $I(0)$ and so we knew in advance that the test was going to give specific results. Although, we do continue with the test in order to have a better understanding on how the test behaves under situations like this. As expected, the results was that the relationship is a degenerate one. Finally, the case of UK is somehow more complex. We conclude that `tis01` is also trend stationary and we present two sets of models. The first one is a not well specified model that suffers from autocorrelation and shows evidence for cointegration. As we can't rely on a model like this, we estimate a second model that seems to be correctly specified without having serial correlation problems but for which the test couldn't reject the hypothesis of no cointegration.

Appendix A

Code Appendix

A.1 R code

The cointegration simulation example

```
set.seed(2018)
n=5000
# y2~RW~I(1)
y2= cumsum(c(0,rnorm(n-1,0,1)))
# y1~I(1)
y1= 0.6*y2 +rnorm(n,0,1)

coint= y1- 0.6*y2
group= data.frame(y1,y2,coint)
groupt= ts(group)

labels= c(expression(y[1][t] %~% I(1)),
          expression(y[2][t] %~% I(1)),
          expression((y[1][t]-gamma*y[2][t]) %~% I(0)))

autoplot(groupt, facets = FALSE) +
    labs(caption=expression('for'~gamma==0.6)) +
    theme(plot.title = element_text(hjust=0.5)) +
    scale_colour_hue(name = "Time Series",
        breaks=colnames(groupt),
        labels=labels)
```

The non-cointegrated & non-unique cointegrating vectors simulation example

```
set.seed(2018)
n=5000
# y2~RW~I(1)
y2= cumsum(c(0,rnorm(n-1,0,1)))
# y1~I(1)
y1= 0.6*y2 +rnorm(n,0,1)

gamma1=0.4
gamma2=0.8
gamma3=0.6
beta=2.7

coint= y1- gamma3*y2
no_coint1= y1- gamma1*y2
no_coint2= y1- gamma2*y2
non_unique_coint= beta*y1 - (beta*gamma3*y2)

group= data.frame(y1,y2,no_coint1,no_coint2,non_unique_coint)
groupt= ts(group)

labels= c(expression(y[1][t] %~% I(1)),
          expression(y[2][t] %~% I(1)),
          expression((y[1][t]-gamma[1]*y[2][t]) %~% I(1)),
          expression((y[1][t]-gamma[2]*y[2][t]) %~% I(1)),
          expression((y[1][t]-gamma[3]*y[2][t]) %~% I(0)))

capt= expression('for:'~gamma[1]==0.4~ ~','~ gamma[2]==0.8 ~','~ gamma[3]==0.6%.%2.7)

autoplot(groupt, facets = FALSE) +
    labs(caption= capt) +
    theme(plot.title = element_text(hjust=0.5)) +
    scale_colour_hue(name = "Time Series",
        breaks=colnames(groupt),
        labels=labels)
```

A.2 EViews code

The codes provided here indicate how each test or equation should be done. The same can be applied to/with other variables just by replacing the variable name with the wanted one.

The following code is supposed produce the correct results using EViews10. In case where one uses a different EViews version or in case of different results, the following actions using the API should work as expected.

For the ARDL model construction, if the following code doesn't converge to the excepted order of ARDL, one should replace the variable name with `@fl(variable_name,n)` where `n` is the number of lags, in order to specify the exact number of lags.

Here we set some examples for the Unit Root tests. Other modifications can be applied as following:

Test in Levels: `dif=0`

Test in first differences: `dif=1`

Augmented Dickey-Fuller test: `adf`

Phillips-Perron: `pp`

Kwiatkowski-Phillips-Schmidt-Shin: `kpss`

no exogenous in the test equation (adf, pp) : `exog=none`

constant exogenous in the test equation (adf, pp, kpss): `exog=const`

constant and linear trend exogenous in the test equation (adf, pp, kpss): `exog=trend`

Innovational Outlier: `type=io`

Additive Outlier: `type=ao`

Break specification: `const/trend/both`

ADF Unit Root tests in Levels without exogenous in test equation

```
TIS01_GRC.uroot(dif=0, adf, none, lagmethod=sic)
```

ADF Unit Root tests in Levels with constant exogenous in test equation

```
CREDIT_GRC.uroot(dif=0, adf, const, lagmethod=sic)
```

Phillips-Perron Unit Root test in Levels withour exogenous in test equation

```
TIS01_GRC.uroot(exog=none, pp)
```

Phillips-Perron Unit Root test in Levels with constant exogenous in test equation

```
EDU_GRC.uroot(exog=const, pp)
```

Kwiatkowski-Phillips-Schmidt-Shin stationarity test in Levels without exogenous in test equation

```
TIS01_GRC.uroot(exog=const, kpss)
```

Breakpoint ADF Unit Root Test

```
TIS01_GRC.buroot(dif=0, type=io, exog=const, break=const,
breakmethod=dfuller, lagmethod=sic)
```

ARDL$(1,0,0,3,0,0,0)$ Case II model, Greece

```
ardl(deplags=4, reglags=3, ic=bic, trend=const)
TIS01_GRC CREDIT_GRC EDU_GRC GDP_GRC INFL_GRC POPG_GRC TRADE_GRC
'{%equation}.rename ardl1003000cii
```

Breusch-Godfrey Serial Correlation LM Test

```
ardl1003000cii.auto(1)
ardl1003000cii.auto(2)
ardl1003000cii.auto(3)
ardl1003000cii.auto(4)
ardl1003000cii.auto(5)
ardl1003000cii.auto(6)
```

Breusch-Pagan-Godfrey Heteroskedasticity Test

```
ardl1003000cii.hettest @regs
```

Create dummy variable

```
series D2014=@year=2014
```

ARDL$(1, 0, 1, 1, 1, 0, 0)$ Case II model with dummy, Greece

```
ardl(deplags=4, reglags=3, ic=bic, trend=const)
TIS01_GRC CREDIT_GRC EDU_GRC GDP_GRC INFL_GRC POPG_GRC TRADE_GRC @D2014
'{%equation}.rename ardl1011100cii
```

LM and BPG Tests

```
ardl1011100cii.auto(1)
ardl1011100cii.auto(2)
ardl1011100cii.auto(3)
ardl1011100cii.auto(4)
ardl1011100cii.auto(5)
ardl1011100cii.auto(6)
ardl1011100cii.hettest @regs
```

CUSUM and CUSUMSQ tests

```
ardl1011100cii.rls(q) c(1) c(2) c(3) c(4) c(5) c(6) c(7) c(8) c(9)
c(10) c(11) c(12)
ardl1011100cii.rls(v) c(1) c(2) c(3) c(4) c(5) c(6) c(7) c(8) c(9)
c(10) c(11) c(12)
```

Conditional Unrestricted ECM and bounds test

```
ardl1011100cii.cointrep
```

Conditional Restricted ECM

```
ardl1011100cii.ecreg
'In case where it produces the ARDL instead of the CRECM use the API
```

Cointegrating Relationship (Long-Run) fit

```
ardl1011100cii.makecoint coint_eq
group group_plots TIS01_GRC (TIS01_GRC - coint_eq)
freeze group_plots.line
```

ARDL fit

```
ardl1011100cii.resids(g)
```

ARDL(1,0,1,1,0,1,0) Case II model, France

```
ardl(deplags=3, reglags=2, ic=bic, trend=const)
TIS01_FRA CREDIT_FRA EDU_FRA GDP_FRA INFL_FRA POPG_FRA TRADE_FRA
'{%equation}.rename ardl1011010cii
```

LM and BPG Tests

```
ardl1011010cii.auto(1)
ardl1011010cii.auto(2)
ardl1011010cii.auto(3)
ardl1011010cii.auto(4)
ardl1011010cii.auto(5)
ardl1011010cii.auto(6)
ardl1011010cii.hettest @regs
```

CUSUM and CUSUMSQ tests

```
ardl1011010cii.rls(q) c(1) c(2) c(3) c(4) c(5) c(6) c(7) c(8) c(9)
c(10) c(11)
ardl1011010cii.rls(v) c(1) c(2) c(3) c(4) c(5) c(6) c(7) c(8) c(9)
c(10) c(11)
```

Conditional Unrestricted ECM and bounds test

```
ardl1011010cii.cointrep
```

Conditional Restricted ECM

```
ardl1011010cii.ecreg
'In case where it produces the ARDL instead of the CRECM use the API
```

Cointegrating Relationship (Long-Run) fit

```
ardl1011010cii.makecoint coint_eq
group group_plots TIS01_FRA (TIS01_FRA - coint_eq)
freeze group_plots.line
```

ARDL fit

```
ardl1011010cii.resids(g)
```

ARDL(1,0,0,1,0,0,0) Case II model, USA

```
ardl(deplags=4, reglags=3, ic=bic, trend=const)
TIS01_USA CREDIT_USA EDU_USA GDP_USA INFL_USA POPG_USA TRADE_USA
'{%equation}.rename ardl1001000cii
```

ARDL(1,0,0,1,0,0,0) Case III model, USA

```
ardl(deplags=4, reglags=3, ic=bic, trend=uconst)
TIS01_USA CREDIT_USA EDU_USA GDP_USA INFL_USA POPG_USA TRADE_USA
'{%equation}.rename ardl1001000ciii
```

ARDL(1,0,0,0,0,0,0) Case IV model, USA

```
ardl(deplags=4, reglags=3, ic=bic, trend=linear)
TIS01_USA CREDIT_USA EDU_USA GDP_USA INFL_USA POPG_USA TRADE_USA
'{%equation}.rename ardl1000000civ
```

Cointegrating Relationship (Long-Run) fit

```
ardl1000000civ.makecoint coint_eq
group group_plots TIS01_USA (TIS01_USA - coint_eq)
freeze group_plots.line
```

ARDL(3,0,1,1,0,0,2) Case I model, United Kingdom

```
ardl(deplags=4, reglags=3, ic=bic, trend=none)
TIS01_GBR CREDIT_GBR EDU_GBR GDP_GBR INFL_GBR POPG_GBR TRADE_GBR
'{%equation}.rename ardl3011002ci
```

ARDL(1,0,0,1,0,0,0) Case I model, United Kingdom

```
ardl(deplags=2, reglags=1, ic=bic, trend=none)
TIS01_GBR CREDIT_GBR EDU_GBR GDP_GBR INFL_GBR POPG_GBR TRADE_GBR
'{%equation}.rename ardl1001000ci
```

Conditional Unrestricted ECM and bounds test for CASE I

```
ardl3011002ci.cointrep
ardl1001000ci.cointrep
```

Cointegrating Relationship (Long-Run) fit

```
ARDL1001000CI.makecoint coint_1001000CI
ARDL3011002CI.makecoint coint_3011002CI
group group_plots TIS01_gbr (TIS01_gbr - coint_1001000CI)
(TIS01_gbr - coint_3011002CI)
freeze group_plots.line
```

ARDL(3,0,0,1,0,0,2) Case II model, United Kingdom

```
ardl(deplags=3, reglags=3, ic=bic, trend=const)
TIS01_GBR CREDIT_GBR EDU_GBR GDP_GBR INFL_GBR POPG_GBR TRADE_GBR
'{%equation}.rename ardl3001002cii
```

ARDL(1,0,0,1,0,0,0) Case II model, United Kingdom

```
ardl(deplags=2, reglags=1, ic=bic, trend=const)
TIS01_GBR CREDIT_GBR EDU_GBR GDP_GBR INFL_GBR POPG_GBR TRADE_GBR
'{%equation}.rename ardl1001000cii
```

Conditional Unrestricted ECM and bounds test for CASE II

```
ardl3001002cii.cointrep
ardl1001000cii.cointrep
```

Cointegrating Relationship (Long-Run) fit

```
ardl1001000cii.makecoint coint_ardl1001000cii
ardl3001002cii.makecoint coint_ardl3001002cii
group group_plots TIS01_gbr (TIS01_gbr - coint_ardl1001000cii)
(TIS01_gbr - coint_ardl3001002cii)
freeze group_plots.line
```

CUSUMSQ test

```
ardl1001000cii.rls(v) c(1) c(2) c(3) c(4) c(5) c(6) c(7) c(8) c(9)
```

ARDL(3,0,0,1,0,0,0) Case IV model, United Kingdom

```
ardl(deplags=4, reglags=3, ic=bic, trend=linear)
TIS01_GBR CREDIT_GBR EDU_GBR GDP_GBR INFL_GBR POPG_GBR TRADE_GBR
'{%equation}.rename ardl3001000civ
```

ARDL(1,0,0,1,0,1,0) Case IV model, United Kingdom

```
ardl(deplags=2, reglags=1, ic=bic, trend=linear)
TIS01_GBR CREDIT_GBR EDU_GBR GDP_GBR INFL_GBR POPG_GBR TRADE_GBR
'{%equation}.rename ardl1001010civ
```

Conditional Unrestricted ECM and bounds test for CASE IV

```
ardl3001000civ.cointrep
ardl1001010civ.cointrep
```

Cointegrating Relationship (Long-Run) fit

```
ardl1001010civ.makecoint coint_ardl1001010civ
ardl3001000civ.makecoint coint_ardl3001000civ
group group_plots TIS01_gbr (TIS01_gbr - coint_ardl1001010civ)
(TIS01_gbr - coint_ardl3001000civ)
freeze group_plots.line
```

Appendix B

Math Appendix

Lag polynomial (*AR* operator) (B.1)

$$\begin{aligned}\psi(L) &= 1 - \psi^*(L) = 1 - \sum_{i=1}^{p} \psi_i L^i \\ &= 1 - \psi_1 L^1 - \psi_2 L^2 - \cdots - \psi_p L^p\end{aligned} \tag{B.1}$$

$$\psi(1) = 1 - \psi^*(1) = 1 - \sum_{i=1}^{p} \psi_i \tag{B.2}$$

$$\psi^*(1) = \sum_{i=1}^{p} \psi_i \tag{B.3}$$

$$\psi^*(L) = \sum_{i=1}^{p} \psi_i L^i \tag{B.4}$$

$$\widetilde{\psi}_i^* = - \sum_{r=i+1}^{p} \psi_r \tag{B.5}$$

$$\widetilde{\psi}^*(L) = \sum_{i=1}^{p-1} \widetilde{\psi}_i^* L^{i-1} \tag{B.6}$$

$$\mathbf{\Phi}(L) = \mathbf{I}_{k+1} - \sum_{i=1}^{p} \mathbf{\Phi}_i L^i \tag{B.7}$$

Lag polynomial (*MA* operator) (B.8)

$$\begin{aligned}\beta(L) &= \sum_{i=0}^{q} \beta_i L^i \\ &= \beta_0 L^0 + \beta_1 L^1 + \cdots + \beta_q L^q\end{aligned} \tag{B.8}$$

The form of the MA operator accounting for more than one variables:

$$\begin{aligned}\beta_j(L) &= \sum_{l_j=0}^{q_j} \beta_{j,l_j} L^{l_j} \\ &= \beta_{j,0}L^0 + \beta_{j,1}L^1 + \cdots + \beta_{j,q_j}L^{q_j}\end{aligned} \tag{B.9}$$

$$\beta(1) = \sum_{l_j=0}^{q_j} \beta_{l_j} \tag{B.10}$$

Open form of general ARDL$(p, q_1, \ldots, q_k)$ **Eq.** (4.2)

$$\begin{aligned}y_t = {} & \alpha_0 + \alpha_1 t \\ & + \psi_1 y_{t-1} + \psi_2 y_{t-2} \ldots \psi_p y_{t-p} \\ & + \beta_{1,0}x_{1,t} + \beta_{1,1}x_{1,t-1} + \cdots + \beta_{1,q_1}x_{1,t-q_1} \\ & + \beta_{2,0}x_{2,t} + \beta_{2,1}x_{2,t-1} + \cdots + \beta_{2,q_2}x_{2,t-q_2} \\ & \vdots \\ & + \beta_{k,0}x_{k,t} + \beta_{k,1}x_{k,t-1} + \cdots + \beta_{k,q_k}x_{k,t-q_k} + \varepsilon_t\end{aligned} \tag{B.11}$$

Beveridge-Nelson decomposition

For an AR process (B.12):

$$\psi(L) = \psi(1) + (1-L)\widetilde{\psi}(L) \tag{B.12}$$

$$\psi^*(L) = (\psi^*(1) + (1-L)\widetilde{\psi}^*(L))L \tag{B.13}$$

For a MA process (B.14):

$$\beta(L) = \beta(1) + (1-L)\widetilde{\beta}(L) \tag{B.14}$$

$$\widetilde{\beta}^*(L) = \widetilde{\beta}(L) - \widetilde{\psi}(L)\psi(L)^{-1}\beta(L) \tag{B.15}$$

Appendix C

Tables Appendix

Table C.1: Conditional ECM, Greece

	Dependent variable: Δtis01			
	Coefficient	Std.Error	t-Statistic	P-Value
Intercept	0.059	0.012	4.985	0.000
tis01_{t-1}	-0.437	0.102	-4.267	0.000
credit_t	2.530e$-$04	7.210e$-$05	3.511	0.001
edu_t	0.292	0.131	2.230	0.033
gdp_{t-1}	$-$2.920e$-$06	5.530e$-$07	-5.283	0.000
infl_t	$-$3.950e$-$04	1.610e$-$04	-2.458	0.020
popg_t	5.510e$-$04	2.079e$-$03	0.265	0.793
trade_t	1.680e$-$04	1.350e$-$04	1.243	0.223
Δgdp_t	$-$1.480e$-$06	9.390e$-$07	-1.578	0.125
Δgdp_{t-1}	3.730e$-$06	1.040e$-$06	3.585	0.001
Δgdp_{t-2}	3.190e$-$06	1.140e$-$06	2.805	0.009

Table C.2: Diagnostic test for ARDL, Greece

Statistic	Serial Correlation LM Test: Breusch-Godfrey lags=1	lags=2	lags=3	lags=4	lags=5	lags=6
Obs∗R^2	1.935e−03	4.633	6.698	7.407	7.579	8.346
Pr($>\chi^2_{d.f.}$)	0.965	0.099	0.082	0.116	0.181	0.214

Note: The *d.f.* for each test is equal to the lags included
LM test H_0: No serial correlation up to order of the *d.f.* of $\chi^2_{d.f.}$

Statistic	Heteroskedasticity Test: Breusch-Pagan-Godfrey Value
Obs∗R^2	34.159
Pr($>\chi^2_{10}$)	0.000

BPG test H_0: No heteroskedasticity

Table C.3: Diagnostic test for ARDL with dummy, Greece

Statistic	Serial Correlation LM Test: Breusch-Godfrey lags=1	lags=2	lags=3	lags=4	lags=5	lags=6
Obs∗R^2	3.773	4.131	4.192	4.221	5.608	8.221
Pr($>\chi^2_{d.f.}$)	0.052	0.127	0.241	0.377	0.346	0.222

Note: The *d.f.* for each test is equal to the lags included
LM test H_0: No serial correlation up to order of the *d.f.* of $\chi^2_{d.f.}$

Statistic	Heteroskedasticity Test: Breusch-Pagan-Godfrey Value
Obs∗R^2	12.112
Pr($>\chi^2_{11}$)	0.355

BPG test H_0: No heteroskedasticity

Table C.4: Levels Equation, Greece

	Dependent variable: tis01			
	Case 2: Restricted Constant and No Trend			
	Coefficient	Std.Error	t-Statistic	P-Value
Intercept	0.135	0.013	10.800	0.000
credit_t	5.800e−04	1.380e−05	41.889	0.000
edu_t	0.668	0.119	5.637	0.000
gdp_t	−6.690e−06	2.540e−07	-26.356	0.000
infl_t	−9.040e−04	2.034e-04	-3.867	0.001
popg_t	1.261e−03	2.521e−03	0.500	0.621
trade_t	3.840e−04	2.020e−03	1.901	0.067

Table C.5: ECM, Greece

	Dependent variable: Δtis01			
	Case 2: Restricted Constant and No Trend			
	Coefficient	Std.Error	t-Statistic	P-Value
Δgdp_t	−1.480e−06	6.760e−07	−2.190	0.036
Δgdp_{t-1}	3.730e−06	8.640e−07	4.315	0.000
Δgdp_{t-2}	3.1904−06	8.980e−07	3.549	0.001
ECT_{t-1}	−0.437	0.055	−7.956	0.000
Observations	41	Residual Std. Error		2.813e−03
R^2	0.667	Log Likelihood		184.743
Adjusted R^2	0.639	AIC		−8.817
BIC	−8.649			

Table C.6: Conditional ECM with dummy, Greece

	Dependent variable: Δtis01			
	Coefficient	Std.Error	t-Statistic	P-Value
Intercept	0.018	0.004	4.751	0.000
tis01_{t-1}	-0.225	0.040	-5.624	0.000
credit_t	-2.18e-05	2.66e-05	-0.818	0.419
edu_{t-1}	-0.089	0.075	-1.166	0.253
gdp_{t-1}	-8.58e-08	2.11e-07	-0.407	0.687
infl_{t-1}	-4.800e-04	8.19e-05	-5.860	0.000
popg_t	-7.130e-04	0.001	-0.710	0.483
trade_t	1.070e-04	6.08e-05	1.757	0.089
Δedu_t	0.267	0.119	2.226	0.033
Δgdp_t	-1.19e-06	4.53e-07	-2.640	0.013
Δinfl_t	-2.050e-04	8.58e-05	-2.393	0.023
D2014_t	0.024	0.002	11.249	0.000

Table C.7: Diagnostic test for ARDL, France

	Serial Correlation LM Test: Breusch-Godfrey					
Statistic	lags=1	lags=2	lags=3	lags=4	lags=5	lags=6
Obs$*$R^2	1.403	3.843	5.367	6.798	9.496	10.045
Pr($>\chi^2_{d.f.}$)	0.236	0.146	0.147	0.147	0.091	0.123

Note: The $d.f.$ for each test is equal to the lags included
LM test H$_0$: No serial correlation up to order of the $d.f.$ of $\chi^2_{d.f.}$

	Heteroskedasticity Test: Breusch-Pagan-Godfrey
Statistic	Value
Obs$*$R^2	7.004
Pr($>\chi^2_{10}$)	0.725

BPG test H$_0$: No heteroskedasticity

Table C.8: Conditional ECM, France

	Dependent variable: Δtis01			
	Coefficient	Std.Error	t-Statistic	P-Value
Intercept	3.868e-03	7.107e-03	0.544	0.590
tis01_{t-1}	-0.358	0.094	-3.815	6.000e-04
credit_t	1.480e-05	5.610e-05	0.264	0.793
edu_{t-1}	0.464	0.309	1.498	0.143
gdp_{t-1}	8.850e-07	6.500e-07	1.361	0.182
infl_t	2.790e-04	3.360e-04	0.830	0.412
popg_{t-1}	0.014	4.944e-03	2.905	0.000
trade_t	-4.510e-04	2.540e-04	-1.779	0.084
Δedu_t	-1.134	0.726	-1.560	0.128
Δgdp_t	6.200e-06	1.330e-06	4.673	0.000
Δpopg_t	2.958e-03	4.967e-03	0.595	0.555

References

Banerjee, A. (1993). *Co-integration, error-correction, and the econometric analysis of non-stationary data.* (C. J. Granger & G. E. Mizon, Eds.) (p. 329). Oxford University Press.

Davidson, J. E. H., Hendry, D. F., Srba, F., & Yeo, S. (1978). Econometric Modelling of the Aggregate Time-Series Relationship Between Consumers' Expenditure and Income in the United Kingdom. *The Economic Journal*, *88*(352), 661–692.

Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, *55*(2), 251–276.

Granger, C. (1983). Co-integrated variables and error-correcting models. *Discussion Paper , Department of Economics, University of California at San Diego*, *13*, 83–13.

Granger, C. W. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, *16*(1), 121–130.

Hamilton, J. D. (1994). *Time Series Analysis* (pp. 1–799). Princeton University Press.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. (C. J. Granger & G. E. Mizon, Eds.) (p. 267). Oxford University Press.

Maddala, G. S., & Lahiri, K. (2009). *Introduction to econometrics* (p. 634). Wiley.

Narayan, P. K. (2005). The saving and investment nexus for China: Evidence from cointegration tests. *Applied Economics*, *37*(17), 1979–1990.

Pesaran, M. H., & Shin, Y. (1998). An autoregressive distributed-lag modelling approach to cointegration analysis. *Econometric Society Monographs*, *31*(7), 371–413.

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, *16*(3), 289–

326.

Phillips, P. C. B., & Ouliaris, S. (1990). Asymptotic Properties of Residual Based Tests for Cointegration. *Econometrica*, *58*(1), 165–193.

Piketty, T. (2001). *Les hauts revenus en France au XXème siècle. Inégalités et redistributions, 1901-1998* (pp. 1–807). Grasset, Paris.