Bulletin of Applied Economics, 2021, 8(1), 29-57
https://doi.org/10.47260/bae/813

Determinants of tax evasion in Greece: Econometric analysis of co-integration and causality, variance decomposition and impulse response analysis

Anastasiou Athanasios[1], Kalligosfyris Charalampos[1] and Kalamara Eleni[1]

Abstract

The purpose of this paper is to examine the causality relationships and the degree of interdependence, between the level of tax evasion in Greece and a set of deterministic factors, using annual data for the period 1995 - 2018. The research methodology employed includes testing for stationarity with the Augmented Dickey-Fuller (ADF) test, cointegration test according to Engle-Granger approach, estimation Error Correction Models (ECM) to investigate the short-run and long-run relationships, variance decomposition and impulse response analysis. The results indicate a significant interdependence, which is an important tool for pursuing a targeted and effective policy to fight tax evasion. More specifically, the survey showed that the level of tax rates, the level of unemployment, the Rule of Laws index, the level of GDP, the level of non-performing loans, the government efficiency, the corruption perception index and the level of final consumption expenditure, affect the size of tax evasion in Greece, significantly. In addition, the results of variance decomposition and impulse response analysis, support the above findings, providing a quantitative representation of the causality relationships between the factors under investigation and tax evasion.

JEL classification numbers: E60, H26, C10
Keywords: Macroeconomics, public economics, applied economics, tax evasion, Greece

1 Introduction

Tax evasion has been the subject of scientific research in many developed economies for a long time (Andreoni Erard & Feinstein 1998, Cuccia 1994, Jackson & Milliron 1986, Long & Swingen 1991, Richardson & Sawyer 2001, Richardson 2006).

Following the theoretical framework introduced by the Allingham & Sandmo model (1972) and its theoretical extensions, as later formulated, empirical investigation of tax evasion and its determinants, has to offered us significant results for both traditional and non-traditional determinants. Nowadays, there are several references in the international literature, including the one by Jackson & Milliron (1986), Cowell (1990), Andreoni (1998), Franzoni (1999) and Torgler (2007). It is worth noting that most of the empirical research has dealt with the subject of tax evasion at the individual level, while similar investigations at business or country level are less frequent. For this reason, the need to investigate the

[1] Department of Management Science and Technology, University of Peloponnese, Greece

Article Info: Received: December 18, 2020 Published online: January 25, 2021

issue beyond the level of individual behavior becomes imperative and has been suggested by notable researchers, several times.

The purpose of this study is to examine the degree of interdependence between the level of tax evasion in Greece and a set of possible determinants, using appropriate econometric methods of time series analysis. Particularly, the main research question which is studied in this paper is what are the causality relationships between tax evasion and a set of factors and to what extent changes in the level of tax evasion in Greece are interpreted by changes in the factors under investigation, in the sense that these changes cause, influence and "lead" the variance of the level of tax evasion in Greece.

It is important to realize that any effort to reduce the size of tax evasion and its consequences in an economy, require an in-depth investigating of those factors that shape a conducive environment to its development. Therefore, the importance of a study aimed at revealing the determinants of tax evasion, would be an important tool of fiscal policy aimed at implementing an effective reform policy that would reduce it.

The study focuses on the Greek economy, mainly due to the recent fiscal crisis experienced by the country, which led to the need to improve the macroeconomic performance of the economy and reduce the extent of tax evasion through targeted policies by improving crucial factors that have been proven to affect the level of tax compliance.

2 Literature Review

The first major literature review of tax evasion was carried out by Jackson & Milliron (1986), which recorded 14 determinants such as age, gender, educational level, employment level (demographic factors), income level, source of income, tax rates, level of penalties imposed and probability of detection (economic factors), complexity of the tax system, fairness, direct contact with tax authorities, the general level of tax compliance of taxpayers and morality or fiscal morality (behavioral factors).

Significant traditional determinants of tax evasion include tax rates, probability of audit and level of fine. In particular, marginal tax rates are an important determinant of tax evasion, but the empirical results of the studies carried out on their impact on the size of tax evasion are controversial. Research studies such as Clotfelter (1983), Mason & Calvin (1978), Alm et al.(1992), Pommerehne & Weck-Hannemann (1996), Sillamaa and Veall (2000), Torgler (2006) and Chiarini et al. (2008), showed a positive relationship between tax rates and tax evasion, unlike studies such as Feinstein (1991), Christian & Cupta (1992), Kamdar (1995) and Joulfaian (2009), which claim that there is a negative relationship linking these two variables. In fact, Richardson & Sawyer (2001) argued that the lack of investigation between the tax rates and the level of income causes this inconsistency of the relationship between tax rates and tax evasion. In addition, case studies have been recorded in relevant international literature (Marelli (1984), Sandmo (2004), Nur-tegin (2008), Reinganum & Wilde (1985), Baldry (1987), Porcano (1988) and Joulfaian and Rider (1996), there was no effect of the tax rates on the size of tax evasion, creating a framework of dubious relations between these two variables.

The second traditional determinant of tax evasion is the probability of detection (Jackson & Milliron, 1986). There is a general coincidence of views on this issue among most of the studies carried out, arguing that the probability of detection and disclosure is linked to a strong, negative and significant relationship to the size of tax evasion, although the crucial question is to what extent the increase in the probability of detecting the illegal actions of taxpayers will increase tax compliance (Andreoni et al. 1998). Witte & Woodbury (1985) analyzed Individual Taxpayer Compliance Measurement. Program

(TCMP) data from 1969 and found that tax evasion is inversely proportional to the probability of audit. Crane & Nourzad (1986) also, using the percentage of tax returns audited by the tax authorities as a measure of the probability of being detected, revealed that an increase in the probability of being detected, on average, leads to a lower level of undeclared incomes. Dubin et al. (1987) analyzed time series data with purpose to further checking the case of the impact of the probability of detection, concluding that a higher probability of detection is related to increased levels of tax compliance as a higher level of declared income. Subsequently, Pommerehne & Weck-Hannemann (1996) compared data from various Swiss cantons and they came to a conclusion that compliance was higher in cantons where more audits were carried out during a fiscal period. However, there are cases where tax compliance is reduced slightly, following a tax audit (Mittone, 2006), thus indicating a positive relationship between tax evasion and tax audits. However, Kastlunger et al. (2009) argued that the reduction in tax compliance observed after an audit is very likely to be caused by a misassessment of the situation. In contrast to the above, the lack of influence of the probability of tax audit over the level of tax evasion has been observed. In particular, Schram & Gerxhani (2006) conducting experiments on this subject in Albania, observed that the participants were not affected by the rate of tax audits (16.6% and 50%). Moreover, Falsetta et al. (2010) argue that the probability of audit affects the decision of taxpayers' compliance only when they accept and support the way the government allocates tax revenues. When taxpayers do not support government programs, their compliance is less, regardless of the possibility of audit.

A second category of factors includes the institutional factors of tax evasion. Among many such factors, we cite variables related to citizens' confidence in state institutions, perception of the level of corruption and the cost of tax compliance. Such factors, which are often found in the literature, cover aspects of the issue relating to justice, the effectiveness of public administration and the state in general, transparency and accountability (Andreoni et. al 1998).

It is generally accepted that the perceptions of tax justice are related to the level of tax evasion in an economy (Jackson & Milliron, 1986), without underestimating the importance of these perceptions (Richardson & Sawyer, 2001). Spicer (1974) discovered a significant negative relationship between tax justice and tax evasion. Song & Yarbrough (1978) also showed that there is a significant negative relationship between these two variables, with 75% of taxpayers arguing that the concept of justice regarding to the "ability to pay" is more important than the "benefit" that tax evaders will enjoy. In addition, Hite & Roberts (1992) found that justice is significantly linked to the perceptions of an improved tax system and that tax justice and tax evasion are negatively related.

Also as tax systems become increasingly complex over the years in many developed economies, this complexity is an important determinant of tax evasion (Jackson & Milliron 1986, Richardson & Sawyer 2001). Previous investigations using methods in both the archive database (Clotfelter 1983, Long & Swingen 1988) and research data (Collins et. al 1992, Milliron & Pu 1988, Vogel 1974) provide strong empirical evidence which demonstrates the existence of a positive relationship between the complexity of the tax system and tax evasion.

In addition, taxpayers' contact with tax authorities is still a determining factor in tax evasion. In a major survey on the structure of the tax system in Europe conducted by Strumpel (1969), was found that as the southernmost European countries have a long history of trying to improve the level of tax compliance of citizens by strengthening their enforcement efforts, these countries have the lowest levels of tax compliance in Europe (Jackson & Milliron 1986). In view of this remark, the investigation by Spicer & Lundstedt's (1976) in the United States. showed that direct contact between taxpayers and tax authorities is likely to be associated with increased tax resistance and hence tax evasion. Wallschutzky

(1984), Klepper & Nagin (1989) and Brooks & Doob (1990) also support this opinion. On the other hand, by reducing the level of contacts between taxpayers and public tax authorities, through a distance service system, the probability of tax evasion is reduced (Tanzi 2000, Torgler & Murphy 2004).

Moreover, while tax morality is an vague concept (Jackson & Milliron 1986), it describes the ethical principles and values of individuals regarding to the payment of taxes (Torgler & Murphy 2004). Research carried out by Spicer (1974), Spicer & Lundstedt (1976) and Tittle (1980) has shown that the tax morality of individuals has a negative connection to tax evasion. In fact, Torgler (2003) showed that tax morality and tax evasion are negatively related. In addition, Riahi-Belkaoiou (2004) found empirical evidence suggesting that tax evasion between countries is negatively related to selected tax morality factors.

Furthermore, tax evasion determinants could be considered those relating to social, economic, demographic and behavioral characteristics of taxpayers. Age, gender, educational level, social norms, income level, source of income, marital status and religious emotion are the most common factors found in the relevant literature. It is worth noting that some of them are less important or even less observable if we try to look at tax evasion not at the level of an individual but at the level of a business or country.

In particular, the age of the taxpayer is one of the most important determinants of tax evasion (Jackson & Milliron (1986)), as studies have shown that older taxpayers generally have a higher level of tax compliance than younger taxpayers (Tittle 1980, Witte & Woodbury 1985, Dubin & Wilde 19199) 88, Feinstein 1991, Hanno & Violete 1996). In fact, Tittle (1980) argues that the relationship between age and this divergent attitude of different age groups towards taxes is due to changes in life cycle and differences between generations, as younger taxpayers are more risk-friendly, they do not consider about the possible effects of their actions, nor about the penalties that may be imposed on them and reflect the social and psychological differences associated with them the period during which they were borned and raised (a generation gap).

Previous studies have shown that the gender of taxpayers is an important determinant of tax evasion. For example, Vogel (1974) and Mason & Calvin (1978) showed that the level of tax compliance of women is higher than that of men. Jackson & Milliron (1986) argues that the tax compliance gap between men and women is shrinking as new generations of liberated women emerge. However, other studies on the gender of taxpayers and its relationship to the level of tax evasion show a tendency for the tax compliance gap between men and women to be maintained (e.g. Brooks & Doob 1990, Collins, Milliron & Toy 1992).

Furthermore, the level of education is considered as another important determinant of tax evasion, as it is usually related to taxpayers' ability to understand tax laws and then to decide whether or not to comply with them (Jackson & Milliron, 1986). In particular, Jackson & Milliron (1986) argued that education had two important elements: the general extent to which taxpayers perceive and understand their tax obligations through tax laws and the specific degree of knowledge about the possibilities of tax evasion. In fact, they argue that by strengthening the level of general tax knowledge, tax compliance is improved because of the positive views and perceptions generated through education about the necessity of taxes and their reciprocal relationship with the standard of living. On the other hand, the increased knowledge of tax evasion possibilities has a negative effect on the level of tax compliance as in essence it enhances tax evasion. However, the majority of studies associating in the influence of the level of education of taxpayers on tax evasion use a general level of education as a measure of citizen education (Richardson & Sawyer 2001). In fact, studies such as Song & Yarbrough (1978), Wallschutzky (1984) and Witte & Woodburry (1985) showed a negative relationship between the general educational level of taxpayers and tax evasion.

Similarly, the level of income is another factor determining the size of tax compliance. We usually refer to adjusted net income or the overall positive income of taxpayers (Jackson & Milliron, 1986). Mason & Lowry (1981) and Witte & Woodbury (1983) demonstrated that middle-income taxpayers generally comply with tax laws, while low-income taxpayers as well as high-income taxpayers generally show a low level of tax compliance. However, Richardson & Sawyer (2001) showed that these empirical results remain confused.

Regarding the sources of taxpayers´ income, these are often referred to as factors determining the size of tax evasion and related to the type or nature of citizens' income (Jackson & Milliron, 1986). Schmolder (1970) showed that when a large part of the labor force of an economy deals with agriculture and small businesses, income and profit taxation is unsuccessful. Wallschutzky's study (1984) showed that those who derive their income from agriculture, trade, and self-employment were more likely to avoid income taxes, while those whose source of income was related to wages or wages subject to withholding, such as the service sector, were less likely to do so.

In addition, these factors include macroeconomic factors, which are found at the level of a study of a group of countries (although rare) or in surveys that use time series at the level of the economy, individually. This category includes per capita income, inflation, unemployment and other macroeconomic indicators.

Specifically, most of the studies conducted on the impact of per capita GDP have shown that by increasing income at national level we are leading to a higher level of economic development of the country, which in turn increases the rate of tax compliance. Chelliah (1971) argued that higher per capita income reflects a higher level of growth, which in turn means not only greater willingness to pay taxes in the form of tax capacity, but also greater capacity to collect taxes as there is a taxable material and income to be taxed. Other studies show that taxpayers with a better standard of living tend to create strong links to behaviors of compliance with social systems (Hinrichs 1966, Tanzi 1987, Ghura 1998). In addition, Frey & Weck-Hanneman (1984) argues that in countries with low per capita income, people tend to work in more than one job, but in their tax return, the income of only one of them is reported. Furthermore, Boame (2009) showed that low per capita income leads to limited liquidity which in turn could create problems with payment and collection of taxes. In fact, Sookram & Watson (2005) found that per capita income has a negative relationship to tax evasion in short run, but in the long run this relationship turns into a positive one. In contrast to the above results, Feige & Cebula (2009) studying the attitude of US taxpayers towards tax compliance for the period 1960-2008, found a positive relationship between GDP per capita and tax evasion, arguing that an increase in per capita income would increase tax evasion.

Furthermore, it is widely known that an increase in unemployment will lead to a decrease in taxpayers' income, which in turn will increase tax evasion. In addition, an increase in cash use caused by rising unemployment as individuals may turn to the black economy or informal economy, may lead to tax collection problems. Alm & Yunus (2009), empirically investigating the issue of the USA in 1979-1997, found a statistically significant and positive relationship between unemployment and the level of tax evasion, arguing that tax evasion is increasing in times of economic downturn. Moreover, they notice that in times of high unemployment, people work in the gray economy and are paid in cash, which then usually do not report to the tax authorities. Similar results were found by Dubin et al. (1987) investigating the issue in the United States with data from the Internal Revenue Service (IRS) 1977-1985, Jou (1992) analyzing data at state level for the period 1976-1989 and Cebula & Feige (2009) with data from the Internal Revenue Service (IRS) for the period 1966-2008. In contrast, Boame (2009), using macroeconomic time series for the period 1987-2003 for Canadian taxpayers, found that an increase in the unemployment rate has a negative and statistically significant effect on the level of tax evasion.

In addition, characteristics related to the business, such as its size, legal form, ownership and activities, could be considered as determinants. However, these factors are rare and are only found in a limited number of studies (Wallace 2002, Slemrod 2007, Rice 1992, Nur-tegin 2008, Joulfaian 2009, Chan et al. 2000, Crocker & Slemrod, 2005, Tedds 2006, Goerke 2006, Negreponte-Delivani, 1990).

3 Data

The purpose of this paper is to investigate, through cointegration theory as developed by Engle and Granger (1987), the long run and short run relationship between tax evasion in Greece and a set of possible determinants of it and to understand the structure of the information transmission mechanism and influence each other. The aim of this effort is to discover and formulate the causality relationships between the level of tax evasion in Greece and its determinants. In this context, we estimate Error Correction Models on a bi-variable level to reveal the causal relationships of the factors under investigation with tax evasion, which allow us to proceed with the imposition of shocks on the variables under investigation, in order the variance decomposition of the forecast error and the estimation of the impulse response functions.

The study covers 12 tax evasion factors under investigation, the selection of which was made based on the conclusions of previous research works in other countries, as developed extensively in the previous section of the literature review. In particular, the analysis includes the following variables:

Table 1: Variables under investigation

Variable	Symbol	Data Source
Tax evasion	TE	Anastasiou A., Kalamara E. and Kalligosfyris C. (2020), Estimation of the size of tax evasion in Greece, Bulletin of Applied Economics, 2020, 7(2), 97-107
Tax Rates	TR	World Bank
Gross Domestic Product (per capita)	GDP	World Bank
Unemployment Rate	U	World Bank
Rule of Law index (as an element of adherence to the rules of law)	RL	World Justice Project
Public Debt (as a percentage of GDP)	PD	World Bank
Private Sector Savings	S	Bank of Greece
Non-Performing Loans Index (as an element of assessment of the business environment)	L	Bank of Greece
Government Effectiveness Index	GE	World Bank
Proportion of Self-Employed	SE	World Bank
Central Government Expenditure	EG	World Bank
Corruption Perceptions Index	C	Transparency International
Level of Education*	ED	World Bank
Final Consumption Expenditure (as a percentage of GDP)	CE	World Bank

**Annual increase in the number of students enrolled in higher education*

The data cover the period from 1995 to 2018 and concern annual observations expressed in logarithmic values of the variables, on the one hand to stabilize the variance, thereby ensuring the stationarity of the series in terms on variance and on the other hand to have values expressed in a common scale.

It is noted that for the estimation of the size of tax evasion in Greece in the period 1995-2018, we used the results of the study of Anastasiou A., Kalamara E. and Kalligosfyris C. (2020) which were based on the application of an indirect approach to the issue, as developed by Tanzi and based on the assumption that estimating the size of the gray economy can lead us to a safe measure of the extent of tax evasion.

4 Methodological issues

4.1 Stationarity test

The investigation of stationarity or non-stationarity of time series is a problem which must be addressed in the early stages of the empirical investigation. The verification of the existence of causality relationships between a set of variables, based essentially on autoregressive vectors, requires the time series under investigation be stationary. In addition, the application of cointegration theory requires the time series, which are involved in the subject of the analysis to be integrated and in particular of the same order that is, to contain at least one unit root. For these reasons, it is necessary to examine the statistical properties of the time series by checking the existence or non-unit root. To check the non-stationarity of series, the augmented Dickey - Fuller test will be used, in the logarithmic values of the variables considered.

4.1.1 Augmented Dickey - Fuller test

Assuming that a given economic time series {Yt} can be described by a first order autoregressive model, AR(1), then we should write the following:

$$Y_t = a Y_{t-1} + e_t \qquad (4.1)$$

In order the series to be stationary, it has to be -1 ‹ α ‹ 1. If a=1 the series is non-stationary, but it becomes stationary if we take first differences. So, to check the existence of a unit root, we are testing the null hypothesis, H_0: α=1.

If in the above relationship (4.1) we subtract from both members the Y_{t-1} the following equivalent form will emerge:

$$\Delta Y_t = \gamma Y_{t-1} + e_t \qquad (4.2)$$

where γ= α-1 and the test of the null hypothesis H_0: a=1 becomes a test of the hypothesis H_0: γ=0. Dickey - Fuller considered three different regressions that can be used to check for the presence or absence of a unit root:

$$\Delta Y_t = \gamma Y_{t-1} + e_t$$
$$\Delta Y_t = a_0 + \gamma Y_{t-1} + e_t$$
$$\Delta Y_t = a_0 + \gamma Y_{t-1} + a_2 t + e_t$$

The difference between these three regressions focuses on the presence of the defining terms a_0 and a_2t. The first form of specialization is a model which, as we have seen above (relationship 4.2), is derived from the random walk model, the second form contains a constant term (a_0) and the third regression form contains both a constant term and a linear time trend (t). The parameter which is most interesting in all regressions is the γ factor. If γ=0 then the series $\{y_t\}$ contains a root unit and therefore is non-stationary. Test is one-sided, so the alternative hypothesis is: H_0: γ<0 since we are not interested in case γ>0 because rejection of the null hypothesis and acceptance of the alternative, it certainly means rejection of the hypothesis γ>0.

In practice, the Augmented Dickey Fuller Test is used, which is based on the assumption that the test of the presence of a unit root can be performed not only by a first order autoregressive model but also by the use of an autoregressive model of higher order. If we consider that a time series is satisfactorily described through a autoregressive p-order model AR(p), it will be written in the following format:

$$Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + ... + a_{p-1} Y_{t-p+1} + a_p Y_{t-p} + e_t \qquad (4.3)$$

If we add and subtract the term $\alpha_p Y_{t-p+1}$ in the previous model (4.3), then we will find:

$$Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + ... + (a_{p-1} + a_p) Y_{t-p+1} - a_p \Delta Y_{t-p+1} + e_t$$

Then if we add and subtract the quantity $(\alpha_{p-1} + \alpha_p) Y_{t-p+1}$, we will take the following:

$$Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + ... + (a_{p-1} + a_p) Y_{t-p+2} - (a_{p-1} + a_p) \Delta Y_{t-p+2} - a_p \Delta Y_{t-p+1} + e_t$$

Following additions and subtractions, the following modified form will be obtained:

$$\Delta Y_t = a_0 + \gamma Y_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta Y_{t-i+1} + e_t \qquad (4.4)$$

where $\gamma = -\left(1 - \sum_{i=1}^{p} \alpha_i\right)$ and $\beta_i = -\sum_{j=i}^{p} a_j$.

And in this case, the most interesting factor is γ. If γ=0 then the series has a unit root and therefore is non-stationary. The stationarity test of the series is equivalent to testing the null hypothesis, H_0: γ=0 versus the alternative H_1: γ<0. To test the above hypothesis, the model in its modified form is estimated by the least squares method and the critical values of the Dickey Fuller (t) are used, which vary depending on the specialization of the model, if it contains a constant term or trend or both (and a constant term and trend). If t<τ, then, we reject the null hypothesis that our series is non-stationary and it is accepted that the series is stationary.

It is obvious that in order to carry out this test, the order (p) of the autoregressive procedure must be known, which is not known in advance before. In practice, what is done is to add so many time lags of Y_t, so that the residuals resulting from the application of the least squares method in the regression (4.4) are not autocorrelated.

4.2 Vector Autoregressive Models

Vector autoregression (VAR) is commonly used for short-run forecasting for a set of endogenous variables. In an autoregressive vector, each endogenous variable of the system is expressed as a function of itself with time lags p periods and p past values of all other endogenous variables of the system. The mathematical representation of the model VAR is:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + ... + A_P Y_{t-P} + \varepsilon_t$$

where Y_t is a k vector of endogenous variables, A_1,A_2, ….A_p are tables with data to be estimated and ε_t is the residuals` vector which can be correlated in the same time period, but we assume that they are not autocorrelated. Also, in addition to the cases concerning the disturbing conditions, the hypothesis of stationarity must be satisfied. This means that the vector Y_t has a constant mean and constant covariances, i.e. the covariances between Y_t and Y_{t+s} are not dependent on t but only on s for s=0,1,2... (i.e. they are independent of time t and only a function of displacement s). When all the above assumptions are applicable, the coefficients can be estimated by applying the method of least squares to each equation separately. The resulting estimators are consistent and asymptotically approximate to normal distribution.

4.2.1 Co-integration theory

When the data used in a regression model do not come from a stationary time series, then the problem of spurious regression occurs, so that the results of estimating such a model are of limited value and are likely to lead to false conclusions about the relationship between the time series under consideration. However, the problem of non-stationarity in the time series can be addressed if we use the first differences rather than its levels.

In many cases, in what economists are interested is the long-run relationship between the variables and this concerns their levels and not their differences. It has also been founded that tackling non-stationarity by taking differences may lead to a loss of information. The problem of the inability to use differences and the inadequacy of levels, when they are integrated, was addressed through the cointegration theory as developed by Engle and Granger(1987). In particular, Engle and Granger have shown that if two time series are integrated of order d and there is a linear combination of these two series which is stationary, then the series are cointegrated. More specifically, even if we have the time series X_{1t}, X_{2t}….X_{nt} which will be in a long-run economic equilibrium when:

$$\beta_1 X_{1t} + \beta_2 X_{2t} + ... + \beta_n X_{nt} = 0 \qquad (4.5)$$

If we consider the vectors $\beta = (\beta_1, \beta_2, .., \beta_n)$ και $X_t = (X_{1t}, X_{2t}, .., X_{nt})'$ then the above relationship (4.5) is written:

$$\beta X_t = 0$$

Furthermore, the relationship $e_t = \beta X_t$ represents the equilibrium error, i.e. the deviation from the long-run equilibrium. The variables $X_{1t}, X_{2t} .... X_{nt}$ will be called co-integrated and will be noted as $X_t \sim CI(d,b)$ if: (a) all variables are integrated of order d, i. e. non-stationary, which become stationary when we take differences of order d, (b) there is a vector $\beta' = (\beta_1, \beta_2, .., \beta_n)$ so that the linear combination $\beta' X_t = \beta_1 X_{1t} + \beta_2 X_{2t} + ... + \beta_n X_{nt}$ is integrated of order (d-b) where b>0. Vector b is called the cointegrating vector.

When the number of variables is only two, then the cointegrating vector is unique, while when the vector X_t includes more than two variables, then the cointegrating vector may not be and usually is not, unique. Generally when the vector X_t includes n elements then there can be up to n-1 linearly independent cointegrating vectors, the number of which is called cointegration order of X_t.

4.2.1.1 Cointegration test - The Engle-Granger method

Assuming we have the variables $Y, X_1, X_2 .... X_k$ and the following long-run equilibrium relationship applies: $Y_t = a_0 + a_1 X_{1t} + a_2 X_{2t} + .. + a_k X_{kt}$ which is also called cointegration regression.

If all the variables are first-order integrated, I(1), then they will be cointegrated if the linear combination: $u_t = Y_t - a_0 - a_1 X_{1t} - a_2 X_{2t} - ... - a_k X_{kt}$ is I(0), which means it's stationary. So the test for cointegration of the variables will be a test whether the residuals that resulted from the estimation of the above cointegration regression are stationary series. Assuming that the residuals are AR(1), that is, they are described by the relationship:

$$u_t = \rho u_{t-1} + e_t \qquad (4.6)$$

then if ρ=1 the residuals will not be a stationary series, whereas $-1 < \rho < 1$ they will be. If we subtract in relation (4.6) the quantity u_{t-1} from both sides we will have:

$$\Delta u_t = \rho_1 u_{t-1} + e_t \qquad (4.7)$$

where $\rho_1 = \rho - 1$ and therefore the stationarity test of the residuals series, i.e. the test of the null hypothesis $H_0 : \rho = 1$, now becomes test of the null hypothesis $H_0 : \rho_1 = 0$. For this test, compare the t statistic calculated on the basis of the results of applying the least squares method to the regression (4.7), with the critical values (t) of Dickey-Fuller but as modified by Engle-Granger. If $t < \tau_\rho$ then the null hypothesis that the residuals series is not stationary and therefore the variables are cointegrated, is rejected.

If the autoregressive procedure which we consider to describe our data, is of higher order, even AR(p), then the augmented Dickey-Fuller test is applied. That is ρ_1 is estimated by the relationship:

$$\Delta u_t = \rho_1 u_{t-1} + \sum_{i=1}^{p-1} \rho_{1i} \Delta u_{t-i} + e_t \qquad (4.8)$$

4.2.2 Error correction model and cointegration

The usefulness of the cointegration theory is that it enables us to develop Error Correction Models. In particular, the formulation of such models is based on or presupposes the existence of an equilibrium relationship between the variables. This means that the variables are cointegrated. Engle and Granger (1987) showed that if two variables are cointegrated, then the short-run imbalance relationship between them as well as the long-run equilibrium relationship, can always be formulated as an error correction model. This result is known as the Granger representation theorem.

The determinants of tax evasion, which are the subject of our analysis, may be removed in the short run, but in the long run they tend to move in a equilibrium with the size of tax evasion. The Error Correction model is able to inform us how the level of tax evasion depends on the changes of the other factors under consideration and at the same time to inform us about the imbalance error of the short run period and the way or the speed with which each factor is corrected to return to the long-run equilibrium relationship.

4.2.2.1 Error Correction Model

Suppose we have two variables for which the following long-run equilibrium relationship holds:

$$Y_t = a_0 + a_1 X_t \tag{4.9}$$

In accordance with the above defining (exact) relationship (4.9), it seems that in the long run there are no deviations which could be represented by some disturbance term. But since the variables are not always in equilibrium, what we are actually observed is an imbalance relationship such as the following:

$$Y_t = \beta_0 + \beta_1 \mathrm{X}_t + \beta_2 X_{t-1} + \gamma_1 Y_{t-1} + e_t \tag{4.10}$$

The above relationship (4.10) can be reformed and expressed as follows:

$$\Delta Y_t = \beta_1 \Delta X_t - (1-\gamma_1)(\mathrm{Y}_{t-1} - a_0 - a_1 X_{t-1}) + e_t \tag{4.11}$$

where $a_0 = \dfrac{\beta_0}{1-\gamma_1}$ and $\alpha_1 = \dfrac{\beta_1 + \beta_2}{1-\gamma_1}$

In this form (4.11) of the imbalance relationship, it is clear that the changes of Y depend on the changes of X and on the imbalance error of the previous period represented by the term $Y_{t-1} - a_0 - a_1 X_{t-1}$. The value of Y is corrected for the imbalance error of the previous period and for this reason the models of this form are also called Error Correction Models. The correction is partial, since its size depends on the factor γ_1 for which we assume it is greater than zero, but smaller than the unit. The imbalance error could be reported not only in the previous period, but also to a longer time lag.

The estimation of such a model is done by the method of Engle and Granger, in two stages. In the first stage, the cointegrating regression (4.9) is estimated by the least squares method and the residuals are calculated $u_t = Y_t - a_0 - a_1 X_t$. In the second step, the actual imbalance errors in the error correction model (i.e. the term $Y_{t-1} - a_0 - a_1 X_{t-1}$) are replaced by the estimated residuals, so that model (4.11) becomes:

$$\Delta Y_t = \beta_1 \Delta X_t - (1-\gamma_1) u_{t-1} + e_t \tag{4.12}$$

In the above model (4.12), all series are stationary since the series Y_t and X_t are cointegrated. Therefore the short-run parameters β_1 and γ_1 can be estimated by the least squares method. Thus we can estimate the imbalance error ($1-\gamma_1$) and the short-run effect of X_t on Y_t.

In this case, assuming that we want to estimate an error correction model to discover the causal relationship between the level of tax evasion y_t and one of its factors under investigation, the z_t, which are cointegrated, then the residuals from the equilibrium regression can be used to estimate the error correction model, which will take the following form:

$$\Delta y_t = a_0 + \sum_{i=1}^{p} a_i \Delta y_{t-i} + \sum_{i=1}^{p} \beta_i \Delta z_{t-i} + \theta_1 (y_{t-1} - k z_{t-1}) + u_{1t} \tag{4.13}$$

$$\Delta z_t = \gamma_0 + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + \sum_{i=1}^{p} \delta_i \Delta z_{t-i} + \theta_2 (y_{t-1} - kz_{t-1}) + u_{2t} \quad (4.14)$$

To estimate the above model, Engle and Granger suggested that the estimated residuals $\{e_{t-1}\}$ represent the deviation from the long-run equilibrium, the period $(t-1)$. So it makes sense to use them as a means of representing expression $y_{t-1} - kz_{t-1}$ in the equations. If this happens, then the error correction model will take the following form:

$$\Delta y_t = a_0 + \sum_{i=1}^{p} a_i \Delta y_{t-i} + \sum_{i=1}^{p} \beta_i \Delta z_{t-i} + \theta_1 e_{t-1} + u_{1t} \quad (4.15)$$

$$\Delta z_t = \gamma_0 + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + \sum_{i=1}^{p} \delta_i \Delta z_{t-i} + \theta_2 e_{t-1} + u_{2t} \quad (4.16)$$

From the estimation of these autoregression models with the least squares method, we can lead to useful conclusions about the relationships between the two time series under investigation.

Granger (1988) showed that in the context of an Error Correction Model (ECM), causality comes from two "sources". One source is through the time lags of differences, which captures the short-run effects of one time series on the other and the second through the error correction term, e_{t-1}, which represent the deviations from the long-run equilibrium relationship between the time series and the rate at which the imbalance error is corrected. Therefore, in order for there to be no causality according to Granger, both 'sources' should be 'closed' (non-significant). If even one of the two sources is 'open', then there is a causality relationship according to Granger, between the time series considered. In addition, the error correction term should be negative as this implies the correction of the short-run imbalance error to the long-run equilibrium. Thus, the Granger causality test in this case consists of the following:

In the first autoregressive model (4.15): In order for there to be no causality according to Granger from the factor under investigation to the level of tax evasion, it must be the $\beta_i = 0, i = 1,2,.., p$ and $\theta_1 = 0$, statistically. In other words, using a X^2- test, is tested the null hypothesis H_0: $\beta_i = 0, i = 1,2,.., p$ and with the help of the statistic t, the statistical significance of the term θ_1 is examined, i.e. the coefficient of the error correction term (e_{t-1}). If it is found that the joint statistical significance of the coefficients of the factor z is zero and at the same time the error correction term is insignificant, then there is no Granger causality from the factor z to the level of tax evasion y. However, if at least one of the above tests shows the opposite, then we should consider that there is a Granger causality from z to y .

In the second autoregressive model (4.16): In the same way, in order for there to be no causality according to Granger, from the level of tax evasion y to the factor z, should be $\gamma_i = 0, i = 1,2,.., p$ and $\theta_2 = 0$, statistically. Again we test the null hypothesis of the joint statistical significance of the coefficients of differences in the level of tax evasion y, H_0: $\gamma_i = 0, i = 1,2,.., p$ with the help of the X^2

statistic and at the same time we examine the statistical significance of the coefficient of the error correction term, H_0: $\theta_2=0$ through t-statistic.

The reason why the error correction term (e_{t-1}) is considered to be a source of causality, is because it arises from the relationship $e_{t-1} = y_{t-1} - kz_{t-1}$. It therefore contains within z which indirectly affects the formation of its prices y. In order not to have causality from z to y, it is necessary to consider that the influence of the term e_{t-1} is insignificant.

4.2.3 Innovation accounting

Innovation accounting is a useful tool to examine the relationship between economic variables. It is distinguished in the process of impulse response analysis and the variance decomposition. What makes the innovation accounting particularly important for the analysis of interactions between economic variables is that it provides a quantitative representation of the relationships and interactions between the under consideration variables, as opposed to the Granger causality analysis, which emerges from both the simple VAR model and the cointegration theory, which is more a qualitative approach to the issue of causal relationships.

4.2.3.1 Impulse response analysis

Impulse response function describes the effect on system variables, for a number of future periods, of a sudden shock to some of the endogenous variables.

Just as an autoregressive model AR(p), it can be written as a moving average model MA(∞), thus, a vector autoregressive VAR(p) is written as a vector moving average (VMA). Suppose we have two time series { y_t } and { z_t } which are the following bi-variable dynamic model:

$$y_t = b_{10} - b_{12} z_t + \gamma_{11} y_{t-1} + \gamma_{12} z_{t-1} + \varepsilon_{yt} \qquad (4.13)$$

$$z_t = b_{20} - b_{21} y_t + \gamma_{21} y_{t-1} + \gamma_{22} z_{t-1} + \varepsilon_{zt} \qquad (4.14)$$

in which we assume that the series { y_t } and { z_t } are stationary and errors ε_{yt} and ε_{zt} is white noise with constant standard deviations σ_y and σ_z , respectively, which are not autocorrelated. The structure of this model incorporates feedback as it allows to series { y_t } and { z_t } to affect each other, simultaneously. Using algebra matrix, we end up the bi-variable autoregressive model to be represented in a moving average model as follows:

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}^{(i)} \phi_{12}^{(i)} \\ \phi_{21}^{(i)} \phi_{22}^{(i)} \end{bmatrix} \begin{bmatrix} \varepsilon_{yt-i} \\ \varepsilon_{zt-i} \end{bmatrix}$$

or simpler,

$$x_t = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i} \qquad (4.15)$$

The above representation of the autoregressive model (4.15), through a moving average model, is particularly useful to examining the interactions between series { y_t } and { z_t }. The coefficients of ϕ_i can be used to examine the effect on endogenous variables of the system of a random, sudden disturbance in variables, which is present through shocks in the ε_{yt} and ε_{zt} . Usually, the shocks are expressed in terms of standard deviations of the disturbance terms ε_{yt} and ε_{zt} , that's half, one, two, and more standard deviations. For example, the coefficient $\phi_{12}^{(0)}$ is the instantaneous effect by of the change by a unit of its standard deviation of ε_{zt} on y_t. In the same way, the terms $\phi_{11}^{(1)}$ and $\phi_{12}^{(1)}$ is the effect it has on its current value y_t, the changes by a unit of standard deviations in ε_{yt-1} and ε_{zt-1}, respectively. However, this change in shocks will not only affect the y_t but also z_t, because y_t affects z_t as seen through the formulation of the autoregressive vector (4.18). In addition, we could say that the terms $\phi_{11}^{(1)}$ and $\phi_{12}^{(1)}$ represent the effects of changes by a unit of standard deviation of ε_{yt} and ε_{zt}, on y_{t+1}. The coefficients $\phi_{11}(i), \phi_{12}(i), \phi_{21}(i), \phi_{22}(i)$ are called impulse response functions.

4.2.3.2 Variance Decomposition

Let's consider the relationship again:

$$x_t = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i}$$

If we want to predict the x_t values that will have after n periods ahead, these will be given by the relationship:

$$x_{t+n} = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t+n-i}$$

And the and the n period forecast error, $x_{t+n} - E_t x_{t+n}$ will be:

$$x_{t+n} - E_t x_{t+n} = \sum_{i=0}^{n-1} \phi_i \varepsilon_{t+n-i}$$

If we focus on the series { y_t } of vector x_t, then we can say that the forecast error in n-step is:

$$y_{t+n} - E_t y_{t+n} = \phi_{11}(0)\varepsilon_{y_{t+n}} + \phi_{11}(1)\varepsilon_{y_{t+n-1}} + ... + \phi_{11}(n-1)\varepsilon_{y_{t+1}} + \phi_{12}(0)\varepsilon_{z_{t+n}} + \phi_{12}(1)\varepsilon_{z_{t+n-1}} + \\ + ... + \phi_{12}(n-1)\varepsilon_{z_{t+1}}$$

If we symbolize the variation of the forecast error y_{t+n}, as $\sigma_y(n)^2$ then we'll have:

$$\sigma_y(n)^2 = \sigma_y^{\,2}\left[\phi_{11}(0)^2 + \phi_{11}(1)^2 + ... + \phi_{11}(n-1)^2\right] + \sigma_z^{\,2}\left[\phi_{12}(0)^2 + \phi_{12}(1)^2 + .. + \phi_{12}(n-1)^2\right]$$

It's possible to decompose the variation of the forecast error and find the part of the variance $\sigma_y(n)^2$ due to series { ε_{yt} }and{ ε_{zt} }. These ratios will be:

$$\frac{\sigma^2{}_y[\phi_{11}(0)^2 + \phi_{11}(1)^2 + ... + \phi_{11}(n-1)^2]}{\sigma_y(n)^2}$$

and

$$\frac{\sigma^2{}_z[\phi_{12}(0)^2 + \phi_{12}(1)^2 + ... + \phi_{12}(n-1)^2]}{\sigma_y(n)^2}$$

The variance decomposition of the forecast error informs us about the effect of the variance of a particular variable on the variance of the forecast error of the change of another variable of the system. Therefore it provides a substantial indication of how strong is the effect of each variable. According to this method, the unexplained part of each variable, i.e. the error of each variable, is individually subject to a shock and the effect of this disorder on the individual variables of the system is studied. In other words with the decomposition of the variance we find what percentage of the variance of the forecast error of the change of the values of a time series, is interpreted by a shock in another variable of the system

To apply both variance decomposition and impulse response functions it requires a transformation of the system into such a form that errors are rectangular. In this way, the effects of the individual shocks on the variables are identified. For this rectangular transformation, the Choleski decomposition, can be used which is the most common one and is based on the assumption that the system of structural equations from which the autoregressive model is derived, is periodic. On the basis of the Choleski decomposition, we set the autoregressive model so that its simultaneous values y_t do not have a simultaneous effect on z_t. This restriction is represented on the autoregressive vector by imposing $b_{21} = 0$ .This means that:

$$e_{1t} = \varepsilon_{yt} - b_{12}\varepsilon_{zt}$$
$$e_{2t} = \varepsilon_{zt} \qquad (4.16)$$

Therefore, the relationship (4.16) shows that all observed errors $\{e_{2t}\}$ attributable to disturbances ε_{zt} . This decomposition defines the system in such a way that a shock to the error ε_{yt} do not directly affect series values $\{z_t\}$. However, there are indirect effects resulting from its prices y_t with time delays, which affect its simultaneous price z_t . Also, a shock to error ε_{zt} effect directly on e_{1t} and e_{2t} , unlike in error ε_{yt} which does not affect e_{2t} . The z_t precedes it y_t . What is achieved through this transformation is that the residuals are unrelated.

5 Empirical results

5.1 Results of Augmented Dickey - Fuller Tests

Tables 1 and 2 below show the results of the Augmented Dickey - Fuller Test (ADF test) of the null hypothesis of unit root at the levels of the time series and their first differences. For the autocorrelation test of the residuals we use the Breusch-Godfrey test which belongs to the category of LM tests (Lagrange Multiplier Tests) and can be done with the distribution X^2 .

The Augmented Dickey - Fuller test for the non-stationarity of time series showed that the null hypothesis of the existence of a unit root at the levels of the variables cannot be rejected for all time series under consideration, which are considered to be non-stationary at the levels. On the contrary, the unit root hypothesis is rejected in the first differences in variable values in all cases. This means that the under consideration time series are integrated of first order I(1) and contain a unit root.

Table 1: Unit root tests (on the level series)

Variable	Lag Length	ADF test	p-value	LM test (X^2)	p-value	Deterministic Terms
TE	1	0.488120	0.8125	1.602981	0.4487	None
TR	0	-2.976808	0.1593	5.193718	0.0745	Trend and Intercept
U	1	-2.277506	0.1873	0.725061	0.6959	Intercept
RL	0	-0.263225	0.5802	0.166109	0.9967	None
PD	0	-2.493850	0.1298	0.165367	0.9206	Intercept
S	1	0.589596	0.8360	2.691412	0.2604	None
L	0	0.871261	0.8908	2.028022	0.3628	None
GE	0	-2.704634	0.2437	2.196024	0.3335	Trend and Intercept
GDP	1	0.415843	0.7945	2.996756	0.2235	None
SE	0	-2.067145	0.2586	0.388597	0.8234	Intercept
EG	0	0.030133	0.6822	2.722950	0.2563	None
ED	0	-2.877598	0.0635	1.446843	0.4851	Intercept
CE	0	0.2001818	0.7357	0.392488	0.8218	None
C	1	-2.124860	0.2375	3.606762	0.1647	Intercept

Table 2: Unit root tests (on the 1st difference series)

Variable	Lag Length	ADF test	p-value	LM test (X^2)	p-value	Deterministic Terms
TE	0	-2.569525*	0.0128	0.857209	0.6514	None
TR	0	-12.35462*	0.0000	7.691513	0.1036	Trend and Intercept
U	0	-1.733258**	0.0786	3.410276	0.1817	None
RL	0	-4.720237*	0.0001	1.069766	0.5857	None
PD	0	-3.798238*	0.0094	0.524987	0.7691	Intercept
S	0	-2.025650*	0.0433	2.613850	0.2707	None
L	0	-3.241841*	0.0025	0.00000	1	None
GE	5	-4.026190*	0.0288	5.525210	0.0631	Trend and Intercept
GDP	0	-1.962964*	0.0494	2.680313	0.2618	None
SE	0	-2.845869*	0.0066	3.046092	0.2180	None
EG	0	-6.491423*	0.0000	0.094138	0.9540	None
ED	0	-3.086463*	0.0425	0.290009	0.8650	Intercept
CE	0	-3.914717*	0.0004	1.291298	0.5243	None
C	0	-3.944323*	0.0004	0.160472	0.9229	None

Notes: The lag length is modeled by Akaike's automatic selection of maximum lags from zero to five lags.
Test critical values at the 5% and 10% level are based on MacKinnon (1996).
* *Indicates statistical significance at 5%.* ** *Indicates statistical significance at 10%.*

5.2 Results of cointegration tests

The cointegration tests between tax evasion and its determinants under investigation, the results of which are presented in Table 3, indicate that in all cases they fail to reject the null hypothesis of non-cointegration of the variables under consideration factors with tax evasion at a bi-variable level, as well as the results of ADF test for stationarity of the residuals series, on a case by case basis, showed that the null hypothesis of the existence of the unit root is rejected and consequently the residuals series is stationary in all cases. In view of these results, we can proceed to the development and estimation of Error

Correction Models to reveal any equilibrium relationship between tax evasion and its determinants under investigation.

Table 3: Test of cointegration between tax evasion (TE) and deterministic factors

Variable	Lag Length	ADF test	p-value	Deterministic Terms	Result
TR	1	-2.667093*	0.0102	None	Cointegrated
U	0	-2.3046*	0.0234	None	Cointegrated
RL	0	-1.951718**	0.0505	None	Cointegrated
PD	1	-1.988342*	0.0468	None	Cointegrated
S	1	-1.919551**	0.0541	None	Cointegrated
L	1	-2.874053*	0.0062	None	Cointegrated
GE	0	-2.448961*	0.0168	None	Cointegrated
GDP	1	-1.717092**	0.0812	None	Cointegrated
SE	1	-1.956137**	0.0501	None	Cointegrated
EG	1	-2.413992*	0.0183	None	Cointegrated
ED	1	-2.056014*	0.0406	None	Cointegrated
CE	1	-2.186477*	0.0306	None	Cointegrated
C	1	-3.126573*	0.0034	None	Cointegrated

Notes: The lag length is modeled by Akaike's automatic selection of maximum lags from zero to five lags. Test critical values at the 5% and 10% level are based on MacKinnon (1996).
** Indicates statistical significance at 5%. ** Indicates statistical significance at 10%.*

5.3 Estimation results of Error Correction Models

Table 4 presents the estimation results of Error Correction Models between tax evasion and its exploratory determinants factors, which show the short-run imbalance relationship and the long-run equilibrium relationship between them, the coefficient of error correction term (ECT) which represent the short-run imbalance error and the way or speed with which it is corrected to reach the long-run equilibrium relationship again.

More specifically, Table 4 presents the results of the short-run impact of the independent variable on the dependent variable, by testing the null hypothesis that the coefficients of the lags of the independent variable are jointly zero, by using the X^2statistic and estimating the statistical significance of the error correction term (ECT) through the t-statistic. It is noted that the error correction term is required to be not only statistically significant but also negative. The last column of Table 4 below shows the causality relationships between the tax evasion in Greece and its determinants under investigation, which emerged by the estimation of the Error Correction Models.

Table 4: Error Correction Models (ECM) and Causality tests

ECM Variables	Dependent Variable	Lags	Wald test X^2-statistic	ECT	t-statistic	Causality
TE, TR	TE	2	7.227998 (0.0269)*	-0.389878	-1.623141 (0.1269)	**TR Granger cause TE short-run**
	TR	2	0.808476 (0.6675)	0.085544	0.726651 (0.4794)	TE does not Granger cause TR
TE, U	TE	3	12.17741	-1.004794	-2.294066	**U Granger cause**

			(0.0068)*		(0.0406)*	**TE short-run and long-run**
	U	3	3.441916 (0.3284)	0.653203	0.601704 (0.5586)	TE does not Granger cause U
TE, RL	TE	1	0.633754 (0.4260)	-0.249561	-2.035406 (0.0568)**	**RL Granger cause TE long-run**
	RL	1	0.120685 (0.7283)	-0.203925	-0.798602 (0.4349)	TE does not Granger cause RL
TE, PD	TE	1	0.631353 (0.4269)	-0.005288	-0.082731 (0.9349)	PD does not Granger cause TE
	PD	1	3.207734 (0.0733)**	-0.192964	-4.614020 (0.0002)*	**TE Granger cause PD short-run and long-run**
TE, S	TE	2	1.967579 (0.3739)	-0.149073	-0.549788 (0.5906)	S does not Granger cause TE
	S	2	0.014104 (0.9930)	-0.700592	-3.268527 (0.0052)*	**TE Granger cause S long-run**
TE, GDP	TE	2	11.66522 (0.0029)*	-1.178323	-5.878717 (0.000)*	**GDP Granger cause TE short-run and long-run**
	GDP	2	2.487314 (0.2883)	0.505010	3.916161 (0.0016)	TE does not Granger cause GDP
TE, L	TE	2	0.966064 (0.6169)	-0.518927	-2.699703 (0.0165)*	**L Granger cause TE long-run**
	L	2	0.904019 (0.6363)	0.105933	0.110299 (0.9136)	TE does not Granger cause L
TE, GE	TE	5	151.1602 (0.000)*	-2.145380	-3.031210 (0.0290)*	**GE Granger cause TE short-run and long-run**
	GE	5	3.036824 (0.6943)	-8.680793	-1.338744 (0.2383)	TE does not Granger cause GE
TE, EG	TE	3	2.721864 (0.4365)	-0.171070	-0.816648 (0.4315)	EG does not Granger cause TE
	EG	3	10.19398 (0.0170)*	-0.531369	-3.395682 (0.0060)*	**TE Granger cause EG short-run and long-run**
TE, C	TE	1	2.867433 (0.0904)**	-0.046244	-0.541129 (0.5951)	**C Granger cause TE short-run**
	C	1	0.155441 (0.6934)	0.220965	2.728534 (0.0138)	TE does not Granger cause C
TE, SE	TE	3	5.759149 (0.1239)	-0.307913	-1.055684 (0.3137)	SE does not Granger cause TE
	SE	3	28.07614 (0.000)*	-0.235211	-4.329843 (0.0012)*	**TE Granger cause SE short-run and long-run**
TE, ED	TE	4	6.093578 (0.6368)	-0.013767	-0.168233 (0.8701)	ED does not Granger cause TE
	ED	4	5.062711	0.098462	2.938486	TE does not

			(0.7509)		(0.0165)	Granger cause ED
TE, CE	TE	2	16.02086 (0.0030)*	-0.458770	-3.807894 (0.0017)*	**CE Granger cause TE short-run and long-run**
	CE	2	8.805027 (0.0662)**	-0.096812	-3.015576 (0.0087)*	**TE Granger cause CE short-run and long-run**

Notes:a. The null hypothesis is that independent variable does not Granger cause dependent variable.
b. The lag length is modeled by Akaike's automatic selection of maximum lags from zero to five lags.
c. The x^2-test is calculated for the joint hypothesis that the coefficients of the lags of the independent variable are zero.
d. The figures in parentheses denote the p-value of test
e. ECT is the Error Correction Term
** Indicates statistical significance at 5%. ** Indicates statistical significance at 10%.*

5.4 Results of variance decomposition

As mentioned above, the variance decomposition of the forecast error informs us about the effect of the variance of a particular variable on the variance on the forecast error of the change of another variable of the system. In this case, the uninterpreted part of each of the system variables (at a bi-variable level), i.e. the error of each variable, is individually subject to a shock and the effect of this disorder on the other variable of the system is studied.

Tables 5[a] and 5[b] below presents the results of the variance decomposition by Choleski, of the forecast error of tax evasion and its determinants under investigation, following the imposition of a shock on the forecast error of a system variable (at a bi-variable level) and the effect of this disorder on the other variable of the system, for a period of 10 years. The study of the results emerge the following useful conclusions:

1. An individual shock in the level of tax rates (TR), unemployment (U), Rule of Law index (RL), gross domestic product (GDP), non-performing loans index (L), government efficiency index (GE) and the level of final consumption expenditure (CE), explains a significant percentage of the variance in the forecast error of the change in the level of tax evasion (TE) in Greece, indicating a significant impact of the above determinants on the level of tax evasion.

2. A sudden shock in the other determinants of tax evasion fails to explain to a significant degree, the variance of the forecast error of the change in the level of tax evasion (TE) in Greece, as the corresponding percentage are low.

3. A sudden shock in the level of tax evasion in Greece (TE) explains a significant percentage of the variance of the forecast error of the change in the level of public debt (PD), private sector savings (S), gross domestic product (GDP), central government expenditure (EG), corruption perception index (C), the level of the self-employed (S), the level of education (ED) and the final consumption expenditure of households (CE), which indicate a significant impact of tax evasion on the levels of the determinants above, short-run or long-run as appropriate.

4. A shock in the level of tax evasion is unable to interpret to a significant extent the variance of the forecast error of the change on the other determinants, as the corresponding percentages are low.

5.5 Impulse response analysis

The impulse response analysis examines the effect on endogenous variables of the system, a random, sudden shock in the variables. A change in the i, for example, variable of the VAR model will not only affect the specific variables, but will also be transmitted to the other endogenous variables of the system through the dynamic form of the model VAR itself, in which the other variables appear to affect the specific variable through their time lags. So the impulse response functions show us the effect that a shock will have on one of the innovations on current and future values of the endogenous variables, each time. It is also through this analysis that we can see how long a shock last and impact the equilibrium.

The coefficients determining the size of the response of the each variable to a shock during period 0, to the other variable of the system for each period, are derived from the VAR model coefficients for which we have only estimates. They will therefore also be subject to a statistical error. Therefore, in order to estimate the response of each variable to an unpredictable change in the other variable of the bi-variable system, it is useful to calculate a two-sided confidence interval. The confidence interval is calculated empirically by the Monte Carlo simulation method, due to the complexity of the impulse response function. If this two-sided confidence interval does not include the zero value, only then can it be argued that the relevant variable, in terms of its significant effect, has a statistically significant non-zero effect on the forecast error of the interpreted variable in the corresponding period.

The following figures 1-13 graphically illustrate the impulse response functions and responses of each variable in question to unpredictable changes in the other system variable, size of one standard deviation. It is noted that each system consists of two variables, one of which is the level of tax evasion and the other one of the factors under investigation of tax evasion. In order to investigate the statistical significance of these responses, each figure also shows the corresponding approximate limits of the confidence interval at a confidence level of 95%. We also set the number of iterations for the application of the Monte Carlo simulation method to 100. From the evaluation of the results of the imposition of impulse response functions, the following conclusions can be drawn:

1. A sudden individual shock in the level of tax rates TR (s.s. first 3 years), unemployment U (s.s. first 5 years), Rule of Law index RL (s.s. first 6 years), Gross Domestic Product GDP (s.s. first 4 years), non-performing loans L (s.s. first 5 years), Government Efficiency Index GE(s.s. 4th -7th year), Corruption Perceptions Index C (s.s. 2nd - 6th year) and final consumption expenditure CE (s.s. 3rd - 8th year), size of one standard deviation, has strong effect on the forecast error of the level of tax evasion in Greece, which is statistically significant (s.s.) during the periods mentioned above in parentheses.

2. A sudden individual shock in the level of public debt PD (s.s. 7th - 17th year) and government expenditure EG (s.s. 4th - 7th year), size of one standard deviation, has weak effect on the forecast error of the level of tax evasion in Greece, which are statistically significant (s.s.) during the periods mentioned above in parentheses.

3. A shock in the other determinants under consideration, does not cause statistically significant upheavals in the size of tax evasion in Greece.

4. A shock in the size of tax evasion TE in Greece, size of one standard deviation, causes strong effect on the size of the public debt PD (s.s. 2nd-5th year), the level of savings of individuals S (s.s. first 5 years) and the number of self-employed SE (s.s. 2nd-3rd year), statistically significant (s.s.) during periods indicated in parentheses. In addition, shocks in the size of tax evasion size of one standard deviation, have a weak effect on the size of the tax rates TR (s.s. first 4 years), the level of government expenditure EG (s.s. 6th year) and final consumption expenditure CE (s.s. 1st year).

Conclusions

The purpose of this paper is to investigate the degree of influence of a set of factors on the level of tax evasion in Greece and to discover the causality relationships between them. In particular, this survey made a significant effort to examine the short-run and long-run relationships between the size of tax evasion and the factors under investigation related to both macroeconomic factors such as the level of tax rates, the size of unemployment, the GDP, as well as socio-economic factors such as the Rules of Law index, the level of non-performing loans, private sector savings, government efficiency, the size of the public debt, the level of government spending, the corruption perception index, the number of self-employed, the final consumption expenditure of individuals and the level of education. For this purpose, annual observations were used, for the period from 1995 to 2018.

This study was based on the cointegration analysis to examine the causal relationships between the level of tax evasion and the factors under investigation, at both short-run and long-run period. The results obtained were illuminating of the interdependence relationships between the variables. We also proceeded to the variance decomposition and the estimation of the impulse response functions with the aim of quantifying the relationships and the duration of the interactions. From the empirical analysis of the behavior of the examined determinants of tax evasion, a number of conclusions emerge.

In particular, the application of the Engle - Granger method and the development of Error Correction Models revealed significant interactions between the examined factors and the size of tax evasion in Greece, as were discovered both short-run interactions and causality relationships that stem from the long-run relationship of the variables and the cointegration relationship. In particular, the results of the implementation of the cointegration theory showed that the level of tax rates, the level of unemployment, the Rule of Law index, the level of the GDP, the level of non-performing loans, the government efficiency, the perception index of corruption and the level of final consumption expenditure, affect the size of tax evasion in Greece.

At the level of economic analysis, the above results could be interpreted as follows:

a) The level of tax rates is related to the size of the tax burden and the amount of disposable income of taxpayers, thus influencing their attitude towards tax evasion.

b) The effectiveness of the government gives the taxpayers the feeling that the taxes they pay, as public revenues, are spent effectively, their needs are largely met by the state and the public goods they enjoy are sufficient. This creates a relationship of trust between the state and the taxpayers that affects the level of voluntary tax compliance.

c) A widespread corruption affects taxpayers' perceptions of institutions and the state, which in turn encourages tax evasion.

d) The increase in income at national level leads to a higher level of economic development of the country, which in turn increases the tax compliance rate.

e) An increase in unemployment rate will lead to a decrease in taxpayers' income, which in turn will increase tax evasion. In addition, an increase in cash use, caused by rising unemployment rate, as individuals may turn to the underground economy, may lead to tax collection problems.

f) The increased adherence to the rules of law creates a social model according to which non-compliance behaviors among taxpayers are not tolerated by society.

g) The improvement of the business climate as reflected in the reduction of the non-performing loans ratio, indicates increased economic activity of companies and possibly higher profitability, which in turn reduces the uncertainty of enterprises and improves their investment prospects, changing their attitude towards the tax evasion.

It was also found that the level of tax evasion in Greece has significant effects on the level of public debt, private sector savings, government spending, the number of self-employed people and final consumption expenditure in the country.

In view of these findings, reducing the level of tax evasion will increase public revenues and reduce the need for external borrowing, thus reducing the level of public debt in the long run. Also, the reduction of tax evasion will lead to a reduction of the tax burden, through the rationalization of tax rates and the abolition of extraordinary tax burdens, leading to an increase in disposable income and household savings. Similarly, a low level of tax evasion will increase public revenues, allowing the state to increase public spending, both at the socio-economic level and at the level of public investment. In addition, the level of tax evasion seems to be linked to the number of the self-employed, mainly through the creation of unequal conditions of competition.

Subsequently, the variance decomposition of the forecast error of tax evasion and of its determinants under investigation, confirms the results of the previous method, as the imposition of an individual shock in the forecast error of a system variable (at a bi-variable level) interprets a significant percentage of the forecast error of the other variable. Regarding to the estimation of the impulse response functions, we point out that the results of the analysis of variance decomposition are confirmed and in addition, are provided evidence of a quantitative representation of the causal relationships of the factors concerned with the size of the tax evasion and the time required for the effect of the measures that may be taken on the above factors of tax evasion, to reduce it.

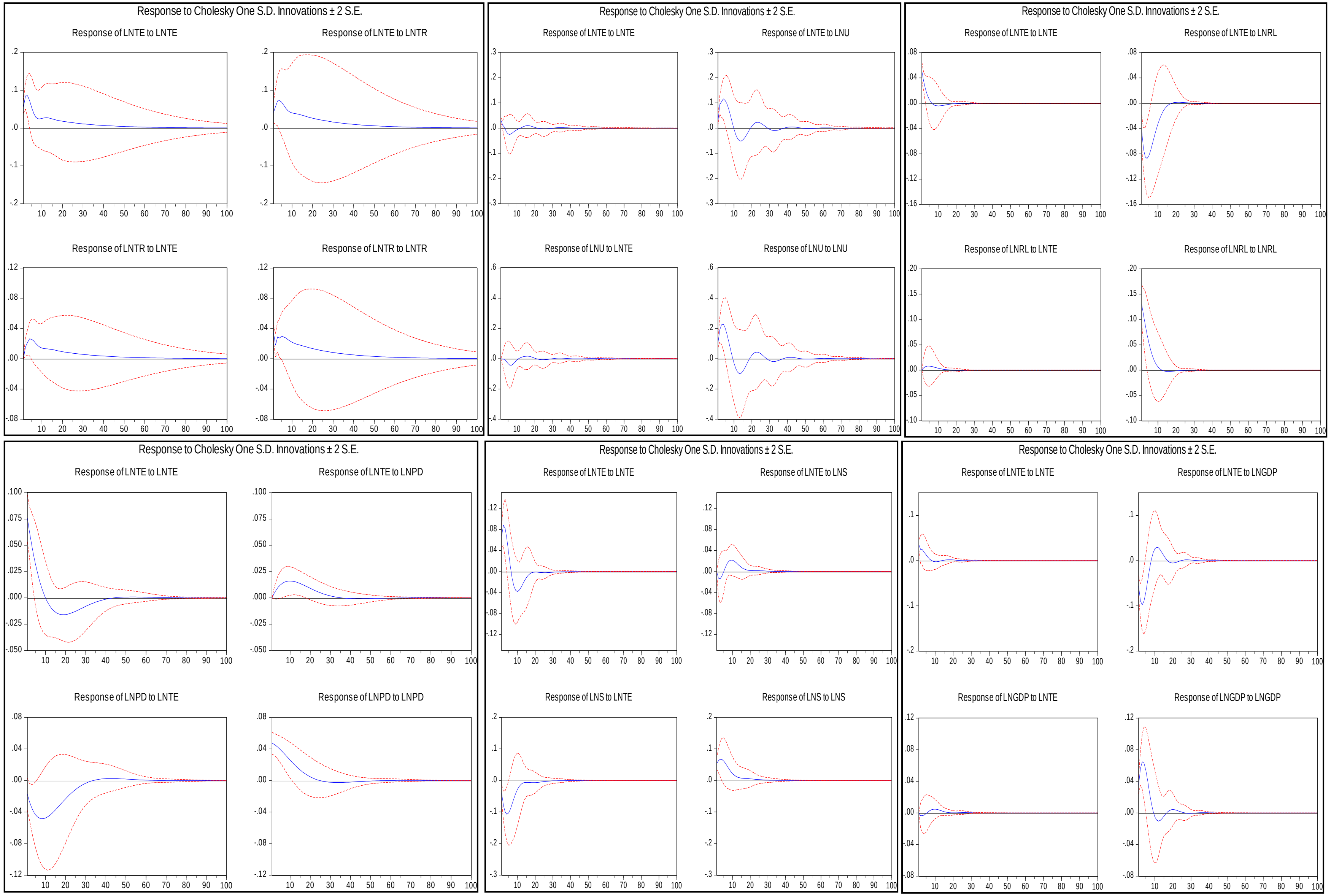
In view of the above results, the Greek tax authorities must, in pursuing the tax policy, take into account of the above critical, as it turned out, macroeconomic and socio-economic factors, both in terms of legislative intervention that aim to improve the level of the above factors, through appropriate fiscal and monetary policy tools, as well as in the evaluation of the time required for the exercise of the policies that will shape the level of these factors and their influence on the size of tax evasion. Furthermore, it is necessary to take into account the importance of the study in how the size of tax evasion affects a number of factors such as the size of public debt and the time lags observed in the exercise of any individual tax evasion reduction policies, on the magnitude of these factors.

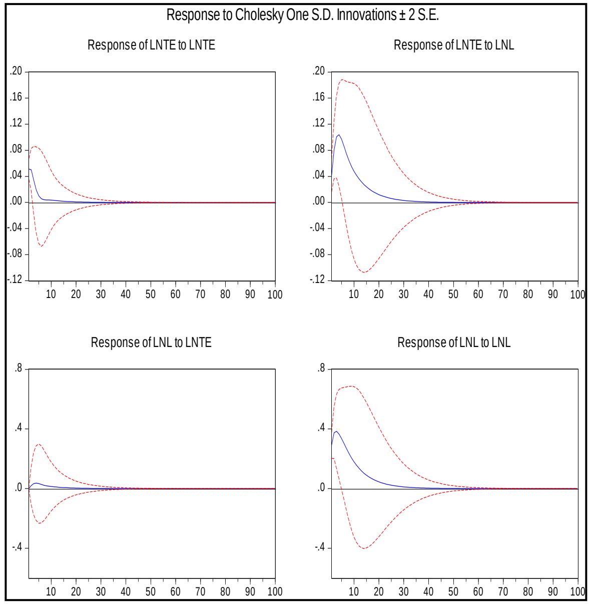
Table 5^{α}: Variance Decomposition (VD) as percentage of forecast error variance

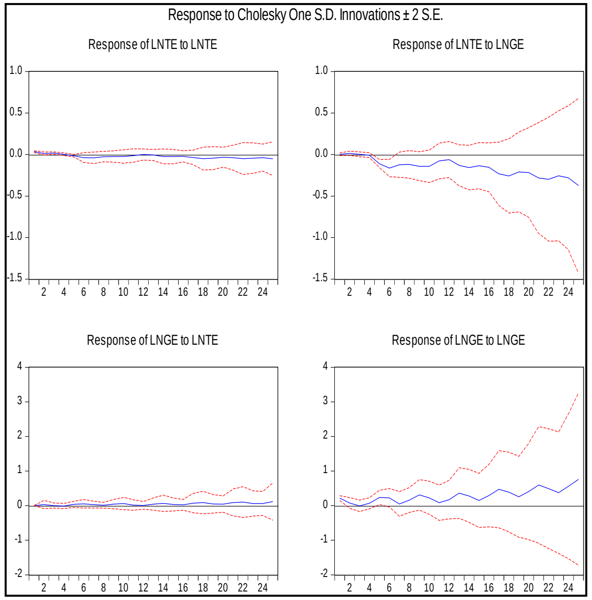
Year	**TR shock**	**TE shock**	**U shock**	**TE shock**	**RL shock**	**TE shock**	**PD shock**	**TE shock**	**S shock**	**TE shock**	**GDP shock**	**TE shock**
	VD of TE	**VD of TR**	**VD of TE**	**VD of U**	**VD of TE**	**VD of RL**	**VD of TE**	**VD of PD**	**VD of TE**	**VD of S**	**VD of TE**	**VD of GDP**
1	69.81796	0.000000	41.42050	0.000000	51.31653	0.000000	0.000000	5.322483	0.000000	50.76433	40.76431	0.000000
2	63.06372	2.747570	89.86378	0.193393	65.25851	0.851273	0.936001	33.14818	1.008961	74.68938	71.65158	3.332055
3	65.41761	4.900736	94.87403	0.500017	77.19627	1.720281	1.230580	53.88180	0.524906	87.87674	81.16142	11.65193
4	70.58309	6.702140	97.17229	0.301176	83.74556	2.292359	1.414162	67.96894	0.438815	93.39497	77.19332	19.15082
5	75.67093	7.304691	98.06681	0.330894	87.23317	2.656436	1.522090	76.92949	0.367877	95.76669	70.43135	26.15470
6	79.06431	7.041269	98.39475	0.484290	89.24469	2.906622	1.591686	82.77984	0.347583	96.83374	64.63190	31.91251
7	80.60889	6.392196	98.29813	0.884012	90.51291	3.091684	1.638323	86.72639	0.358832	97.52213	61.27258	35.99507
8	81.25926	5.728168	97.58192	1.679222	91.37355	3.235098	1.670857	89.48354	0.361223	97.94642	60.27836	38.40381
9	81.91358	5.191656	96.56313	2.451135	91.99158	3.349608	1.694253	91.47156	0.381840	98.17303	60.34295	39.55694
10	82.65556	4.788304	95.95785	2.958624	92.45509	3.443111	1.711512	92.94544	0.455206	98.26090	60.38368	39.95892

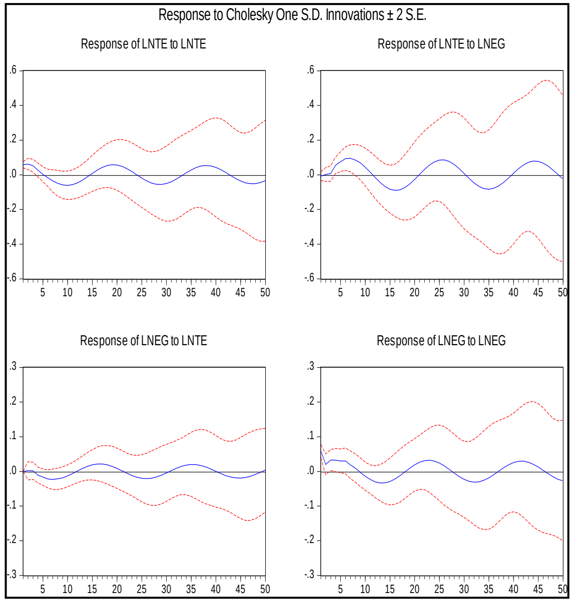
Table 5^{b}: Variance Decomposition (VD) as percentage of forecast error variance

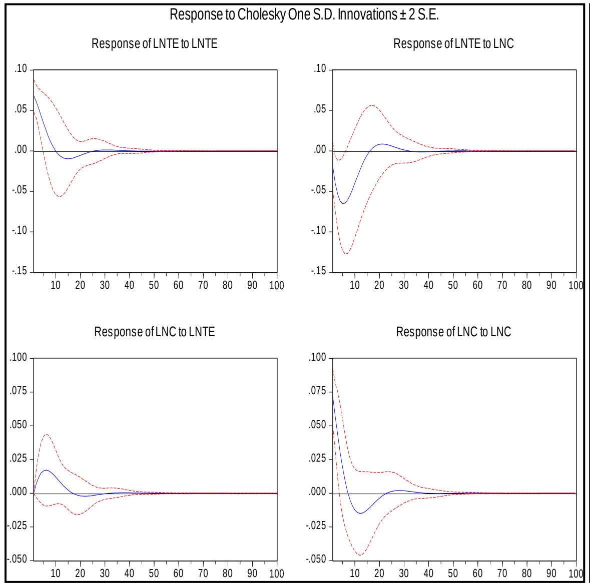
Year	**L shock**	**TE shock**	**GE shock**	**TE shock**	**EG shock**	**TE shock**	**C shock**	**TE shock**	**SE shock**	**TE shock**	**ED shock**	**TE shock**	**CE shock**	**TE shock**
	VD of TE	**VD of L**	**VD of TE**	**VD of GE**	**VD of TE**	**VD of EG**	**VD of TE**	**VD of C**	**VD of TE**	**VD of SE**	**VD of TE**	**VD of ED**	**VD of TE**	**VD of CE**
1	50.97129	0.000000	20.57704	0.000000	0.000000	0.912357	6.120086	0.000000	0.000000	4.975070	0.000000	39.81671	0.000000	1.824724
2	73.30245	0.014652	20.58142	14.22558	0.938198	9.534477	12.75543	1.016946	2.989807	8.485387	0.563064	61.87285	1.932088	11.54151
3	87.52118	0.026805	45.46199	12.92027	0.663709	10.17377	13.25221	8.945011	2.124909	13.54597	1.322395	61.45772	4.015915	22.70731
4	92.57095	0.040135	73.18821	14.17425	1.920415	19.72209	11.60115	24.71108	1.768540	27.03236	2.295691	64.68960	30.03162	25.85281
5	94.69007	0.050545	97.16051	8.646147	1.812753	44.45656	9.840320	39.41796	1.641761	59.57115	5.005010	67.16748	54.21043	29.04512
6	95.74229	0.057663	93.92698	12.95306	1.809973	64.08425	8.573708	49.61388	1.676163	75.77716	6.907688	71.45747	62.64694	32.16357
7	96.37244	0.062019	89.87258	10.88857	1.944271	73.66518	7.682590	56.46954	1.808461	84.15276	7.950841	74.93205	63.98761	33.65121
8	96.80579	0.064599	90.07651	12.90083	1.980027	79.40109	6.966573	61.24292	1.699914	88.61777	8.133612	78.52673	62.91937	34.30910
9	97.13435	0.066197	90.23765	12.34909	2.333638	82.61119	6.353309	64.75033	1.584600	91.02623	8.020513	80.93551	61.62950	34.68446
10	97.40056	0.067315	89.52236	14.38082	2.572768	84.63408	5.834283	67.52714	1.500005	92.06496	7.853671	82.85496	60.84553	34.73557

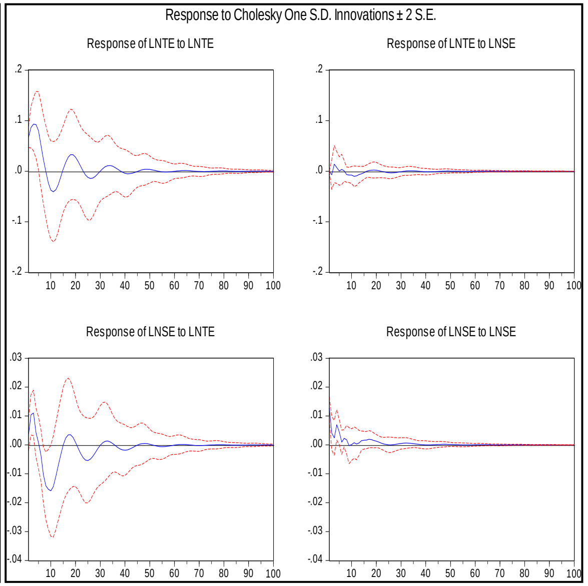

Response to Cholesky One S.D. Innovations ± 2 S.E.
Response of LNTE to LNTE
Response of LNTE to LNTR
Response of LNTR to LNTE
Response of LNTR to LNTR
Response to Cholesky One S.D. Innovations ± 2 S.E.
Response of LNTE to LNTE
Response of LNTE to LNU
Response of LNU to LNTE
Response of LNU to LNU
Response to Cholesky One S.D. Innovations ± 2 S.E.
Response of LNTE to LNTE
Response of LNTE to LNRL
Response of LNRL to LNTE
Response of LNRL to LNRL
Response to Cholesky One S.D. Innovations ± 2 S.E.
Response of LNTE to LNTE
Response of LNTE to LNPD
Response of LNPD to LNTE
Response of LNPD to LNPD
Response to Cholesky One S.D. Innovations ± 2 S.E.
Response of LNTE to LNTE
Response of LNTE to LNS
Response of LNS to LNTE
Response of LNS to LNS
Response to Cholesky One S.D. Innovations ± 2 S.E.
Response of LNTE to LNTE
Response of LNTE to LNGDP
Response of LNGDP to LNTE
Response of LNGDP to LNGDP

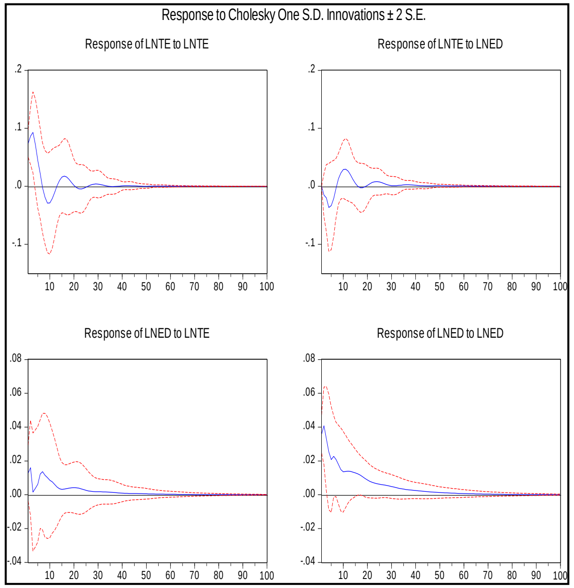

Response to Cholesky One S.D. Innovations ± 2 S.E.
Response of LNTE to LNTE
Response of LNTE to LNL
Response of LNL to LNTE
Response of LNL to LNL

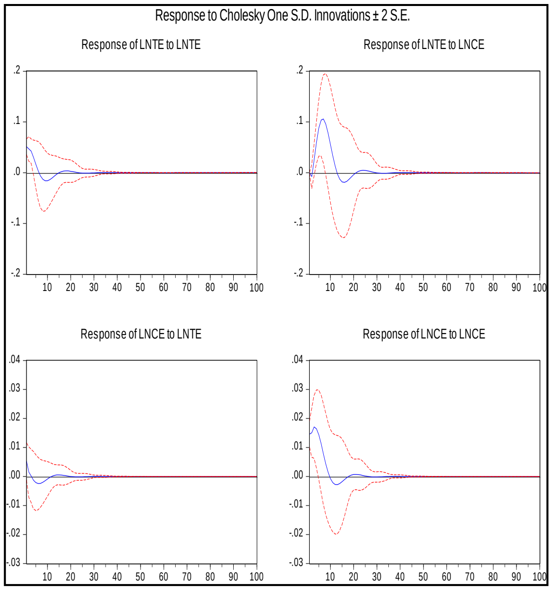

Response to Cholesky One S.D. Innovations ± 2 S.E.
Response of LNTE to LNTE
Response of LNTE to LNGE
Response of LNGE to LNTE
Response of LNGE to LNGE


Response to Cholesky One S.D. Innovations ± 2 S.E.
Response of LNTE to LNTE
Response of LNTE to LNEG
Response of LNEG to LNTE
Response of LNEG to LNEG


Response to Cholesky One S.D. Innovations ± 2 S.E.
Response of LNTE to LNTE
Response of LNTE to LNC
Response of LNC to LNTE
Response of LNC to LNC


Response to Cholesky One S.D. Innovations ± 2 S.E.
Response of LNTE to LNTE
Response of LNTE to LNSE
Response of LNSE to LNTE
Response of LNSE to LNSE


Response to Cholesky One S.D. Innovations ± 2 S.E.
Response of LNTE to LNTE
Response of LNTE to LNED
Response of LNED to LNTE
Response of LNED to LNED



Figures 1-13: Impulse Response Analysis

References

Allingham, M. and Sandmo, A. (1972). Income Tax Evasion: A Theoretical Analysis, Journal of Public Economics, Vol. 1, 323-338.

Alm, J., McClelland, H. and Schulze, W. (1992). Why Do People Pay Taxes?, Journal of Public Economic, Vol. 48, 21-28.

Alm, J. And Yunus, M. (2009). Spatially and Persistence in U.S. Individual Income Tax Compliance, National Tax Journal, Vol. 62(1), 101-24

Anastasiou A., Kalamara E. and Kalligosfyris C. (2020). Estimation of the size of tax evasion in Greece, Bulletin of Applied Economics, Vol. 7(2), 97-107

Andreoni, J., Erard, B. and Feinstein, J. (1998). Tax Compliance, Journal of Economic Literature, Vol. 36, 818-860

Baldry, J.C. (1987). Income Tax Evasion and the Tax Schedule: Some Experimental Results, Public Finance, Vol. 42(3), 357-83

Boame, A. (2009). Individual Tax Compliance: A Time-Series Regression Using Canadian Data 1987-2003, Compliance Research and Measurement Section, Baseline Research Paper.

Brooks N. & A.H. Doob (1990). Tax evasion: Searching for a theory of compliant behavior. In Securing compliance: Seven case studies ed. ML Friedland 122-164 Toronto, University of Toronto Press.

Chan, C.W., Troutman, C.S. and O'Bryan, D. (2000). An Expanded Model of Taxpayer Compliance: Empirical Evidence from the United States and Hong Kong, Journal of International Accounting Auditing and Taxation, Vol. 9, 83-103.

Chelliah, R.J. (1971). Trends in Taxation in Developing Countries, International Monetary Fund Staff Papers,, Vol. 18, 254-331.

Chiarini, B., Marzano, E. and Schneider, F. (2008). Tax Rates and Tax Evasion: An Empirical Analysis of the Structural Aspects and Long-Run Characteristics in Italy, IZA Discussion Paper, No. 3447.

Christian, C. W. and Gupta, S. (1992). New Evidence on Secondary Evasion, The Journal of the American Taxation Association, Vol. 15(1) 72 – 93.

Christou G. (2003). Introduction to econometrics, Gutenberg, Athens

Clotfelter, C. (1983). Tax Evasion and Tax Rates: An Analysis of Individual Returns, Review of Economics and Statistics, Vol. 65(3), 363-73

Collins, J.H., Millron, V.C., and Toy, R.D. (1992). Determinants of Tax Compliance: A Contingency Approach. The Journal of the American Taxation Association, Vol.14, 1-29.

Cowell, F. (1990). Cheating the government: The economics of evasion, Cambridge, Mass. and London: MIT Press.

Crane, S. and Nourzad, F. (1985). Time Value of Money and Income Tax Evasion Under Risk Averse Behavior: Theoretical Analysis and Empirical Evidence, Public Finance, Vol. 40(3), 381-94

Crocker, K., and Slemrod., J. (2005). Corporate Tax Evasion with Agency Costs, Journal of Public Economics, Vol. 89(9–10), 1593–1610.

Cuccia A. (1994). The economics of tax compliance : what do we know and where do we go?. Journal of accounting literature**,** Vol. 13.1994, p. 81-116

Dubin, J. A., Michael J. G. and Louis L. W. (1987). Are We a Nation of Tax Cheaters? New Econometric Evidence on Tax Compliance, American Economic Review, Vol. 77(2), 240-245S.

Dubin, J. A., and Wilde, L. L. (1988). An empirical analysis of federal income tax auditing and compliance, National Tax Journal, Vol. 41, 61–74.

Engle και Granger (1987. Co-Integration and Error Correction: Representation, Estimation, and Testing, Econometrica Vol. 55, No. 2 (Mar., 1987), pp. 251-276

Falsetta, D., Schafer, J. K., and Tsakumis, G. T. (2010). Tax Evasion: Audit Probability and the Moderating Role of Goal Conflict. Working Paper.

Feige, E.L. and Cebula, R. (2009). America's Underground Economy: Measuring the Size, Growth and Determinants of Income Tax Evasion in the U.S.

Feinstein, J. S. (1991). An Econometric Analysis of Income Tax Evasion and its Detection, Rand Journal of Economics, Vol. 22(1), 14-35.

Franzoni, L.A. (1999). Tax Evasion and Tax Compliance in B. Bouckaert and G. De Geest (eds), Ecyclopedia of Law and Economics, Cheltenham, UK and Northampton, MA, USA: Edward Elgar

Franzoni, L. A. (2008). Tax Compliance, Encyclopaedia of Law and Economic, Boudewijn Bouckaert and Gerrit De Geest, eds., Edward Elgar.

Frey, B. and Weck-Hanneman, H. (1984). The Hidden economy as an 'unobservable variable, European Economic Review, Vol.26, 33-35.

Ghura, D. (1998). Tax Revenue in sub-Sahara Africa: Effects of economic policies and corruption. International Monetary Fund. Working Paper, 1998/135.

Goerke, L. and M. Runkel (2006). Profit Tax Evasion under Oligopoly with Endogenous Market Structure, National Tax Journal, Vol. 59, 851-857.

Granger C. (1988). Causality, cointegration and control. Journal of Economic Dynamics and Control, 1988, vol. 12, issue 2-3, 551-559

Hanno, M. D. and Violette, G. R. (1996). An Analysis of Moral and Social Compliance on Taxpayer Behaviour, Behavioural Research in Accounting, Vol.8, 163-189.

Hinrichs, H. H. (1966). A General Theory of Tax Structure Change During Economic Development, Cambridge, Mass: Harvard Law School International Tax Program

Hite & Roberts (1992). Progressive Taxation, Fairness, and Compliance, Law and Policy Vol 16

Jackson, B. R., and Milliron, V. C. (1986). Tax compliance research: Findings, problems and prospects, Journal of Accounting Literature, Vol. 5, 125–165.

Jou, J. (1992). Income Tax Evasion: Theory and Empirical Evidence from the U.S. State-Level Data, Tax Year 1976-1989. Ph.D. Dissertation, University of Chicago.

Joulfaian, D. (2009). Bribes and Business Tax Evasion, The European Journal of Comparative Economics, Vol. 6(2), 227-244.

Joulfaian, D. and Rider, M. (1996). Differential Taxation and Tax Evasion by Small Business, National Tax Journal, Vol. 51(4), 676-87.

Kamdar, N. (1995). Information Reporting and Tax Compliance: An Investigation Using Individual TCMP Data, Atlantic Economic Journal, Vol. 23 (4), 278-92.

Kastlunger, B., Kirchler, E, Mittone, L, and Pitters, J. (2009). Sequences of Audit, Tax Compliance, and Taxpaying Strategies, Journal of Economic Psychology, Vol. 30(3), 405-418.
Kastlunger, B., Dressler, S., Kirchler, E., Mittone, L., and Voracek, M. (2010). Sex differences in tax compliance: Differentiating between demographic sex, gender-role orientation, and prenatal masculinization (2D:4D), Journal of economic psychology, Vol.31, 542-552.
Klepper, S. And Nagin, D. (1989). Certainty and Severity of Punishment Revisited. Criminology, Vol.27(4), 721-746.
Long, S. and Swingen, J. (1991). The Conduct of Tax-Evasion Experiments: Validation, Analytical Methods, and Experimental Realism, in: P. Webley, H. Robben, H. Elffers and D. Hessing, Tax Evasion: An Experimental Approach. Cambridge University Press, Cambridge: 128-138
Marrelli, M. (1984). On Indirect Tax Evasion, Journal of Public Economics, Vol. 25, 181-196
Mason, R. and Calvin, L. (1978). A study of admitted income tax evasion, Law and Society Review, Vol. 12, 73–89.
Mason & Lowry (1981). How Do We Affect Taxpayer Behavior?: The Case for Positive Incentives, Assistance or Enforcement, IRS 1990 Research Conference
Milliron,V. and Toy, G. R. (1988). Tax Compliance: An Investigation of Key Features, The Journal of the American Taxation Association, Vol.41, 84 -103.
Mittone, L. (2006). Dynamic behaviour in tax evasion: An experimental approach, The Journal of Socio-Economics, Vol. 35(5), 813-835.
Negreponti-Delivani M. (1990). The Problematic Greek Industry And Some Of Its Solutions
Nur-tegin K. D. (2008). Determinants of Business Tax Compliance. The B.E. Journal of Economic Analysis and Policy, Vol. 8(1), 1-26.
Porcano, M. (1988). Correlates of tax evasion, Journal of Economic Psychology, Vol. 9,47-56
Pommerehne, W. and Weck-Hannemann, H. (1996). Tax Rates, Tax Administration and Income Tax Evasion in Switzerland, Public Choice, Vol. 88, 161-170.
Reinganum, J.F and Wilde, L.L. (1985). Income Tax Compliance in a Principal-Agent Framework, Journal of Public Economics, Vol. 26, 1-18.
Riahi-Belkaoiu, A. (2004). Relationship between tax compliance internationally and selected determinants of tax morale, Journal of International Accounting, Auditing and Taxation, Vol. 13, 135– 143.
Rice, E. M. (1992). The Corporate Tax Gap: Evidence on Tax Compliance by Small Corporations, in Why People Pay Taxes, edited by Joel Slemrod. Ann Arbor: University of Michigan Press, 1992
Richardson, G. (2006). Determinants of tax evasion: A cross-country investigation, Journal of International Accounting, Auditing and Taxation, Vol. 15, 150–169.
Richardson, M. and Sawyer, A. J. (2001). A taxonomy of the tax compliance literature: Further findings, problems and prospects, Australian Tax Forum, Vol. 16, 137–320.
Sandmo, A. (2004). The Theory of Tax Evasion: A retrospective view, Norwegian School of Economics and Business Administration, Discussion Paper 31/04.
Schram, A. and Gërxhani, K. (2006). Tax Evasion and Income Source: A Comparative Experimental Study. Journal of Economic Psychology, Vol.27(3), 402-422.
Sillamaa, M. and Veall, M.R. (2000). The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1988 Tax Flattening in Canada, Quantitative Studies in Economics and Population Research Reports, No 354, McMaster University.
Slemrod, J. (2007). Cheating Ourselves: The Economics of Tax Evasion, Journal of Economic Perspectives, Vol. 21(1), 25-48.
Song, Y. and Yarbrough, T. E. (1978). Tax Ethics and. Taxpayer Attitudes: A Survey, Public Administration Review, Vol. 38(5), 442–52
Sookram, S. and Watson, P.K. (2005). Tax evasion, Growth and the Hidden Economy in Trinidad and Tobago, Working Paper, Sir Arthur Lewis Institute of Social and Economic Studies.
Spicer, M.W. (1974). A Behavioral Model of Income Tax Evasion. Dissertation, Ohio State University

Spicer, M. W. and Lundstedt, S. B. (1976). Understanding Tax Evasion, Public Finance, Vol.21(2), 295-305.

Strumpel, B. (1969). The contribution of survey research to public finance, in: A. T. Peacock (Ed.), Quantitative Analysis in Public Finance, 13–22. New York: Praeger.

Tanzi, V. (1987), Quantitative Characteristics of the Tax Systems of Developing Countries, in The Theory of Taxation for Developing Countries, ed. by David Newbery and Nicholas Stern (New York: Oxford University Press), pp. 205-41.

Tanzi, V. and Zee, H.H. (2000). Tax Policy for Emerging Markets: Developing Countries, National Tax Journal, Vol.53 (2), 299-322.

Tedds (2006). Tax non-compliance and corporate governance: a comparative study. Working paper (2006)

Tittle, C. (1980). Sanctions and social deviance: The question of deterrence, New York: Praeger.

Torgler, B. (2003). Tax Morale in Transition Countries, Post-Communist Economies, Vol. 15, 357-381.

Torgler, B. (2006). The Importance of Faith: Tax Moral and Religiosity, Journal of Economic Behavior and Organization. 61(1) 81-109

Torgler, B. (2007a). Tax Compliance and Tax Morale: A Theoretical and Empirical Analysis, Cheltenham, UK: Edward Elgar.

Torgler, B. (2007b). Tax Morale in Central and Eastern European Countries, in: Nicolas Hayoz and Simon Hug (eds.), Tax Evasion, Trust and State Capacities. How Good Is Tax Morale in Central and Eastern Europe? Bern: Peter Lang: 155-186.

Torgler, B. (2011). Tax Morale and Compliance, The World Bank, Europe and Central Asia Region, Working Paper Series 5922.

Torgler, B. and Murphy, K. (2004). Tax Morale in Australia: What Factors Shape It and Has It Changed Over Time?, Journal of Australian Taxation. Vol. 7, 298-335.

Vogel, J. (1974). Taxation and Public Opinion in Sweden: An interpretation of recent survey data, National Tax Journal, Vol. 27(4), 499-513.

Wallace, S. (2002). Imputed and Presumptive Taxes: International Experiences and Lessons for Russia, Andrew Young School of Policy Studies, Georgia State University, International Studies Program, Working Paper 02-03,

Wallschutzky, I. (1984). Possible Causes of Tax Evasion, Journal of Economic Psychology, Vol. 5(4), 371-384.

Witte, A. D., and Woodbury, D. F. (1985). The effect of tax laws and tax administration on tax compliance: The case of the U.S. individual income tax, National Tax Journal, Vol. 38, 1–13.