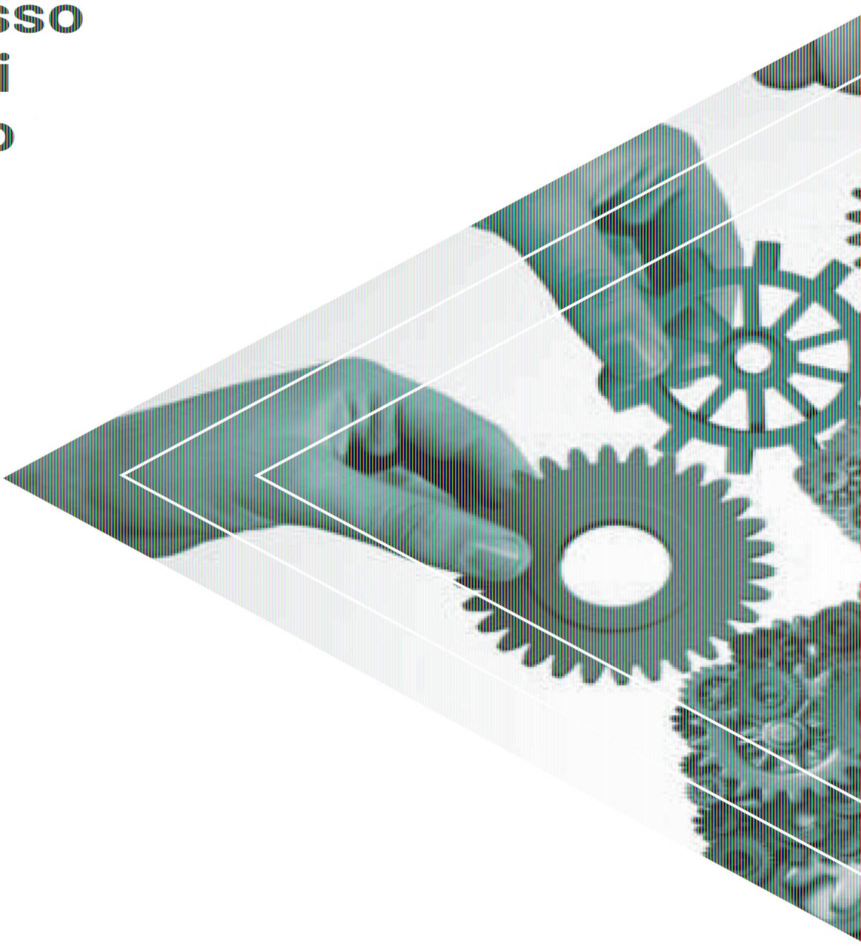
CNR – DSU Working Paper
CNR – Dipartimento Scienze Umane e Sociali I

Tax Evasion, Technological Progress ar Expdenditure: Theory and Empirical Ev

Maria Rosaria | **Alfano**
Salvatore | **Capasso**
Salvatore | **Ciucci**
Nicola | **Spagnolo**

CNR
DSU



ISSN 3035-224X
3 | 2025

workingpaper.dsu@cnr.it

www.dsu.cnr.it/working-paper/

Tax Evasion, Technological Progress and R&D Expdenditure: Theory and Empirical Evidence

Maria Rosaria Alfano
Department of Economics, University of Campania "Luigi Vanvitelli", ITALY
mariarosaria.alfano@unicampania.it

Salvatore Capasso
Department of Business and Economics, University of Naples Parthenope, Italy
Department of Human and Social Studies, CNR, ITALY
salvatore.capasso@uniparthenope.it

Salvatore Ciucci
Department of Economics, University of Campania "Luigi Vanvitelli", ITALY
salvatore.ciucci@unicampania.it

Nicola Spagnolo
Department of Economics, University of Campania "Luigi Vanvitelli", ITALY
nicola.spagnolo@unicampania.it

ABSTRACT

This paper introduces a novel theory of tax evasion, suggesting that the optimal decision to evade taxes is influenced by agents' portfolio technologies. The theory argues that evading taxes involves the use of specific production technologies ("hidden technologies") with returns that differ from those of technologies that require tax compliance ("visible technologies"). As technological progress affects the gap and returns between these types of production technologies, the optimal level of tax evasion shifts. Furthermore, considering social customs, the paper concludes that the overall level of tax evasion is determined by the prevalent nature of technologies in the economy. This theory is empirically tested using a panel cointegration framework, analysing data from 67 countries between 2001 and 2020 to ensure broad relevance. The empirical results confirm the theoretical model, even when accounting for variations in R&D expenditure levels.

Questo articolo introduce un nuovo modello teorico di evasione fiscale, suggerendo che l'evasione fiscale è influenzata dal progresso tecnologico. Tale teoria stabilisce che gli individui possono scegliere di evadere le tasse attraverso la scelta della tecnologia da impiegare nel processo produttivo, le tecnologie produttive disponibili differiscono sia nel rendimento marginale, che nell'opportunità di nascondere i redditi generati all'autorità fiscale. Utilizzando dati panel, raccolti per 67 paesi, fra il 2001 e il 2020, viene condotta un'analisi empirica che conferma i risultati teorici.

Keywords: Tax evasion, labour effort, Technological progress, R&D expenditure, Shadow economy, Panel cointegration.
DOI: 10.36134/CNRDSUWP-2025-3

Sottomesso il 22/10/2025 e pubblicato il 19/01/2026

1. Introduction

Although there is no consensus on the role tax evasion plays in the economy, it is widely acknowledged as a negative side effect of taxation that creates adverse effects on economic systems (Slemrod and Yitzahi, 2002). Consequently, it is important to limit tools that alter taxpayers' behavior. Issues such as tax evasion and the shadow economy's size and impact are key economic policy objectives worldwide. While the terms "tax evasion" and "shadow economy" are often used interchangeably, the shadow economy encompasses all economic activities not registered by authorities and avoiding governmental regulation or remaining hidden from the state for tax purposes (Schneider 2012). The shadow economy involves not only tax evasion but also non-compliance with labour market regulations and avoidance of other administrative requirements, meaning that factors affecting tax evasion also impact the shadow economy.

The literature on tax evasion and the shadow economy is extensive, considering the numerous factors that theoretically influence individual tax compliance decisions. Dell'Anno (2021a) categorizes the causes of tax evasion into: the taxation system (including tax burden, marginal tax rate, and complexity of tax regulations); labour force participation and composition (such as self-employment rate, unemployment rate, and percentage of illegal immigrants); tax enforcement (cost and probability of audits, penalty schemes); regulatory system (rule of law, labour market regulation, restrictions on immigrants); institutional quality (corruption, public policy fairness, media and economic freedom, public goods and services quality); and tax morale (the rejection of tax fraud by taxpayers). Empirical studies find correlations between the shadow economy and various economic factors, such as financial development (Bittencourt *et al.*, 2014; Berdiev and Saunoris, 2016), inflation (Bittencourt *et al.*, 2014), business cycles (Elgin *et al.*, 2021), economic growth (Alfonso *et al.*, 2020), inequality (Dell'Anno, 2021b), poverty (Bonnet and Venkatesh, 2016), economic development (Goel and Nelson, 2016), and education (Uyar *et al.*, 2022).

In our view, technological progress, a key driver of economic growth, significantly impacts tax evasion, influencing the variables identified in the literature as main causes of the shadow economy's size. R&D expenditure serves as a proxy for technological progress, yet few studies have explored the impact of R&D expenditure on the shadow economy.

Our analysis extends the existing literature in several ways. First, we propose a theoretical model examining the relationship between technological progress and individual tax compliance choices. We suggest that technological progress affects labour/leisure choices by altering returns on different production technologies, with advanced technologies being harder to conceal from authorities. We then conduct a panel cointegration analysis to investigate the long-term equilibrium between R&D expenditure and the shadow economy size, confirming our theoretical findings. To our knowledge, this is the first attempt to explain the link between technological progress and tax evasion. Unlike most studies that apply OLS or GLS to non-stationary panel variables, generating spurious estimates, we use FMOLS and PMG procedures to produce consistent and robust results, addressing potential endogeneity, serial correlation, heteroscedasticity, and cross-sectional dependence.

The rest of the paper is structured as follows. Section 2 reviews the related literature. In Section 3, we introduce the theoretical model and interpret the results. Section 4 outlines the empirical methodology, discusses the baseline findings, and presents various robustness checks. Section 5 offers some concluding remarks.

2. Related literature

The literature on the economics of tax evasion is extensive, covering both theoretical and empirical aspects, and has been reviewed multiple times (Dell'Anno, 2022).

The economic analysis of tax compliance began with the seminal study by Allingham and Sandmo (1972). This study combined the concepts of the economics of crime with the principles of optimal portfolio theory, examining a risk-averse taxpayer with exogenous taxable income making portfolio choices under uncertainty to maximize expected utility. The risk in this context arises from the possibility that tax authorities may conduct random audits. If evasion is detected, a fixed penalty rate greater than the tax rate is applied to the hidden income.

In 1974, Yitzhaki modified the model by assuming that if evasion is detected, a fine proportional to the tax rate is applied. This modification, known as the "standard model," provides clear comparative static predictions. According to the Allingham-Sandmo model, and supported by empirical evidence, the optimal level of tax evasion decreases when either the audit probability or the penalty rate increases. Additionally, as absolute risk aversion decreases, tax evasion becomes positively correlated with income. However, the standard model has faced criticism due to its foundational assumptions, limited real-world applicability, and the prediction that optimal evasion decreases as tax rates increase. Yitzhaki (1987) challenged this prediction by suggesting that the probability of detection increases with evaded income, especially for risk-neutral agents.

Some authors have relaxed the assumption of risk-neutrality in their studies (Cremer and Gahvari, 1994). The incorrect predictions of the standard model are often attributed to the reliance on expected utility theory, as proposed by Von Neumann and Morgenstern (1947). To address the limitations of expected utility theory, several authors have incorporated non-expected utility theories into the analysis of tax compliance decisions. For example, Dhami and Al-Nowaihi (2007) have employed prospect theory, which introduces a reference point against which gains and losses are evaluated, resulting in a different payoff function structure for gains and losses.

The use of weighted probabilities in prospect theory offers greater flexibility, leading to more accurate predictions. Arcand and Graziosi (2005) have applied Rank-Dependent Expected Utility, generating values of the weighted probability function recursively. Snow and Warren (2005) introduced the concept of ambiguity to account for situations where an individual's information is partial, increasing uncertainty in decision-making. In practice, different probability distributions can be observed, but selecting the appropriate distribution can be challenging.

The standard model has been extended in various ways to enhance its accuracy. For instance, some researchers have made the taxpayer's income endogenous by introducing labour supply decisions into the model (Cremer and Gahvari, 1995; Schroyen, 1997). Other studies have highlighted the significance of social norms and social interactions in individual tax evasion decisions. Cowell and Gordon (1988) introduced the concept of fairness, linking the provision of public goods to tax compliance decisions. They assumed that taxpayers consider the services they receive in return for their tax payments, with changes in the tax rate significantly impacting the supply of public goods and influencing tax evasion decisions.

Gordon (1989) argued that tax evasion carries a psychic cost in the form of a loss of social prestige if detected. Dell'Anno (2009) extended Gordon's model to analyse how tax morale – defined as the intrinsic motivation to pay taxes - affects tax evasion decisions.

Only few empirical studies have investigated the relationship between technological R&D spending and the shadow economy. Markellos *et al.* (2016) noted a negative statistical relationship between the share of R&D expenditures over GDP and the scale of the shadow economy. Nevzorova *et al.* (2018) conducted a correlation analysis using a sample of 402 observations from 2010 to 2015, finding that countries with a high level of shadow economy tend to have extremely low R&D expenditure shares, indicating a negative relationship. Amanova (2022) empirically analysed the effect of R&D expenditure and digitalization on Uzbekistan's shadow economy, finding an inverse correlation between shadow economy size and R&D expenditure.

3. The Model

To endogenize individuals' gross income, several authors, have attempted to extend the standard Allingham and Sandmo model to incorporate the labor supply decision. Cowell (1985), in particular, was among the first to explore the idea that taxpayers could allocate their labor supply between legal and illegal activities. Building on this literature, we consider an economy populated by a given number of representative rational individuals who have access to two technologies: a primary production technology, which allows each of them to earn an income $w_p = \theta_p e_p$ and a subsidiary production technology which generates an income $w_s = \theta_s e_s$. Where θ_p and θ_s represent the individual's earning ability in the primary and subsidiary production technologies, respectively, and e_p and e_s denote the labour efforts exerted by the individual in each respective technology.

The tax authority is tasked with tax collection and conducting audits to ensure compliance. It can directly observe individuals' income only from the primary production technology. Individuals are subject to an income tax rate *t*. However, income generated from the subsidiary technology is not directly observable by the tax authority and can only be accessed through audits. In this context, individuals can evade taxes by reallocating their labour efforts to the subsidiary production technology. To combat tax evasion, the tax authority employs a random auditing process with an auditing probability $p \in [0,1]$. If an individual is audited and found to have evaded taxes, they are liable to pay a penalty *F*.[1]

The economic literature has extensively demonstrated that technological advancements significantly enhance productivity, leading to a more skilled workforce with higher earning potential, thereby fostering long-term economic growth. Moreover, studies have shown that new technologies can coexist alongside older ones for extended periods (Battisti and Stoneman, 2005).

It is plausible to expect that technological change amplifies the disparity in returns between primary and subsidiary production technologies. The rationale is that as technological progress unfolds, the adoption of more advanced technologies is likely to become widespread, making it increasingly difficult to conceal income generated from these labour technologies.

Now, let's assume individuals have quasi-linear preferences, represented as follows:

$$u(c; e_p; e_s): c - \frac{(e_p)^2}{2} - \frac{(e_s)^2}{2} \qquad (1)$$

where c denotes consumption and e denotes labour effort. Note that $u(c; e_p; e_s)$ is twice continuously differentiable, concave[2], increasing in consumption and decreasing in labour efforts[3]. The individual's consumption c, which can take any non-negative value, is given by:

$$c = w_p(1-t) + w_s(1-p) + pw_s(1-F) \qquad (2)$$

Therefore, the agent consumes all net income and acts rationally. The model under consideration is set in a single period where the individual's sole decision variable is the level of labour effort allocated to each production technology. This decision aims to maximize the utility function while respecting the budget constraint (2). We can formalize the individual's maximization problem as follows:

$$\max_{e_p; e_s} u(c; e_p; e_s): c - \frac{(e_p)^2}{2} - \frac{(e_s)^2}{2} \qquad (3)$$

subject to

$$c = w_p(1-t) + w_s(1-p) + pw_s(1-F) \qquad (4)$$

$$e_p \geq 0; e_s \geq 0; e_p + e_s = TL \qquad (5)$$

Eq. (4) ensures that labour efforts cannot be negative and their sum must equal to the total working time TL.

The optimal labour efforts that solve problem (3) are as follow (refer to the proof in Appendix A):

$$e_p^* = 0;\ e_s^* = TL \qquad (6)$$

if

$$\theta_p \leq \frac{\theta_s(1-pF)-TL}{1-t}$$

[1]Please note that tax evasion is a feasible option, given that $pF \in [0,1]$, where both p and F are less than one.

[2]This property is detailed in Appendix A.

[3] The simpler case of separable utility functions is commonly employed in the literature on optimal labor income taxation and contract theory to describe the preferences of agents.

$$e_p^* = \frac{TL + \theta_p(1-t) - \theta_s(1-pF)}{2}; \; e_s^* = \frac{TL - \theta_p(1-t) + \theta_s(1-pF)}{2} \quad (7)$$

if

$$\frac{\theta_s(1-pF)-TL}{1-t} < \theta_p < \frac{\theta_s(1-pF)+TL}{1-t}$$

$$e_p^* = TL; \; e_s^* = 0 \quad (8)$$

if

$$\theta_p \geq \frac{\theta_s(1-pF)+TL}{1-t}$$

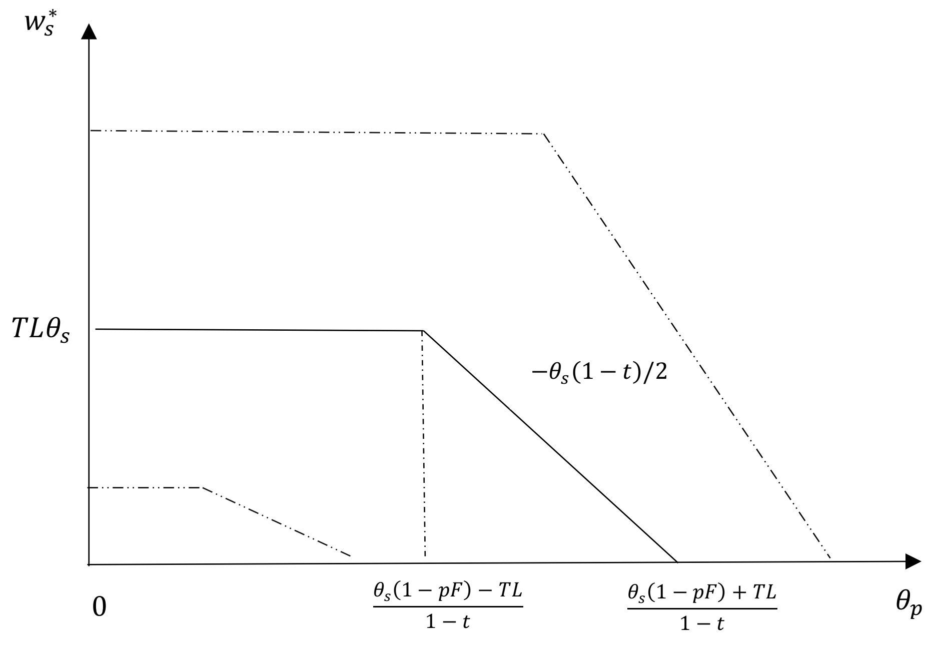
Therefore, the optimal amount of concealed income is:

$$w_s^* = TL\theta_s \text{ if } \theta_p \leq \frac{\theta_s(1-pF)-TL}{1-t} \quad (9)$$

$$w_s^* = \frac{TL\theta_s - \theta_p\theta_s(1-t) + \theta_s^2(1-pF)}{2} \text{ if } \frac{\theta_s(1-pF)-TL}{1-t} < \theta_p < \frac{\theta_s(1-pF)+TL}{1-t} \quad (10)$$

$$w_s^* = 0 \text{ if } \theta_p \geq \frac{\theta_s(1-pF)+TL}{1-t} \quad (11)$$

Fig. 1 – The tax evasion – Primary technology productivity curve. The optimal amount of evaded income w_s^* in relation to the productivity of the primary technology θ_p. The slope of the curve is $-\theta_s(1-t)/2$



Source: Authors' own elaboration

Figure 1 graphically illustrates the above statement. When the productivity level of the subsidiary technology, θ_s, is given and the productivity of the primary technology, θ_p, is low, the individual opts for maximum effort in the subsidiary production technology, leading to maximum tax evasion. Conversely, at high productivity levels of θ_p, full tax compliance is observed. Within a specific range of θ_p, as detailed in Appendix B, the optimal concealed income, w_s^*, shows a negative correlation with the productivity of the primary technology. Hence, an increase (or decrease) in θ_p results in a decrease (or increase) in the amount of tax evasion. This relationship stems from $\frac{\partial w_s^*}{\partial \theta_p} = -\frac{\theta_s(1-t)}{2} < 0$.

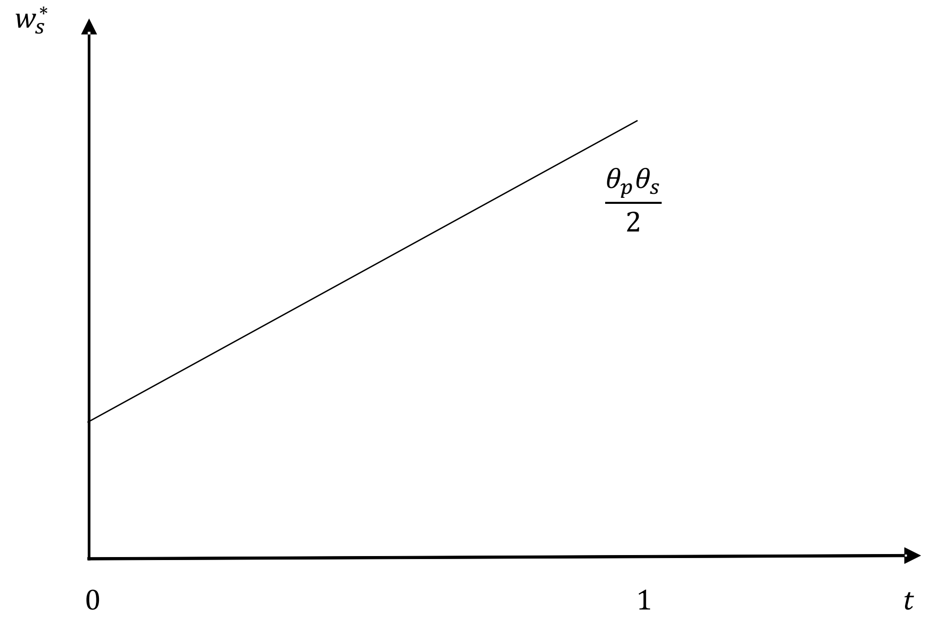
Examining equations (8) and (9), it is evident that the optimal amount of tax evasion is positively affected by the productivity of the subsidiary technology, θ_s. This relationship is evident from the fact that $\frac{\partial w_s^*}{\partial \theta_s} > 0$ (refer to Appendix B for the proof). An increase (or decrease) in the productivity of the subsidiary technology, θ_s, causes the tax evasion-primary technology productivity curve in Figure 1 to shift upward (or downward), resulting in a steeper (or flatter) slope.

Furthermore, the extent of tax evasion is inversely influenced by the difference in productivity between the primary and subsidiary technologies$\left(\theta_p - \theta_s\right)$. Additionally, the negative impact of primary technology productivity on tax evasion depends on the levels of subsidiary technology productivity and the tax rate. A higher subsidiary technology productivity (or a lower tax rate) strengthens the negative effect of primary technology productivity on tax evasion, and conversely[4].

From eq. (9) it can be noted that the optimal amount of concealed income is positively related with the tax rate, since $\frac{\partial w_s^*}{\partial t} = \frac{\theta_p \theta_s}{2} > 0$ (refer to Appendix B for the proof).

Figure 2 illustrates the relationship between the optimal amount of tax evasion and the tax rate. The slope of this curve is positive and is influenced by the productivity levels of both the primary and subsidiary technologies. Therefore, an increase (or decrease) in θ_p and/or θ_s results in the Tax Evasion - Tax Rate curve becoming steeper (or flatter). Consequently, the productivity levels of both technologies, regardless of their distance from each other, impact the positive correlation between the tax rate and individuals' tax compliance decisions. When the distance between production technologies is minimal and both technologies have low productivity, the effect of a change in the tax rate on evasion is minimal because both the potential taxes owed and the opportunity cost of evasion are low. Conversely, if the distance between production technologies remains minimal but the productivity of both technologies is high, the impact of the tax rate on tax evasion is more pronounced.

Fig. 2 – The Tax evasion – Tax rate curve. The relationship between the optimal amount of tax evasion w_s^* and the tax rate t.



Source: Authors' own elaboration.

[4] The mechanism is straightforward to explain: higher values of θ_s lead to increased tax evasion. Therefore, an increase in θ_p has a more pronounced effect on evasion when θ_s is high because evasion levels are already elevated. Conversely, when the tax rate is high, an increase in θ_p has a smaller impact on tax evasion because the potential tax liability is greater if individuals choose to evade fewer taxes.

To summarize:

- The larger the gap between the productivity of the primary and subsidiary technologies $(\theta_p - \theta_s)$, the lower the level of tax evasion in the economy, and conversely;
- The negative influence of primary technology productivity on tax evasion is positively correlated with the magnitude of subsidiary technology productivity;
- The responsiveness of individuals in terms of tax compliance to changes in the tax rate depends on the productivity levels of both technologies.

3.1 The effect of Social Customs

Many behavioural economics authors emphasize that tax compliance decisions are influenced by social interactions rather than being made in isolation. Individuals are embedded in various social contexts that shape their choices. Gordon (1989) was an early advocate for integrating insights from social customs literature into the standard model of tax evasion. This integration aimed to explain the additional costs associated with tax evasion, known as 'psychic costs'. These costs include the loss of social standing or reputation when caught evading taxes. They escalate with the level of evasion and vary across individuals, leading to a segmentation of the population into evaders and non-evaders.

In this section, we incorporate social customs and conformity into our tax evasion model using the framework developed by Myles and Naylor (1996). According to this framework, individuals derive social utility when taxes are paid honestly but incur social disutility when engaging in evasion.

The utility of an individual who chooses not to evade taxes is:

$$u_{ne}(c; e_{p,ne}) = c - \frac{(e_{p,ne})^2}{2} + aR(1-\mu) + b \tag{11}$$

with $a \geq 0$ and $b \geq 0$. We assume that $R' > 0$. You can think of $aR(1-\mu)$ as the utility of conforming to the group of full tax compliance individuals. In fact, $(1-\mu)$ is the proportion of population not evading tax, and b the utility obtained from following the social custom; while $e_{p,ne}$ is the labour effort exerted in the primary production technology, when the individual chooses not to use also the subsidiary production technology, and therefore it is equal to TL.

We can rewrite eq. (11) as:

$$u_{ne}^*(c, e_{p,ne}^*) = \theta_p(1-t)TL - \frac{TL^2}{2} + aR(1-\mu) + b \tag{12}$$

Since the return from the social custom and from conformity is derived only from the honest behaviour, the individual's utility level of who chooses evasion[5], is given by:

$$u_e^*(c; e_p^*; e_s^*) = \frac{1}{2}TL\left[\theta_p(1-t) + \theta_s(1-pF) - \frac{1}{2}TL\right] + \frac{1}{4}\left[\theta_p(1-t) - \theta_s(1-pF)\right]^2 \tag{13}$$

(See Appendix C for the proof).

Tax evasion takes place if $u_{ne}^* < u_e^*$, while full tax compliance occurs if $u_{ne}^* \geq u_e^*$[6]

As in Myles and Naylor (1996), it is possible to derive the critical proportion of evaders μ^*:

[5] We consider the case in which individual chooses to exert effort in both production technologies.
[6] The weak inequality in this expression arises from the assumption that the individual prefers to adhere to social customs unless there is a utility loss associated with doing so.

$$\mu^* = 1 - \frac{TL^2 - 2TL\left[\theta_p(1-t) - \theta_s(1-pF)\right] + \left[\theta_p(1-t) - \theta_s(1-pF)\right]^2 - 4b}{4aR} \quad (14)$$

(See Appendix C for the proof).

If $\mu > \mu^*$then an individual of parameters $(a; b)$ chooses to evade taxes, instead if $\mu \leq \mu^*$then an individual of parameters $(a; b)$ chooses full tax compliance.

Since $\frac{\partial\mu^*}{\partial t} < 0$; $\frac{\partial\mu^*}{\partial\theta_p} > 0$ and $\frac{\partial\mu^*}{\partial\theta_s} < 0$ (see Appendix C for the proof), an increase (decrease) in the tax rate leads to a decrease (increase) in the critical proportion μ^*[7], but what we want to emphasize is that, this critical value changes as the distance between the two production technologies $(\theta_p - \theta_s)$ changes, in particular the greater (lower) the distance the greater (lower) the critical proportion of population evading tax, such that the individual of parameters $(a; b)$ begins to evade.

Therefore, since a decrease in the distance between the two production technologies could induce more taxpayers to evade the relationship between this distance and tax evasion is strengthened.

A branch of behavioral economy suggests that (Myles *et al.,* 1996; Bethencourt *et al.*, 2019) different tax compliance choices are just due to different parameters of social customs and social conformity, if so, it would be difficult to explain that different industries of the same economy have a different level of tax evasion, because this would mean that individuals with higher social customs and social conformity parameters, would all be employed in a specific industry, and it is very unrealistic. Hence, the intuition is that, holding the negative correlation between social customs parameters and tax evasion (as well as the positive relationship between tax rate and tax evasion) there exists another variable which strongly affects individual tax compliance decision, and it is represented by the distance between the accessible production technologies.

4. Empirical analysis

To test our theoretical framework empirically, we conduct an investigation using panel data. The econometric model aims to analyse the impact of R&D expenditure, used as a proxy for technological progress, on the shadow economy, which serves as a proxy for tax evasion. The model is specified as follows:

$$lnSHA_{it} = \beta_0 + \beta_1 lnRD_{it} + \beta_2 lnGDP_{it} + \beta_3 lnEDU_{it} + \varepsilon_{it} \quad (15)$$

where $i = 1; 2; \ldots.; N$ and $t = 1; 2; \ldots.; T$ identify the country and time dimensions respectively. We collect annual data for 67 countries (refer to Appendix D for the full list). The dependent variable, Shadow Economy (SHA), is quantitatively estimated as a percentage of GDP based on Elgin *et al.* (2021), using the Multiple Indicators Multiple Causes (MIMIC) approach. The independent variable, R&D expenditure (RD), is expressed as a percentage of GDP. Following the existing literature, we also include control variables to mitigate any potential omitted variable bias: GDP, representing per capita gross domestic product (Schneider and Williams, 2013; Hassan and Schneider, 2016;); and EDU, indicating the global education index (Uyar *et al.,* 2022).

The Shadow Economy (SHA) data is a time series available until 2020, hence all observations span from 2001 to 2020 (20 observations). Table 1 outlines the variables, their definitions, and measurements, sources, and data descriptions. Table 2 provides summary statistics for the entire sample, while Table 3 illustrates the unconditional correlation matrix, highlighting the negative correlation between R&D expenditure and the size of the shadow economy.

[7] Note that in Myles and Naylor (1996), the sign of the derivative $\frac{\partial\mu^*}{\partial t}$ can be both positive and negative. However, empirical experiments (Kleven *et al.,* 2011) suggest that as the tax rate increases, the critical proportion of the population evading taxes tends to increase.

Tab. 1 – Data Description and Sources

Variable	Notation	Description	Source
Shadow Economy	*SHA*	Size of the shadow economy as a percentage of GDP.	Elgin, C., M. A. Kose, F. Ohnsorge, and S. Yu. 2021. "Understanding Informality." CERP Discussion Paper 16497, Centre for Economic Policy Research, London. Informal Economy Database (worldbank.org).
R&D expenditure	*RD*	Gross domestic expenditures on research and development (R&D), expressed as a percentage of GDP. They include both capital and current expenditures in the four main sectors: Business enterprise, Government, Higher education and Private non-profit. R&D covers basic research, applied research, and experimental development.	World Development Indicators; World Bank
GDP per capita	*GDP*	GDP per capita based on purchasing power parity (PPP). PPP GDP is gross domestic product converted to international dollars using purchasing power parity rates. An international dollar has the same purchasing power over GDP as the U.S. dollar has in the United States. Data are in constant 2017 international dollars.	World Development Indicators; World Bank
Education Index	*EDU*	Education index is quantified based on two variables: mean years of schooling ("MYS") and expected years of schooling ("EYS"). MYS represents the years of education for a person aged 25 and older in their lifetime and EYS represents the years of schooling a child is expected to attend. The index is published by Human Development Report and each country is ranked from approximately 0 to	UNESCO Institute for Statistics and other sources. https://globaldatalab.org/shdi/metadata/edindex/

Note: The sample consists of 67 countries over the period 2001-2020.

Tab. 2 – Summary Statistics

Variable	Obs.	Mean	Median	Std. Dev.	Min	Max
$lnSHA$	1,340	3.214	3.268	0.491	2.088	4.219
$lnRD$	1,340	-0.329	-0.153	1.146	-3.767	1.741
$lnGDP$	1,340	0.813	0.931	0.791	-1.960	2.490
$lnEDU$	1,340	4.385	4.403	0.124	3.903	4.566

Note: Summary statistics for all 67 countries for the period 2001 to 2020.
Source: Authors' calculation.

Tab. 3 – Correlation Matrix

Variables	$lnSHA$	$lnRD$	$lnGDP$	$lnEDU$
$lnSHA$	1			
$lnRD$	-0.663	1		
$lnGDP$	-0.666	0.666	1	
$lnEDU$	-0.609	0.7153	0.9223	1

Note: Correlation matrix for all 67 countries for the period 2001 to 2020.
Source: Authors' calculation.

Given the variables, we aim to explore their short- and long – run relationship using panel cointegration techniques. The coefficients β in Equation (15) capture the long – run effects between the variables, the error term is denoted by ε.

As pointed out by Shahbaz (2010), the log linear specification provides efficient results compared to a linear specification. Therefore, all variables are transformed into natural logarithms to ensure consistency, accuracy, and robustness in the estimators. [8]

The cointegrating regression considers only the long – run property of a given model, and does not deal with the short – run dynamics explicitly. Clearly, a good panel modelling should describe both short – run dynamics and the long – run equilibrium simultaneously. For this purpose, a panel vector error correction model (VECM) was developed. The specification of the Panel VECM model can be written as follows:

$$\Delta lnSHA_{it} = \alpha_0 + \alpha_1 \Delta lnRD_{it} + \alpha_2 \Delta lnGDP_{it} + \alpha_3 \Delta lnEDU_{it} + \alpha_4 ECT_{it-1} + \upsilon_{it} \qquad (16)$$

Where Δ is the first difference operator, $ECT_{it-1} = lnSHA_{it} - \beta_0 - \beta_1 lnRD_{it-1} - \beta_2 lnGDP_{it-1} - \beta_3 lnEDU_{it-1}$ is the lagged error correction term (ECT) derived from the long – run cointegrating relationship of equation (15). Thus, ECT_{it-1} is deviation from long – run equilibrium at time $t-1$, while, α_1, α_2 and α_3 are the short – run parameters, the adjustment parameter is denoted by α_4.

The long-run relationship described by Equation (15) holds if the time series for each of the four variables are non-stationary, integrated of the same order, and the variables form a cointegrated system. By definition, two or more non-stationary variables are cointegrated if there exists a linear combination of these variables that is stationary. Therefore, cointegration in the traditional sense indicates that the long-run relationship between the variables is linear (in our case, log-linear).

An important consideration in panel data analysis is cross-sectional dependency (CSD), where the behaviour of one country may influence that of another. Traditional panel data estimation methods often assume cross-sectional independence, but the presence of CSD can lead to unreliable results. Using the Pesaran (2004) test for CSD, we reject the null hypothesis of cross-sectional independence at the 1% significance level. [9]

Prior to conducting cointegration tests, it is essential that all variables possess the same time series properties. Specifically, they should exhibit a unit root in levels and be integrated of the same order, denoted as I(d). The choice of unit root test in panel data analysis can be influenced by cross-sectional dependency (CSD). According to Hoechle (2007), first-generation panel unit root tests may yield biased results in the presence of CSD. Therefore, in our study, we employ second-generation panel unit root tests. Specifically, the CIPS test proposed by Pesaran (2007). In Table 4, the results of the unit root panel test using CIPS are presented. Based on the statistics, all variables are non-stationary at the level. The CIPS test rejects the null hypothesis of non-stationarity at the 1% significance level when the variables are I(1).

Tab. 4 – Panel Unit Root Test

	CIPS	
	Level	**First Difference**
$lnSHA$	-1.437	-3.061***
$lnRD$	-1.603	-3.694***
$lnGDP$	-1.633	-2.593***
$lnEDU$	-2.036	-3.612***

Note: critical values for CIPS test (Pesaran, 2007) are -2.36 for 1% significance, -2.2 for 5% significance and -2.11 for 10% significance. In the CIPS test, the null hypothesis is that the panel is homogeneously non-stationary. ***, ** and * indicate the significance levels of 1%, 5% and 10%, respectively.
Source: Authors' calculation.

To explore the possibility of long-run convergence among our data series, we employ a panel cointegration test.

To confirm the cointegration relationship between the variables, we conduct two types of tests: the Pedroni (1999, 2004) and Westerlund (2005) tests. According to the results presented in Table 5, all tests reject the null hypothesis of no cointegration at the 1% significance level.

[8] None of the variables take on negative values.
[9] This result is not presented here to save space. Test results are available from the authors upon request.

Tab. 5. – Pedroni and Westerlund Cointegration Tests

	Pedroni cointegration test		Westerlund cointegration test	
	t Statistics	*p* value	Variance ratio	*p* value
Modified Phillips – Perron	2.69	0.003	-4.124	0.000
Phillips – Perron	-5.229	0.000		
Augmented Dickey – Fuller	-3.356	0.000		

Note: The null hypothesis for both the Pedroni and Westerlund tests is: No cointegration exists across panels. The alternative hypothesis in the Pedroni test is: All panels are cointegrated. The alternative hypothesis in the Westerlund test is: Some panels are cointegrated.
Source: Authors' calculation.

Since all variables in the model are cointegrated and exhibit a long-run association, we proceed to estimate the coefficients.

4.1 Empirical results

Once we confirm that the variables are cointegrated, we proceed to establish the long – run equilibrium relationship using the FMOLS (Fully Modified Ordinary Least Squares) estimation methods suitable for heterogeneous panels. The FMOLS, initially developed by Phillips and Hansen (1990) and Philips and Moon (1999), and introduced by Pedroni (2000), employs a non-parametric approach, such as Newey-West, to tackle issues of serial correlation and endogeneity. Using FMOLS ensures that we mitigate endogeneity and serial correlation problems in our model.

Although FMOLS estimation method examines only long run parameters, the PMG (the Pooled Mean Group) estimation method introduced by Pesaran *et al.* (1999) calculates both long and short run parameters, including the adjustment for long – run equilibrium (speed of adjustment) and error variance to be heterogeneous. PMG constraints the long – run coefficients to be identical in an error correction framework, but allows the short run coefficients and error variances to differ across groups.

Table 6 presents the results of the FMOLS and PMG estimations.

Tab. 6 – FMOLS and PMG Estimation Results

	Long – run	
	FMOLS	**PMG**
$lnRD$	-0.203***	-0.022***
	(0.046)	(0.006)
$lnGDP$	-0.384***	-0.165***
	(0.123)	(0.022)
$lnEDU$	1.193	-0.574***
	(0.835)	(0.105)
Adjusted R – squared	0.525	
Countries	67	67
Observations	1,339	1,237
	Short – run	
	PMG	
ECT	-0.123***	
	(0.017)	
$\Delta lnRD$	0.004	
	(0.005)	
$\Delta lnGDP$	-0.474***	
	(0.033)	
$\Delta lnEDU$	-0.073	
	(0.087)	
Countries	67	
Observations	1,273	

Note: ***, ** and * indicate the significance levels of 1%, 5% and 10%, respectively. The standard errors are reported in brackets.
Source: Authors' calculation.

Since the variables are expressed in natural logarithms, the coefficients can be interpreted as elasticities. The results from both long – run estimations indicate that R&D expenditure *(lnRD)* has a negative and statistically significant effect on the shadow economy *(lnSHA)*. According to the FMOLS model, in the long – run, a 1% increase in R&D expenditure is associated to a 0.203% decrease in the size of the shadow economy, while the long – run coefficient of PMG estimation is equal to -0.022. In the short – run the impact of R&D expenditure is not statistically significant. The error correction term is correctly negatively signed and highly significant, its magnitude is -0.123 suggesting a speed adjustment process, which means that, if shadow economy is 1% out of equilibrium, a 12.3 percent adjustment towards equilibrium will take place within the first year.

Regarding the control variables, the negative impact of GDP *(lnGDP)* is confirmed, both in the long- and short – run, which is consistent with findings in the existing literature. However, the impact of education *(lnEDU)* is not statistically significant.

4.2 Robustness checks

We assess the robustness of our results using alternative methods. First, we employ The DOLS estimator, which was introduced by Stock and Watson (1993) and later extended by Kao and Chiang (2000), the DOLS follows a parametric approach and counters these issues through lags and adding the leads of explanatory variables. To estimate the short – run model, we employ the MG (Mean Group) estimation, which imposes no restrictions on coefficient, both in the long- as well as in the short – run.

Additionally, we differentiate between countries with high and low R&D expenditures to account for heterogeneity related to the level of R&D expenditure in our sample. Specifically, we divide our sample based on whether R&D expenditure (RD) is above the median (RD > 0.875%; high R&D) or below the median (RD < 0.875%; low R&D). [10]

Table 7 presents the results of short- and long – run estimations using DOLS and MG estimators.

Tab. 7 – DOLS and MG Estimation Results

	Long – run	
	DOLS	**MG**
$lnRD$	-0.2***	0.128
	(0.052)	(0.087)
$lnGDP$	-0.412***	-0.064
	(0.137)	(0.207)
$lnEDU$	1.324	-2.246
	(0.937)	(1.583)
Adjusted R – squared	0.542	
Countries	67	67
Observations	1,337	1,237
	Short – run	
	MG	
ECT	-0.513***	
	(0.045)	
$\Delta lnRD$	0.000	
	(0.004)	
$\Delta lnGDP$	-0.316***	
	(0.032)	
$\Delta lnEDU$	-0.067	
	(0.087)	
Countries	67	
Observations	1,273	

Note: ***, ** and * indicate the significance levels of 1%, 5% and 10%, respectively. The standard errors are reported in brackets.
Source: Authors' calculation.

[10] The mean of R&D expenditure is 1.187, with minimum and maximum values of 0.23 and 5.705, respectively.

In the long – run, the negative and significant influence of R&D expenditure on the size of shadow economy, is confirmed. In the short – run, the estimated coefficient is not statistically significant. The value of the error correction term is statistically significant at 1% level and equal to -0.513.

The list of countries clustered according to R&D is presented in Table 8, whereas Table 9 shows the results of long- and short – run estimation for high and low R&D expenditure countries.

Tab. 8 – Countries Clustered by R&D

High R&D Expenditure
Australia; Austria; Belgium; Brazil; Canada; China; Croatia; Czech Republic; Denmark; Estonia; Finland; France; Germany; Hungary; Iceland; Ireland; Israel; Italy; Japan; Korea, Rep.; Luxembourg; Malaysia; Netherlands; New Zealand; Norway; Portugal; Russian Federation; Singapore; Slovenia; Spain; Sweden; United Kingdom; United States.
Low R&D Expenditure
Armenia; Azerbaijan; Belarus; Bulgaria; Chile; Colombia; Costa Rica; Cyprus; Egypt, Arab Rep.; El Salvador; Georgia; Greece; Guatemala; India; Iran, Islamic Rep.; Kazakhstan; Kuwait; Kyrgyz Republic; Mexico; Moldova; Mongolia; Paraguay; Peru; Poland; Romania; Slovak Republic; South Africa, Tajikistan; Thailand; Trinidad and Tobago; Tunisia; Turkey; Ukraine; Uruguay.

Note: we divide our sample based on whether R&D expenditure (RD) is above the median (RD > 0.875%; high R&D) or below the median (RD < 0.875%; low R&D).

Tab. 9 – Sub-samples Estimation Results

	Long – run			
	High R&D countries		**Low R&D countries**	
	FMOLS	**PMG**	**FMOLS**	**PMG**
$lnRD$	-0.323***	-0.032***	-0.112*	0.000
	(0.120)	(0.009)	(0.059)	(0.002)
$lnGDP$	-0.463**	-0.265***	-0.429***	-0.194***
	(0.194)	(0.026)	(0.136)	(0.010)
$lnEDU$	1.154	-0.373***	2.111**	0.007
	(1.358)	(0.112)	(0.904)	(0.054)
Adj. R – squared	0.339		0.148	
Countries	33	33	34	34
Observations	659	627	679	646

	Short – run	
	High R&D countries	**Low R&D countries**
	PMG	**PMG**
ECT	-0.133***	-0.26***
	(0.022)	(0.039)
$\Delta lnRD$	0.002	0.010**
	(0.009)	(0.005)
$\Delta lnGDP$	-0.582***	-0.32***
	(0.052)	(0.034)
$\Delta lnEDU$	-0.155	-0.096
	(0.135)	(0.088)
Countries	33	34
Observations	627	646

Note: ***, ** and * indicate the significance levels of 1%, 5% and 10%, respectively. The standard errors are reported in brackets.
Source: Authors' calculation.

In the long – run the R&D expenditure negatively affects the shadow economy in each subsample. The R&D expenditure is more elastic with shadow economy, in high R&D countries. In the short – run the R&D expenditure has a positive effect on the size of the shadow economy only for low R&D countries, this impact is not statistically significant for high R&D countries.

5. Conclusions

The economic literature on tax evasion and shadow economy has flourished for over half a century, and understanding this phenomenon is vital for policymakers due to its negative impact on the economy. Economists widely agree that the primary driver of tax evasion is the tax burden. Both theoretical and empirical literature supports a positive relationship between tax burden and evaded income, which is intuitively logical. Nonetheless, the literature has found evidence of other factors that could have an important influence on the size of shadow economy. The technological progress, which is considered the main factor driving economic growth, must necessarily have an important influence on tax evasion and shadow economy. The connection between technological progress and the shadow is still unexplored in the literature.

This paper has set out a model of tax evasion, linking technological progress, labour efforts and tax evasion. It has been shown that, consistent with tax evasion literature, there exists a positive relationship between taxation and tax evasion. The main result is that, both the individual tax compliance decision, and therefore the optimal level of tax evasion, and the individual's sensitivity to tax rate changes, is strongly affected by the return gap on the accessible production technologies' changes.

Since many authors claim that, individual tax evasion choice is affected not only by tax rate but also by tax morale, social customs, social conformity and other psychics variables, we have extended our framework with social customs, and it has been shown that the correlation between the return gap on the accessible production technologies' changes and the optimal level of tax evasion, it is even stronger.

Furthermore, we conduct an empirical analysis which provides consistent support for the theoretical findings. We investigate the role of R&D expenditure used as a proxy for the technological progress, in determining the size of the shadow economy, in a sample of 67 countries from 2001 to 2020, using the Pedroni (1999, 2004) and Westerlund (2005) panel cointegration tests. Both the panel cointegration tests provides evidence of a long – run equilibrium among our variables. This long – run cointegration relationship is confirmed using alternative estimation method in form of the DOLS and FMOLS. A VECM is employed to analyze the short – run effect, which is estimated using PMG and MG estimators; the short – run parameters are not statistically significant. The negative impact of R&D expenditure on the size of the shadow economy is larger in high R&D expenditure countries; instead for the low R&D countries, in the short – run, R&D expenditure positively affects the size of the shadow economy.

Considering that spending in research and development not only provides many social and economic benefits but is also an important driver of economic growth, from a policy perspective, the role of R&D expenditure in curbing the size of shadow economy is of paramount importance. These findings are essential for policymakers to formulate the right policies.

References

Afonso, O., Neves, P. C., Pinto, T. (2020). The non-observed economy and economic growth: A meta-analysis. *Economic Systems*, 44(1), 100746.

Aghion, P., Howitt, P. (1992). A model of growth through creative destruction. *Econometrica* 60(2), 323-351.

Akerlof, G. A. (1980). A Theory of Social Custom, of Which Unemployment May be One Consequence. *The Quarterly Journal of Economics,* 94(4), 749-775

Allingham, M. G., Sandmo, A. (1972). Income tax evasion: a theoretical analysis. *Journal of Public Economics,* 1, 323-338.

Alm, J., Jackson, B. R., McKee, M. (1992). Estimating the Determinants of Taxpayer Compliance with Experimental Data. *National Tax Journal*, 45(1), 107-114.

Amanova, M. (2023). Do R&D and digitalization affect shadow economy in Uzbekistan? *In Proceedings of the 6th International Conference on Future Networks & Distributed Systems (ICFNDS '22).* Association for Computing Machinery, New York, NY*, USA*, 258–264. https://doi.org/10.1145/3584202.3584240

Andersen, P. (1977). Tax evasion and labour supply. *Scandinavian Journal of Economics*, 79, 375-383.

Andreoni, J., Erard, B., Feinstein, J. (1998). Tax compliance. *Journal of Economic Literature* 36(2), 818-860.

Arcand, J. L., Graziosi, G. R. (2005). Tax Compliance and Rank Dependent Expected Utility. *The Geneva Risk and Insurance Review*, 30(1), 57-69.

Babic, V., Zarić, S. (2022). Tax evasion and educational level: evidence from the European Union Countries. *KNOWLEDGE – International Journal*, 51(1), 17-21.

Battisti, G., Stoneman, P. (2005). The intra-firm diffusion of new process technologies. *International Journal of Industrial Organization* 23, 1-22.

Berdiev, A. N., Saunoris, J. W. (2016). Financial development and the shadow economy: A panel VAR analysis. *Econ. Model.* 57, 197-207.

Bernasconi, M., Zanardi A. (2004). Tax evasion, tax rates, and reference dependence. *Finanzarchiv*, 60, 422-455.

Bethencourt, C., Kunze, L. (2019). Tax evasion, social norms, and economic growth. *Journal of public economic theory*, 21(2), 332-346.

Bittencourt, M., Gupta, R., Stander, L. (2014). Tax evasion, financial development and inflation: Theory and empirical evidence. *J. Bank. Financ.,* 41, 194-208.

Blackburn, K., Bose, N., Capasso, S. (2012). Tax evasion, the underground economy and financial development. *J. Econ. Behav. Organ.,* 83(2), 243-253.

Bonnet, F., Venkatesh, S. (2016). Poverty and informal economies. In D. Brady & L. Burton (Eds.), *Oxford handbook of the social science of poverty* (pp. 637-659), Oxford University Press.

Buehn, A., Farzanegan, M. R. (2013). Impact of education on the shadow economy: Institutions matter. *Economics Bulletin*, 33, 2052-2063.

Caballe, J., Panades, J. (2004). Inflation, tax evasion, and the distribution of consumption. *Journal of Macroeconomics, Elsevier,* 26(4), 567-595.

Capasso, S., Jappelli, T. (2013). Financial development and the underground economy. *J. Dev. Econ.,* 101, 167-178.

Chong, A., & Gradstein, M. (2007). Inequality and informality. *Journal of Public Economics*, 91(1–2), 159-179.

Clotfelter, C. T. (1983).Tax Evasion and Tax Rates: An Analysis of Individual Returns. *Review of Economics and Statistics* (Cambridge, Massachusetts), 65, 363-373.

Cowell, F. (1985). Tax evasion with labour income. *Journal of Public Economics*, 26(1), 19-34.

Cowell, F. (1990). Tax Sheltering and the Cost of Evasion. *Oxford Economic Papers*, 42(1), 231-243.

Cowell, F., Gordon, J. P. F. (1988). Unwillingness to pay: Tax evasion and public good provision. *Journal of Public Economics*, 36(3), 305-321.

Cremer, H., Gahvari, F. (1994). Tax evasion, concealment and the optimal linear income tax. *Scandinavian Journal of Economics,* 96, 219-239.

Dell'Anno, R. (2007). The Shadow Economy in Portugal: An Analysis with the MIMIC Approach. *Journal of Applied Economics*, 10, 253-277.

Dell'Anno, R. (2021a). Theoretical approaches to the phenomenon of informality. In S. M. Ruesga, J. Baquero, & J. L. Delgado (Eds.), *Dialogues on Socioeconomy: Informality in Latin America* (chapter 3) (pp. 117-134). Tirant lo Blanc.

Dell'Anno, R. (2021b). Inequality and informality in transition and emerging countries. *IZA World of Labor*, *325*, https://doi.org/10.15185/izawol.325.v2

Dell'Anno, R., Gomez-Antonio, M., Alanon Pardo A. (2007). Shadow Economy in three different Mediterranean Countries: France, Spain and Greece. A MIMIC Approach. *Empirical Economics*, 33, 51-84.

Dell'Anno, R. (2009). Tax evasion, tax morale and policy maker's effectiveness. *Journal of Behavioral and Experimental Economics*, 38(6), 988-997.

Dell'Anno, R. (2022). Theories and Definitions of the Informal Economy: A Survey. *Journal of Economic Surveys*. (doi: 10.1111/joes.12487).

Dhami, S., Al-Nowaihi, A. (2007). Why do people pay taxes? Prospect theory versus expected utility theory. *Journal of Economic Behavior and Organization*, 64(1), 171-192.

Elgin, C., Kose, M. A., Ohnsorge, F., Yu, S. (2021). Growing apart or moving together? Synchronization of informal and formal economy business cycles. In F. Ohnsorge & S. Yu (Eds.), *The long shadow of informality: Challenges and policies.* World Bank.

Elgin, C., M. A. Kose, F. Ohnsorge, and S. Yu. (2021). Understanding Informality. *CERP Discussion Paper 16497, Centre for Economic Policy Research*, London.
Ferreira-Tiryaki, G. (2008). The informal economy and business cycles. *Journal of Applied Economics*, *11*, 91-117.
Fortin, B., Lacroix, G., Villeval, M. C. (2007). Tax evasion and social interactions. *Journal of Public Economics*, 91(11-12), 2089-2112.
Franzoni, L. A. (2008). Tax Compliance. In B. Bouckaert and G. De Geest (Eds.), *Encyclopedia of Law and Economics*. Edward Elgar.
Giles, D. E. A. (1999a). Measuring the Hidden Economy: Implications for Econometric Modelling. *Economic Journal*, 109/3, 370-380.
Gilligan, G., Richardson, G. (2005). Perceptions of tax fairness and tax compliance in Australia and Hong Kong- a preliminary study. *Journal of Financial Crime*, 12(4), 331-343.
Goel, R. K., Nelson, M. A. (2016). Shining a light on the shadows: Identifying robust determinants of the shadow economy. *Economic Modelling*, 58(C), 351-364.
Gordon, J. P. F. (1989). Individual morality and reputation costs as deterrents to tax evasion. *European Economic Review,* 33, 797-805.
Granda-Carvajal, C. (2010). The unofficial economy and the business cycle: A test for theories. *International Economic Journal*, 24(4), 573-586.
Grossman, G. M., Helpman, E. (1991). Innovation and Growth in the Global Economy. *MIT Press*.
Hashimzade, N., Myles, G. D., Tran-Nam, B. (2013). Applications of behavioural economics to tax evasion. *Journal of Economic Surveys*, 27(5), 941-977.
Hassan, M., Schneider, F. (2016). Size and Development of the Shadow Economies of 157 Worldwide Countries: Updated and New Measures from 1999 to 2013. *J Glob Econ*, 4(218), 2.
Hoechle, D. (2007). Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence. *The Stata Journal*, 7(3), 281-312.
Immervoll, H. (2000). The impact of inflation on income tax and social insurance contributions in Europe. *EUROMOD Working Papers* EM2/00.
Isachsen, A. J., Samuelson, S. O., Strom, S. (1985). The Behavior of Tax Evaders*. The Economics of the Shadow Economy*, 1.
Isachsen, A. J.,Strom, S. (1980). The hidden economy: The labour market and tax evasion. *Scandinavian Journal of Economics*, 82, 304-11
Jackson, B. R., Milliron, V. C. (1986). Tax compliance research: Findings, problems, and prospects. *Journal of Accounting Literature*, 5(1), 125-165.
Johnson, S., Kaufmann, D., Zoido-Lobatón, P. (1998a). Regulatory Discretion and the Unofficial Economy. *The American Economic Review*, 88(2): 387-392.
Kao, C., Chiang, M. H. (2000). On the estimation and inference of a cointegrated regression in panel data. *Advances in Econometrics*.
Kaplow, L. (1990). Optimal taxation with costly enforcement and evasion. *Journal of Public Economics*, 43(2), 221-236.
Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., Saez, E. (2011). Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark. *Econometrica*, 79(3), 651-92.
Kwok, B. Y. S., Yip, R. W. Y. (2018). Is tax education good or evil for boosting tax compliance? Evidence from Hong Kong. *Asian Economic Journal*, 32(4), 359-386.
Layson, S. (1983). Homicide and deterrence: another view of the Canadian time-series evidence. *Can. J. Econ.,* 16, 52-73.
Mansfield, E. (1963). Intra-Firm Rates of Diffusion of an Innovation. *The Review of Economics and Statistics,* 45.
Markellos, R. N., Psychoyios, D., Schneider, F. (2016). Sovereign debt markets in light of the shadow economy. *European Journal of Operational Research*, 252(1), 220-231.
McGee, R. W., Tyler, M. (2006). Tax Evasion and Ethics: A Demographic Study of 33 Countries. Available at SSRN: https://ssrn.com/abstract=940505 or http://dx.doi.org/10.2139/ssrn.940505

Medina, L., Schneider, F. (2018). Shadow Economies around the World: What Did We Learn over the Last 20 Years?. *African Department, IMF Working Papers* 18/17.

Myles, G., Naylor, R. (1996). A model of tax evasion with group conformity and social customs. *European Journal of Political Economy* 12(1), 49-66.

Naylor, R. (1990). A social custom model of collective action. European *Journal of Political Economy,* 6(2), 201-216.

Nevzorova, E. N., Kireenko, A. P., Leontyeva, Y. V. (2018). If Development of Technologies can Win Shadow Economy. In *2018 XIV International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE),* 422-428. IEEE.

Pedroni, P. (1999). Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics,* 61, 653-670.

Pedroni, P. (2000). Fully modified OLS for heterogenous cointegrated panels. In B. H. Baltagi, T. B. Fomby, & R. C. Hill (Eds.), *Advances in econometrics* (99-130). Amsterdam: JAI Press.

Pedroni, P. (2004). Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. *Econometric Theory* 20, 597-625.

Pencavel, J. H. (1979). A note on income tax evasion, labor supply, and nonlinear tax schedules. *Journal of Public Economics, Elsevier,* 12, 115-124.

Pesaran, H., Shin, Y., Smith, R. P. (1997). Pooled Estimation of Long-Run Relationships in Dynamic Heterogeneous Panels. *Department of Applied Economics University of Cambridge.*

Pesaran, H., Shin, Y., Smith, R. P. (1999). Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American Statistical Association*, 94(446), 621-634.

Pesaran, M. H. (2004). General diagnostic tests for cross section dependence in panels. *IZA Discussion Paper,* 1240. Cambridge.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, 22(2), 265-312.

Philips P. C. B., Hansen, B. E. (1990). Statistical Inference in Instrumental Variables Regression with I(1), Processes. *Review of Economic Studies*, 57, 99-125.

Phillips, P. C. B., Moon, H. R. (1999). Linear Regression Limit Theory for Nonstationary Panel Data. *Econometrica*, 67, 1057-1111.

Pyle, D. J. (1991). The Economics of Taxpayer Compliance. *Journal of Economic Surveys*, 5(2), 163-98.

Romer, P. M. (1990). Endogenous technological change. *J. Polit. Econ.,* 98(5, Part 2). S71- S102.

Ross, A. M., McGee, R. W. (2012). Education level and ethical attitude toward tax evasion: A six-country study. *J. Legal Ethical & Regul. Isses*, *15*, 93.

Sandmo, A. (1981). Income tax evasion, labour supply, and the equity-efficiency trade-off. *Journal of Public Economics*, 16(3), 265-288.

Sandmo, A. (2005). The Theory of Tax Evasion: A Retrospective View. *National Tax Journal*, 58(4), 643-63.

Sandmo, A. (2012). An evasive topic: theorizing about the hidden economy. *Int Tax Public Finance,* 19, 5-24.

Schneider, F. (2003). The shadow economy. In , C. K. Rowley and F. Schneider (Eds.), *Encyclopedia of Public Choice*, Kluwer Academic Publishers, Dordrecht.

Schneider, F. (2005). Shadow economies around the world: what do we really know?. *European Journal of Political Economy*, 21(4), 598-642.

Schneider, F. (2012). The Shadow Economy and Work in the Shadow: What Do We (Not) Know? IZA Discussion Papers 6423.

Schneider, F., Buehn, A., Montenegro, C. E. (2010). New Estimates for the Shadow Economies all over the World. *International Economic Journal*, 24(4), 443-461.

Schneider, F., Williams, C. (2013). The Shadow Economy. *Institute of Economic Affairs*.

Schroyen, F. (1997). Pareto efficient income taxation under costly monitoring. *Journal of Public Economics*, 65(3), 343-366.

Shahbaz, M. (2010). Income inequality-economic growth and non-linearity: a case of Pakistan. *Int. J. Soc. Econ.,* 37(8), 613-636.

Slemrod, J. (1985). An empirical test for tax evasion. *Review of Economics and Statistics*, 67(2), 232-238.
Slemrod, J. (2007). Cheating ourselves: the economics of tax evasion. *J. Econ. Perspect*, 21(Winter (1)), 25-48.
Slemrod, J., Yitzhaki, S. (2002). Tax Avoidance, Evasion, and Administration. *Handbooks in Economics*, 4. Amsterdam; London and New York: Elsevier Science, North-Holland.
Snow, A., Warren, R. (2005). Ambiguity about Audit Probability, Tax Compliance, and Taxpayer Welfare. *Economic Inquiry*, 43(4), 865-871.
Stock, J. H., Watson, M. W. (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica,* 61(4), 783e820.
Tanzi, V. (1999). Uses and Abuses of Estimates of the Underground Economy. *Economic Journal*, 109(3), 338-347.
Thomas, J. J. (1992). Informal Economic Activity, *LSE, Handbooks in Economics*, Harvester Wheatsheaf, London.
Torgler, B. (2004). Cross-culture comparison of tax morale and tax compliance: evidence from Costa Rica and Switzerland. *International Journal of Comparative Sociology*, *45*, 17-43.
Uyar, A., Benkraiem, R., Nimer, K., Schneider F. (2022). Education, Digitalization and Tax Evasion: International Evidence. *International Journal of Public Administration*.
Von Neumann, J., Morgenstern, O. (1947). Theory of games and economic behavior (2nd rev. ed.). *Princeton University Press*.
Westerlund, J. (2005). New simple tests for panel cointegration. *Econometric Reviews,* 24, 297-316.
Yaniv, G. (1999). Tax Compliance and Advance Tax Payments: A Prospect Theory Analysis. *National Tax Journal*, 52(4), 753-64.
Yitzhaki, S. (1974). A note on income tax evasion: a theoretical analysis*. Journal of Public Economics,* 3, 201-202.
Yitzhaki, S. (1987). On the Excess Burden of Tax Evasion. *Public Finance Quarterly*, 15(2), 123-137.

Appendix A

The individual maximization problem is of the following form:

$$\max_{e_p;e_s} u(c; e_p; e_s): c - \frac{(e_p)^2}{2} - \frac{(e_s)^2}{2}$$

subject to

$$c = w_p(1-t) + w_s(1-p) + pw_s(1-F)$$

$$e_p \geq 0; e_s \geq 0; e_p + e_s = TL$$

Since $w_p = \theta_p e_p$ and $w_s = \theta_s e_s$, we can rewrite the utility function as:

$$u(\cdot) = \theta_p(1-t)(TL - e_s) + \theta_s(1-pF)e_s - \frac{(TL - e_s)^2}{2} - \frac{(e_s)^2}{2}$$

The first order condition is:

$$\frac{\partial u}{\partial e_s} = -\theta_p(1-t) + \theta_s(1-pF) + TL - e_s - e_s$$

Equating the first order condition to zero, and solving for e_s:

$$e_s{}^* = \frac{TL - \theta_p(1-t) + \theta_s(1-pF)}{2}$$

$\frac{\partial u}{\partial e_s} > 0$ when $e_s < \frac{TL-\theta_p(1-t)+\theta_s(1-pF)}{2}$ and $\frac{\partial u}{\partial e_s} < 0$ when $e_s > \frac{TL-\theta_p(1-t)+\theta_s(1-pF)}{2}$, therefore $e_s{}^*$ is a point of maximum.

Checking the constraints:

$$e_s{}^* > 0 \Rightarrow \theta_p < \frac{\theta_s(1-pF) + TL}{(1-t)}$$

$$e_s{}^* < TL \Rightarrow \theta_p > \frac{\theta_s(1-pF) - TL}{(1-t)}$$

Therefore:

If $\frac{\theta_s(1-pF)-TL}{(1-t)} < \theta_p < \frac{\theta_s(1-pF)+TL}{(1-t)}$, then $e_s{}^* = \frac{TL-\theta_p(1-t)+\theta_s(1-pF)}{2}$, and $e_p{}^* = TL - \frac{TL-\theta_p(1-t)+\theta_s(1-pF)}{2} = \frac{TL+\theta_p(1-t)-\theta_s(1-pF)}{2}$;

If $\theta_p \leq \frac{\theta_s(1-pF)-TL}{(1-t)}$ then $e_s{}^* = TL$ and $e_p{}^* = TL - TL = 0$;

If $\theta_p \geq \frac{\theta_s(1-pF)+TL}{(1-t)}$ then $e_s{}^* = 0$ and $e_p{}^* = TL$.

Appendix B

Given that $w_s^* = \theta_s {e_s}^* = \frac{\theta_s TL - \theta_p \theta_s (1-t) + {\theta_s}^2 (1-pF)}{2}$; the first derivates of w_s^* with respect to all other parameters are:

$$\frac{\partial w_s^*}{\partial \theta_p} = -\frac{\theta_s (1-t)}{2}$$

$$\frac{\partial w_s^*}{\partial \theta_s} = \frac{TL - \theta_p (1-t) + 2\theta_s (1-F)}{2}$$

$\frac{\partial w_s^*}{\partial \theta_s} > 0$ when $\theta_p < \frac{TL + 2\theta_s (1-pF)}{(1-t)}$; which is always satisfied since the existence condition of ${e_s}^*$: $\theta_p < \frac{\theta_s (1-pF) + TL}{(1-t)}$.

$$\frac{\partial w_s^*}{\partial t} = \theta_p \theta_s$$

Appendix C

Now, we have to show that:

$$u_e^*(c;e_p^*;e_s^*) = \tfrac{1}{2}TL\left[\theta_p(1-t)+\theta_s(1-pF)-\tfrac{1}{2}TL\right]+\tfrac{1}{4}\left[\theta_p(1-t)-\theta_s(1-pF)\right]^2$$

Starting from the utility function:

$$u(c;e_p;e_s): c-\frac{(e_p)^2}{2}-\frac{(e_s)^2}{2}$$

And given that:

$$c=\theta_p e_p(1-t)+\theta_s e_s(1-pF)$$

$$e_p^*=\frac{TL+\theta_p(1-t)-\theta_s(1-pF)}{2};\ e_s^*=\frac{TL-\theta_p(1-t)+\theta_s(1-pF)}{2}$$

Then:

$$\begin{aligned}u_e^*(c;e_p^*;e_s^*) &= \theta_p(1-t)\frac{TL+\theta_p(1-t)-\theta_s(1-pF)}{2}\\ &+\theta_s(1-pF)\frac{TL-\theta_p(1-t)+\theta_s(1-pF)}{2}-\frac{1}{2}\left(\frac{TL+\theta_p(1-t)-\theta_s(1-pF)}{2}\right)^2\\ &-\frac{1}{2}\left(\frac{TL-\theta_p(1-t)+\theta_s(1-pF)}{2}\right)^2\end{aligned}$$

$$\begin{aligned}u_e^*(c;e_p^*;e_s^*) &= \left[\theta_p(1-t)+\theta_s(1-pF)\right]\frac{TL}{2}\\ &+\frac{1}{2}\left[{\theta_p}^2(1-t)^2-2\theta_p\theta_s(1-t)(1-pF)+{\theta_s}^2(1-pF)^2\right]\\ &-\frac{1}{8}\left[TL^2+2TL\left(\theta_p(1-t)-\theta_s(1-pF)\right)+\left(\theta_p(1-t)-\theta_s(1-pF)\right)^2\right]\\ &-\frac{1}{8}\left[TL^2-2TL\left(\theta_p(1-t)-\theta_s(1-pF)\right)+\left(\theta_p(1-t)-\theta_s(1-pF)\right)^2\right]\end{aligned}$$

$$\begin{aligned}u_e^*(c;e_p^*;e_s^*) &= \left[\theta_p(1-t)+\theta_s(1-pF)\right]\frac{TL}{2}+\frac{1}{2}\left(\theta_p(1-t)-\theta_s(1-pF)\right)^2\\ &-\frac{1}{8}\left[2TL^2+2\left(\theta_p(1-t)-\theta_s(1-pF)\right)^2\right]\end{aligned}$$

$$\begin{aligned}u_e^*(c;e_p^*;e_s^*) &= \left[\theta_p(1-t)+\theta_s(1-pF)\right]\frac{TL}{2}-\frac{TL^2}{4}+\frac{1}{2}\left(\theta_p(1-t)-\theta_s(1-pF)\right)^2\\ &-\frac{1}{4}\left(\theta_p(1-t)-\theta_s(1-pF)\right)^2\end{aligned}$$

Therefore:

$$u_e^*(c;e_p^*;e_s^*) = \tfrac{1}{2}TL\left[\theta_p(1-t)+\theta_s(1-pF)-\tfrac{1}{2}TL\right]+\tfrac{1}{4}\left[\theta_p(1-t)-\theta_s(1-pF)\right]^2$$

Since:

$$u_{ne}^*(c,e_{p,ne}^*)=\theta_p(1-t)TL-\frac{TL^2}{2}+aR(1-\mu)+b$$

We have that $u_{ne}^*=u_e^*$; if:

$$\theta_p(1-t)TL-\tfrac{TL^2}{2}+aR(1-\mu)+b=\tfrac{1}{2}TL\left[\theta_p(1-t)+\theta_s(1-pF)-\tfrac{1}{2}TL\right]+\tfrac{1}{4}\left[\theta_p(1-t)-\theta_s(1-pF)\right]^2$$

Solving for μ:

$$aR(1-\mu)+b$$
$$=\frac{1}{2}\theta_p(1-t)TL-\theta_p(1-t)TL-\frac{1}{4}TL^2+\frac{1}{2}TL^2+\frac{1}{2}\theta_s(1-pF)TL$$
$$+\frac{1}{4}\left[\theta_p(1-t)-\theta_s(1-pF)\right]^2$$
$$aR(1-\mu)+b=\frac{1}{4}TL^2-\frac{1}{2}\theta_p(1-t)TL+\frac{1}{2}\theta_s(1-pF)TL+\frac{1}{4}\left[\theta_p(1-t)-\theta_s(1-pF)\right]^2$$
$$aR(1-\mu)=\frac{1}{4}\left[TL^2-2TL\left(\theta_p(1-t)-\theta_s(1-pF)\right)+\left[\theta_p(1-t)-\theta_s(1-pF)\right]^2\right]-b$$
$$\mu^*=1-\frac{TL^2-2TL\left[\theta_p(1-t)-\theta_s(1-pF)\right]+\left[\theta_p(1-t)-\theta_s(1-pF)\right]^2-4b}{4aR}$$

Now we prove that $\frac{\partial\mu^*}{\partial t}<0$; $\frac{\partial\mu^*}{\partial\theta_p}>0$ and $\frac{\partial\mu^*}{\partial\theta_s}<0$.

$\frac{\partial\mu^*}{\partial t}=-\frac{1}{4aR}\left[2TL\theta_p-2\theta_p\left(\theta_p(1-t)-\theta_s(1-pF)\right)\right]$ which is lower than zero when: $TL-\theta_p(1-t)+\theta_s(1-pF)>0\ \Rightarrow\ \theta_p<\frac{\theta_s(1-pF)+TL}{1-t}$; which is the existence condition of $e_s^*=\frac{TL-\theta_p(1-t)+\theta_s(1-pF)}{2}$; therefore $\frac{\partial\mu^*}{\partial t}<0$.

$\frac{\partial\mu^*}{\partial\theta_p}=-\frac{1}{4aR}\left[-2TL(1-t)+2(1-t)\left(\theta_p(1-t)-\theta_s(1-pF)\right)\right]$ which is greater than zero when: $-TL+\theta_p(1-t)-\theta_s(1-pF)<0\Rightarrow\theta_p<\frac{\theta_s(1-pF)+TL}{1-t}$; which is the existence condition of $e_s^*=\frac{TL-\theta_p(1-t)+\theta_s(1-pF)}{2}$; therefore $\frac{\partial\mu^*}{\partial\theta_p}>0$.

$\frac{\partial\mu^*}{\partial\theta_s}=-\frac{1}{4aR}\left[2TL(1-pF)-2(1-pF)\left(\theta_p(1-t)-\theta_s(1-pF)\right)\right]$ which is lower than zero when: $TL-\theta_p(1-t)+\theta_s(1-pF)>0\Rightarrow\theta_p<\frac{\theta_s(1-pF)+TL}{1-t}$; which is the existence condition of $e_s^*=\frac{TL-\theta_p(1-t)+\theta_s(1-pF)}{2}$; therefore $\frac{\partial\mu^*}{\partial\theta_s}<0$.