



JOEEP

e-ISSN: 2651-5318

Journal Homepage: <http://dergipark.org.tr/joep>

Araştırma Makalesi • Research Article

Evaluating Tax Compliance Performance in Neo-Liberal Türkiye: A Non-Parametric Efficiency Analysis from 1982 to 2021

Neo-Liberal Türkiye'de Vergi Uyum Performansının Değerlendirilmesi: 1982'den 2021'e Parametrik Olmayan Bir Etkinlik Analizi

Öner Gümüş^{a,*}

^aDoç. Dr., Kütahya Dumlupınar University-Bilecik Şeyh Edebalı University, Kütahya/ Türkiye

ORCID: 0000-0002-3231-3293

MAKALE BİLGİSİ

Makale Geçmişi:

Başvuru tarihi: 7 Mayıs 2025

Düzeltilme tarihi: 1 Şubat 2026

Kabul tarihi: 24 Mayıs 2026

Anahtar Kelimeler:

Vergi uyumu

Caydırıcılık teorisi

Mali mübadele teorisi

Kaygan zemin çerçevesi

Vergi tahsil performansı

ARTICLE INFO

Article history:

Received: May 7, 2025

Received in revised form: Feb 1, 2026

Accepted: May 24, 2026

Keywords:

Tax compliance

Deterrence theory

Fiscal exchange theory

Slippery slope framework

Tax collection performance

ÖZ

Bu çalışma 1982-2021 yılları arasında parametrik olmayan yöntemlerin kullanılması suretiyle Türkiye'de vergi uyum etkinliğini analiz etmektedir. Çalışmada veri zarflama analizi ve Malmquist endeksi uygulanarak onar yıllık dört dönem ele alınmıştır. Bulgular, ekonomik büyüklüğün vergi tahsilat performansını doğrudan arttırmadığını göstermektedir. Teknolojik ilerleme yüksek olmasına rağmen operasyonel etkinlik düşüş göstermiştir. Bootstrap analizi ile yapılan yanlışlık düzeltmesinde ise etkinliğin gerilediği görülmüştür. Sonuç olarak kurumsal güven, yönetim kalitesi ve kamu hizmetlerinin niteliğinin vergi uyumunun temel belirleyicileri olduğu ortaya çıkmıştır.

ABSTRACT

This study analyzes tax compliance efficiency in Türkiye between 1982 and 2021 using non-parametric methods. Data envelopment analysis and the Malmquist index were applied across four ten-year periods. The results indicate that economic size does not directly increase tax collection performance. Despite substantial technological progress, operational efficiency has declined. Bias correction through bootstrap analysis further revealed a decline in efficiency. Overall, institutional trust, governance quality, and the quality of public services have emerged as key determinants of tax compliance.

1. Introduction

Türkiye's neoliberal transformation after 1980 represents not only economic liberalization but also a radical

restructuring of public finances and the tax system. This new era, which began with the decisions of January 24, 1980, and was institutionalized with the military intervention of September 12, narrowed the state's role in the economy while making tax collection more dependent on market

* Sorumlu yazar/Corresponding author.

e-posta: oner.gumus@dpu.edu.tr - oner.gumus@bilecik.edu.tr

Atf/Cite as: Gümüş, Ö. (2026). Evaluating Tax Compliance Performance in Neo-Liberal Türkiye: A Non-Parametric Efficiency Analysis from 1982 to 2021. *Journal of Emerging Economies and Policy*, 11(1), 189-207.

This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors.

dynamics and voluntary compliance mechanisms. In this context, this study, covering the period from 1982 to 2021, evaluates the impact of the aforementioned structural transformation on tax compliance performance within an efficiency analysis framework, using macroeconomic and demographic capacity indicators as inputs and the collection/accrual ratio as the output.

Gross Domestic Product (GDP) is included in the model because it reflects the tax capacity of the country and population reflects the number of potential taxpayers. In this study, GDP and population are conceptualized as input indicators representing the macro-fiscal capacity (potential tax base, administrative-financial capacity, and scale) in which the tax administration operates. The literature on tax capacity and tax effort emphasizes that a country's potential level of tax collection can be captured by structural indicators such as income level and demographic scale (Fenochietto & Pessino, 2013). Similarly, in frontier-based tax performance studies, the extent to which macro capacity indicators can be effectively converted into actual collection performance is evaluated using DEA/SFA approaches (Afonso et al., 2025; Anastasiou et al., 2023).

Tax policies in Türkiye have undergone significant changes over time. Post-1980 reforms, the predominance of indirect taxes, policies to combat the informal economy, and tax amnesties have made Türkiye a country that needs to be analyzed in terms of tax compliance. The data spans from 1982 to 2021. We selected 1982 as the initial year because it marked the transition from coup d'état to democracy. This shift occurred because elections were held in 1982, leading to administrations assuming office through electoral processes. The study ends in 2021 because sudden policy changes and structural breaks following the pandemics could affect the data envelopment analysis results and yield misleading conclusions.

The aim of the study is to assess the overall effectiveness of tax compliance in Türkiye. To this end, data envelopment analysis was employed. The limitations of the study are discussed in detail in the data set and methodology section. This study evaluates the efficiency of tax compliance in Türkiye and makes a methodological contribution to the existing literature through the use of data envelopment analysis. In addition, it provides policy-relevant findings and inferences that may help to enhance the effectiveness of tax policies.

In this context, we seek answers to the following research questions:

- What is the level of tax compliance efficiency in Türkiye between 1982 and 2021?
- How has the efficiency of tax compliance changed over the years, and to what extent is it associated with the inputs used in the study?
- Which variables not used in the study may be related to the effectiveness of tax compliance?

- In which periods has Türkiye's tax collection efficiency increased or decreased in terms of the efficiency of tax compliance?

Answering these questions will provide insights for evaluating the effectiveness of the policy in Türkiye. We tested the following hypotheses to address these questions:

- **H₀**: GDP and population are not direct determinants of tax compliance.
- **H₁**: GDP and population have a direct and significant effect on tax compliance.
- **H₂**: Technological developments are a critical factor that increases tax compliance independent of economic growth.
- **H₃**: Tax administration and scale efficiency are more important factors than economic growth.
- **H₄**: Tax compliance is affected by structural factors as well as economic indicators.

The hypotheses were derived by jointly evaluating two literature axes: (i) the tax capacity/effort approach and (ii) behavioral-institutional alignment approaches. Tax capacity literature suggests that increases in economic and demographic capacity can support collection performance by expanding the potential tax base. Therefore, H₁ is based on testing this capacity–performance expectation (Fenochietto & Pessino, 2013; Afonso et al., 2025). In contrast, deterrence, fiscal exchange, and slippery slope frameworks emphasize that compliance is shaped not only by economic size but also by structural elements such as regulatory power, trust in institutions, perceptions of justice, and governance quality. Therefore, H₀ and H₄ represent the assumption that economic indicators alone may not be explanatory (Buchanan, 1976; Prinz et al., 2014; Muehlbacher et al., 2011). H₂ is based on the argument that digitalization can create an impact independent of growth by both increasing deterrence capacity and strengthening voluntary compliance by reducing compliance costs (OECD, n.d.; Martínez et al., 2022). H₃ is consistent with the literature on tax administration efficiency, which emphasizes that tax administration performance may depend not only on capacity but also on how this capacity is utilized through managerial efficiency and scale efficiency (Moesen & Persoon, 2002; Anastasiou et al., 2023).

The data will be made available on reasonable request. There is no conflict of interest. This study did not receive any external funding. The study does not require ethical approval, as it does not involve any work with human or animal subjects. This article has been proofread and linguistically edited using GenAI tools to improve grammar and fluency. No part of the article, including data analysis, interpretations, or conclusions, was generated by AI tools. The author takes full responsibility for the content of this publication.

2. Theoretical Framework

To understand the magnitude and effects of this problem, the concept of tax compliance should be addressed. People

generally understand tax compliance as adhering to tax laws, which includes accurate reporting and timely payment of taxes. It can be considered tax risk management in a narrow sense or compliance with tax laws in a broad sense (Boll, 2014; Randlane, 2015; Tsaregradskaya, 2021). When tax compliance is low, public revenues decrease, the informal economy increases, and fiscal discipline deteriorates. In such a case, the provision of public goods and services deviates from its optimal level. Therefore, in this study, tax compliance is treated as the dependent variable and is represented by the collection/accrual ratio, which reflects the tax administration's success in converting assessed taxes into actual collections.

Tax compliance can be considered as a multidimensional phenomenon within the framework of established theoretical approaches in public finance and behavioral economics. In this context, firstly, deterrence theory assumes that individuals or institutions shape their behavior rationally according to cost-benefit analysis. In the tax context, it is predicted that individuals will be less likely to engage in tax avoidance or evasion if the probability of audit is high and penalties are aggravated (Kogler et al., 2022; Tabbach, 2002; Berejikian, 2002).

Second, the fiscal exchange theory argues that individuals' willingness to pay taxes to the government is directly proportional to the quality of goods and services they receive. According to this theory, a fair and transparent public expenditure system is an important factor that increases tax compliance (Buchanan, 1976).

Third, the slippery slope framework argues that trust in tax authorities and their enforcement power should be considered together. While trust increases voluntary compliance, power reinforces compulsory compliance (Prinz et al., 2014; Muehlbacher et al., 2011). In the tax compliance literature, payment compliance measures a taxpayer's psychological commitment to fulfilling their financial obligations. In the context of the slippery slope framework, the inability to collect an accrued tax reflects the taxpayer's weak belief in the system's legitimacy (trust) rather than the administration's insufficient coercive power. Therefore, the collection/accrual rate is not only an administrative success in collection, but also a behavioral proxy variable indicating the level of operation of the fiscal contract between the state and citizens (Lisi, 2019; Kirchler et al., 2008; Kogler et al., 2023).

These theories allow tax compliance to be understood not merely as an outcome of economic factors, but also as a function of governance quality, institutional trust, and perceived fairness of the tax system. In this study, tax compliance is measured on the basis of efficiency, and the effects of GDP, and population (number of taxpayers) on tax compliance behavior over time are interpreted in line with this theoretical framework.

3. Literature Review

Tax compliance literature is clustered into two main lines. The first line focuses on behavioral/psychological mechanisms that explain taxpayer compliance (deterrence, tax ethics, perception of fairness, trust, and norms). The second line addresses the problem of measuring and comparing the performance of the tax system or tax collection agencies. Concepts such as potential tax capacity, tax effort, or tax administration efficiency are evaluated using frontier approaches (DEA, SFA, Malmquist, and bootstrap derivatives). This study aims to provide a long-term effectiveness assessment for Türkiye by examining both literatures together, based on the assumption that tax compliance is shaped not only by economic factors but also by structural/institutional transformations.

3.1. Behavioral approaches to tax compliance: the balance between mandatory and voluntary compliance

While deterrence theory explains tax compliance through the likelihood of inspection and the severity of penalties, newer approaches emphasize that compliance behavior is determined not only by fear of detection but also through channels such as norms, moral motivations, perceptions of justice, and trust in institutions. In this context, the literature shows that compliance can differ even between two periods with the same audit capacity, because institutional trust and legitimacy can have a strong impact on tax compliance. This approach is consistent with studies suggesting that tax compliance is related not only to economic capacity but also to institutional quality and governance.

At this point, the Slippery Slope framework offers an important approach that bridges the literature by explaining compliance through the interaction between the power of authority and trust in authority. Voluntary compliance is associated with elements such as transparency, fair treatment, predictability, and qualified public services, while mandatory compliance is more closely linked to mechanisms such as oversight, sanctions, and third-party information reporting.

3.1.1. Audit-deterrence/enforcement and sanctions

Witte and Woodbury (1985) argue that increased audit probability, information reporting, and tax deductions increase tax compliance, while moral dilemmas and decreased audit rates decrease compliance. A model was developed, and an estimation method was used in the study.

Dubin and Wilde (1988) argue that audits conducted by the Internal Revenue Service (IRS) significantly deter noncompliance with the Federal Income Tax and support an economic approach to tax compliance. The study utilized survey and data analysis methods.

Slemrod (2019) has demonstrated that a new randomized controlled trial has shown promising results, arguing that audits, information reporting, and remittance regimes effectively reduce tax evasion. This study is a comprehensive review.

Bruttel and Friche (2014) emphasize that tax compliance depends on past tax practices. In other words, it is argued that previous tax audits and practices may influence taxpayers' current income declarations. A controlled experiment was conducted in the study.

De Neve et al. (2021) argue that simplifying communication and deterrent messages significantly increases tax compliance, while tax ethics is ineffective and therefore less cost-effective than standard enforcement measures. The study used natural field experiments as its methodology.

When these studies are evaluated together, there is a strong common finding that control and enforcement mechanisms increase tax compliance; however, the magnitude and persistence of the effect may vary depending on the methods used (field experiment, survey, observational data), the nature of the control (information reporting/third-party reporting, withholding, messaging design), and contextual conditions. Therefore, the literature shows that compliance is not a single-channel output; deterrence tools can have a high impact under some conditions but remain limited under others, and therefore the simultaneous role of institutional/technological factors must be discussed.

3.1.2. Tax ethics and voluntary compliance motivation

Kornhauser (2006) argues that tax ethics, influenced by personal values, social norms, and cognitive processes, plays a significant role in voluntary tax compliance and that the IRS can influence this through education, fair procedures, and research-based policies. The study is a comprehensive review.

Alm and Torgler (2011) argue that ethical norms play an important role in individual tax compliance decisions and that a comprehensive approach involving sanctions, services, and trust is necessary to combat tax evasion. The study attempts to prove its arguments through mathematical notation.

The tax ethics literature emphasizes that voluntary compliance cannot be explained solely by financial incentives; some studies argue that moral motivation is a strong determinant, while others suggest that certain interventions can yield more cost-effective results compared to standard enforcement tools. This differentiation suggests that tax behavior is sensitive to variables such as the country context, level of trust in institutions, and quality of administration–taxpayer interaction, and therefore that trust/governance components should not be neglected in the interpretation of macro-level results.

3.1.3. Trust and authority go hand in hand

Kirchler et al. (2008) contend that people's compliance with tax rules depends on both the power of tax authorities and trust in them, and that the interplay between these two factors plays a crucial role in shaping compliance strategies. The study is comprehensive in nature.

Gangl et al. (2015) argue that the dynamics between power and trust in tax offices can influence cooperation and that

the interaction environment can be transformed from a hostile environment to one based on service and trust. This study is a comprehensive review.

Batrancea et al. (2019) argue that trust in authorities and their power significantly increases tax compliance intentions and reduces tax evasion intentions in 44 countries with different channels for voluntary and mandatory compliance. The study used surveys and data analysis methods.

Bezerra et al. (2024) argue that tax compliance measures implemented by tax authorities can effectively encourage voluntary compliance with tax obligations, leading to an increase in taxes paid and a reduction in tax gaps. The study utilized surveys and data analysis methods.

The literature on trust–authority power shows that enforcement/power and trust are usually complementary, not alternatives. When these studies are evaluated together, it is understood that while increasing the intensity of oversight may improve short-term compliance, channels of trust such as the legitimacy of the administration, fair treatment, and service quality are critical for sustainable compliance. Therefore, differences in inter-period coordination performance should be interpreted within a framework sensitive to the evolution of the trust–power mix over time, rather than solely based on economic magnitudes.

3.1.4. Breach of tax compliance and the overall situation

Graetz and Wilde (1985) argue that tax compliance in the United States is rapidly deteriorating and that unreported income accounts for 10-15 percent of total taxable income, causing concern among critics. This study is a review.

Tilahun (2019) states that tax compliance is influenced by factors such as penalty levels, the fairness of the tax system, tax rates, and the likelihood of detection and audit, and that a responsible citizen approach is necessary for the collection of state revenues. The study is a systematic review.

Studies on breach of compliance show that tax compliance can fluctuate not only with micro-level behavioral parameters but also with factors such as macro shocks, institutional breakdowns, and policy consistency. For this reason, approaches that can measure not only the level of alignment in period-to-period comparisons but also the mechanisms by which alignment changes over time (e.g., efficiency and productivity decompositions) are becoming increasingly meaningful in the literature.

3.2. Structural/institutional determinants: financial exchange, perception of justice, service quality, and digitalization

The common emphasis of approaches that establish tax compliance within a corporate framework is that taxpayers can view tax payment not only as an obligation but also as a financial exchange in the state-citizen relationship. The efficiency of public expenditures, the visibility of public services, perceptions of corruption, accountability, and belief in the fairness of the system are discussed as channels that can either increase or decrease compliance.

Another area that has strongly contributed to this debate in recent years is digitalization and the technological transformation of tax administration. Digitalization can strengthen the deterrence channel by increasing oversight capacity and traceability. On the other hand, digitalization can also promote voluntary tax compliance by simplifying processes, reducing costs, and improving the quality of public services provided by the administration. The OECD's assessments of digital transformation in tax administrations particularly emphasize the potential of this transformation to redesign processes and reduce compliance costs (OECD, n.d.).

In this context, there are also empirical studies in the literature that directly link digitization with collection/accrual activities. For example, a study on OECD countries highlights the importance of technological and managerial transformation by showing that tax collection efficiency can be improved through reform components such as decentralization, simplification, digitization, and education (Martinez et al., 2022).

3.2.1. Fiscal exchange/public expenditure-tax compliance link

Buchanan (1976) noted that if political processes deviate from the fundamental logic of change, difficulties will arise in the system and essentially emphasized the origin of tax evasion in an indirect manner. From this perspective, tax compliance is based on the legitimacy of the constitutional contract between the state and the individual. The state's excessive fiscal exploitation is interpreted by taxpayers as a breach of contract, and this situation makes tax noncompliance a legitimate constitutional response in the eyes of the individual.

3.2.2. The relationship between the right to have a say in the use of taxes and tax compliance

Casal et al. (2016) argue that having a say in tax contributions and allocation leads to higher compliance, especially when avoiding the tax framework. The study used a controlled experimental methodology.

3.2.3. Perception of justice

Marshall et al. (2024) argue that perceptions of distributive justice are the most decisive factor influencing tax compliance, while procedural and interactive justice have less impact. The study employed a meta-analysis method.

Nartey (2023) argues that isomorphic forces and tax fairness positively influence small and medium enterprises' tax compliance, while strategic responses to institutional pressures have negligible effects. The study utilized surveys and data analysis methods.

When fiscal exchange and justice literature are evaluated together, it is seen that taxpayers view taxation not merely as an obligation but within the framework of reciprocity (service-tax balance) in the state-citizen relationship, and that under conditions where this framework is undermined,

discord may increase. This result suggests that explanations linking tax compliance solely to increases in economic capacity may be insufficient in some periods; perceived fairness, accountability, and quality of public services may be decisive factors in compliance behavior.

3.2.4. Service quality/trust in the state/administration-taxpayer relationship

Dularif and Rustiarini (2021) argue that improving tax services and trust in the state are the most important factors in increasing tax compliance; however, personal norms, social norms, and religiosity do not encourage tax compliance as strongly as expected. This study used a systematic review methodologically.

Akdoğan and Akdoğan (2023) argue that high tax compliance is associated with economic freedom, capital markets, effective competition laws, low crime rates, and trust in authorities. The study is comprehensive in nature.

Syamsudin et al. (2023) argue that tax compliance involves risk reduction, conformity, and adherence to laws, and that it applies to taxpayers, tax offices, judges, and tax advisors. A phenomenological method was used in the study.

The literature on perceptions of justice and service quality shows that distributive justice and the quality of interaction with the administration can play a systematic role in voluntary compliance. However, these effects may emerge in interaction with structural conditions such as countries' institutional capacity and level of informality. Therefore, the result that increases in economic size alone do not translate into higher collection performance (or do so in some periods) provides an explanatory framework consistent with the institutional trust–legitimacy–service quality channels emphasized in the literature.

3.3. Macro-level tax performance literature and the measurement problem

A broad body of literature evaluating tax performance at the macro level attempts to estimate the potential tax revenues that countries can collect and determine the position of actual tax revenues relative to this potential. This approach emphasizes the concepts of tax effort and tax capacity. In particular, using methods such as stochastic frontier analysis (SFA), the estimated frontier is calculated by considering the structural characteristics of countries. IMF studies show that this approach is used to assess how close countries are to their tax capacity and that the factors determining capacity include a broad set of structural variables. (Fenochietto and Pessino, 2013).

While this literature is important, it often operates from a proximity to potential perspective and may be limited in directly capturing the administrative efficiency dimension due to the lack of detailed data on the tax administration's operational processes. For this reason, a second sub-literature has focused on measuring the effectiveness of tax administrations and tax offices as if they were production units.

3.4. Comparative studies measuring tax administration/collection efficiency using DEA and Malmquist

Efforts to measure tax administration efficiency have diversified over time, despite the difficulty of accessing data. In this context, DEA and its derivatives have been found useful because they allow for the assessment of tax administration performance without reducing it to a single ratio, particularly due to their multi-input-multi-output structure.

At this point, it should be noted that DEA applications in the literature are grouped into two different levels:

- In studies at the tax office/institution level, DMUs are generally tax offices, and inputs are defined in terms of administrative resources (staff, budget, audit capacity, etc.).
- In country- or period-level studies, DMUs can be countries or time periods, and inputs are selected primarily from indicators representing macro-fiscal capacity.

This study is close to the second approach. The purpose of the study is to measure the extent to which the macro-fiscal capacity (GDP and population) available in Türkiye during specific periods can be effectively converted into collection/accrual performance. Therefore, the studies summarized below, even if they do not directly use the same input set, provide a methodological basis demonstrating that tax collection/administration performance can be measured from a frontier/efficiency perspective.

3.4.1. The studies at the tax office level (micro/corporate)

Moesen and Persoon (2002) evaluated numerous regional tax offices in Belgium using a DEA/FDH-based approach, discussing differences in efficiency between offices and the role of managerial/organizational factors.

Førsund et al. (2006) conducted a study on local tax offices in Norway and examined changes in performance over time using DEA and the Malmquist index. They also drew attention to the importance of bias correction and confidence intervals using bootstrap.

Katharaki and Tsakas (2010) studied tax offices in Greece and produced efficiency scores using DEA; they analyzed the role of environmental factors and scale size using tools such as window analysis and Tobit.

Fuentes and Lillo-Bañuls (2015) conducted a study on SUMA tax offices in Spain. In this study, efficiency change was measured using the DEA-based Malmquist index, and confidence intervals were generated using smoothed bootstrap to strengthen the distinction between technological change and efficiency change.

3.4.2. Country comparative (macro/institutional) studies

Anastasiou et al. (2023) measured the efficiency of tax administrations in European countries and revealed the

relative efficiency differences between countries using CCR-based DEA applications in this study.

Afonso et al. (2025) present a dynamic efficiency analysis from the perspective of potential tax burden/tax effort using both DEA and SFA on panel data covering the period 2000-2021, which demonstrates that frontier approaches are increasingly being used at the macro level.

Martínez et al. (2022) demonstrate that tax collection efficiency can be improved through reform components such as decentralization, simplification, digitalization, and training by measuring efficiency based on DEA in a sample of OECD countries; this result provides comparative evidence that digital transformation can be associated with increased efficiency.

When these studies are evaluated together, two points stand out:

- Performance metrics show how close the best possible outcome is achieved with available resources, rather than whether performance is high or low in a country/institution.
- It is important to determine how the efficiency changed in order to interpret the results. Therefore, Malmquist decomposition and bootstrap-based inference tools do much more than just produce scores.

3.5. The gap in the literature and the contribution of this study

The comparative literature above demonstrates that tax compliance is a multidimensional phenomenon and that performance/efficiency measurement offers a different complementary perspective. However, a significant portion of studies using frontier methods are limited to country comparisons or the tax office level. In addition, studies that relate tax compliance to a direct compliance indicator, such as the collection/assessment ratio over a very long period of time for a single country, and discuss periodic disruptions together are relatively limited.

The contribution of this study to the literature is evident in three areas:

- Providing an efficiency measurement for Türkiye that considers periodic disruptions over a long time horizon such as 1982–2021,
- Being able to distinguish productivity dynamics by tracking efficiency changes and technological changes together using Malmquist decomposition,
- Demonstrating the statistical robustness of findings by reporting the sampling sensitivity of efficiency scores using the Simar-Wilson bootstrap approach.

When examining studies conducted specifically for Türkiye, it is observed that tax efficiency is generally limited to the regional level or the impact of a specific reform on the tax audit process. This study differs from the existing literature

in that it analyzes Türkiye's forty-year neoliberalism journey from 1982 to 2021 from a long-term time series perspective and using the Malmquist productivity index, rather than focusing on a single institutional reform. Thus, the study achieves a unique academic position by separating the historical trend of Türkiye's tax compliance performance into its technological progress and managerial efficiency components.

4. Methodology and Data Set

4.1. Research Design

Many different methods have been used in the literature to evaluate the efficiency of tax compliance. However, this study combines data envelopment analysis and the Malmquist Index to measure efficiency. The reason for choosing these methods is to measure tax collection efficiency over time and to analyze efficiency changes in the tax system. The Malmquist Index is an indicator that measures changes in efficiency and technological progress over time. When used in conjunction with data envelopment analysis, it helps to determine how tax collection efficiency has changed over the years and which factors are responsible for the increase in efficiency. For example, this method can distinguish whether the increase in tax collection efficiency in a given year is due solely to better administration or to overall economic growth.

The Malmquist Index is a measure of efficiency change that captures both efficiency change and technological progress. It is usually calculated using data envelopment analysis, a method that evaluates how efficiently different units operate by comparing them to a set standard of efficiency (Kneip et al., 2020; Shojaie et al., 2024). The index can be decomposed into various components and provides a comprehensive view of efficiency dynamics. Recent developments have led to global and cyclical Malmquist efficiency indices that address issues such as different profit frontiers over time and provide a single measure of efficiency change. These indices are particularly useful for comparing productivity across different periods and units without relying on a geometric mean (Tohidi et al., 2012; Tohidi and Razavyan, 2013).

The Malmquist Index has been modified for two-stage network systems, prevalent in sectors like banking and insurance. This method considers data uncertainty by improving productivity measurements using fuzzy mathematical programming within the data envelopment analysis framework (Shojaie et al., 2024). The Malmquist Index has been applied to assess environmental productivity, such as in the construction sector in China, which shows regional differences in productivity trends. This application emphasizes the usefulness of the index in assessing sector-specific productivity changes over time (Liu et al., 2022).

Inference about efficiency change using the Malmquist Index has been challenging, particularly in dynamic environments. Recent studies have refined the index for

applied research by developing new methods and statistical procedures that improve its accuracy (Kneip et al., 2020). Innovations such as data envelopment analysis models have expanded the capabilities of the Malmquist Index. These methods enable detailed analysis of specific variables and the use of time series data, giving a clearer picture of how productivity changes over time (Asmild et al., 2016; Sanchez, 2018).

The integration of the Malmquist Index with data envelopment analysis provides a robust framework for analyzing efficiency changes in various fields. Recent methodological developments have improved the applicability and reliability of the index in dynamic and uncertain environments by addressing previous limitations. These developments continue to support informed decision-making and resource allocation in various fields. Informed decision-making requires clear efficiency assessments. This study employs the Malmquist index in conjunction with data envelopment analysis.

The main reason for choosing data envelopment analysis instead of econometric models in the study is the necessity of conducting a direct efficiency analysis to measure the efficiency of tax compliance. While econometric models usually look at how one thing affects another and the relationships between different factors, data envelopment analysis is a method that assesses how well inputs (like GDP and population) are turned into output (tax compliance).

The determinants of tax compliance could be analyzed using econometric models, but the purpose of the study is not only to see which variables affect tax compliance but also to measure whether the tax collection processes in the country are efficient. We observe the change in efficiency over time and the effect of technological progress using data envelopment analysis and the Malmquist index.

For this reason, data envelopment was chosen instead of econometric modeling because it is the only way to assess how tax compliance has changed in terms of technical and scale efficiency over the years analyzed. This method also allows one to show changes in performance between specific years and to provide more concrete recommendations for policymakers.

4.2. Data Selection

We considered macroeconomic and fiscal factors determining tax compliance when constructing the data envelopment analysis model. The following context determines the variables used in the model: GDP and population are considered inputs; tax collection efficiency is considered an output, and thus, it is tried to determine how effective tax compliance is.

One of the most widely used indicators for measuring tax compliance is the ratio of tax collection to accrued tax. This indicator reflects the extent to which accrued taxes can be converted into actual collections and is therefore also used in studies representing tax compliance/performance at the

macro level in the Turkish context (Gündem, 2022). This ratio helps measure the efficiency of tax collection by revealing the difference between the expected tax revenues of the state and the actual revenues collected. This ratio, which is used as the dependent variable (output) in the study, allows for the analysis of tax collection efficiency in Türkiye between 1982 and 2021.

Population is a demographic factor. We expect an increase in taxpayers as the population grows. However, factors such as the expansion of the informal economy may also negatively affect tax compliance. The population variable is therefore crucial to comprehending structural shifts in tax compliance. In addition, an increase in population means an increase in the number of taxpayers. This is because not only income earners but also non-income earners are taxpayers. Even if a taxpayer does not work, he/she must make expenditures to survive. This means that the taxpayer indirectly pays taxes.

Table 1 summarizes the dependent and independent variables used in the model. We consider tax compliance as the dependent variable and GDP and population as the independent variables. While GDP is expected to have a positive effect since it is an indicator of economic size, the population's effect on tax compliance can be both positive and negative.

Table 1. Dependent Variable (output) and Independent Variables (inputs)

Variable	Type	Description	Expected Impact
Tax Compliance	Dependent (output)	Tax collection performance (ratio of tax collection to accrued tax)	This is the efficiency score calculated by the DEA model based on the inputs
GDP	Independent (input)	Economic size indicator	Positive (+)
Population	Independent (input)	Demographic size may affect the number of taxpayers.	Positive (+) or negative (-)

The collection/accrual rate is frequently defined in the literature as payment compliance. This rate is not merely an indicator of administrative performance; it represents the taxpayer's final behavioral decision at the stage of actually paying the declared tax debt. A taxpayer who has passed the registration and declaration stages is more likely to pay the assessed tax, not only due to liquidity constraints but also because of their trust in the public authority and their perception that the tax paid will be used for public services (fiscal exchange). In this context, the ratio in question is considered a financial reflection of voluntary compliance (Ay & Ay, 2025; Kogler et al., 2023; Hocaoglu et al., 2023; Lisi, 2019; Ogbolu et al., 2024).

4.3. Data Sources

The data on tax compliance, GDP, and population used in the study are taken from the Turkish Statistical Institute's "Indicators of 100 Years" (TÜİK, 2023).

4.4. Selection of Data Envelopment Analysis Models

Two different data envelopment analysis models were used in this study. The first is the CCR model. CCR assumes constant returns to scale (CRS) and measures technical efficiency (CRSTE); it also calculates absolute efficiency by assuming the decision-making unit operates at optimal scale (Charnes et al., 1978). The second is the BCC model. This model assumes variable returns to scale (VRS) and measures pure technical efficiency (VRSTE); it also evaluates managerial performance without considering scale effects (Banker et al., 1984). Scale efficiency is calculated as $SCALE=CRSTE/VRSTE$.

4.4.1. The CCR Model

The model uses a fractional linear programming approach in which inputs are scaled by weights (v) and outputs scaled by weights (u). The efficiency score (Θ) is calculated as the ratio of weighted outputs to weighted inputs, with the constraint that all efficiency scores remain bounded between 0 and 1. The CCR model can be expressed as:

$$\max \frac{\sum_r u_r y_{rj}}{\sum_i v_i x_{ij}}$$

Subject to:

$$\frac{\sum_r u_r y_{rk}}{\sum_i v_i x_{ik}} \leq 1, \quad \text{for all } k$$

$$u_r, v_i \geq \epsilon > 0$$

Where;

- y_{rj} = output r for period j (tax collection performance)
- x_{ij} = input i for period j (GDP, population)
- u_r, v_i = weights for output and input

A CRSTE score of 1.000 indicates that the period operated on the efficiency frontier with optimal scale. Values below 1.000 indicate inefficiency-resources were not optimally utilized given their scale. This model reveals absolute efficiency, combining both pure technical efficiency and scale efficiency into a single measure.

4.4.2. The BCC Model

The BCC model adds a convexity constraint to the CCR model, allowing the efficiency frontier to be piecewise linear rather than a single ray from the origin. This permits the model to account for periods or units that may operate at increasing, decreasing, or constant returns to scale. The BCC model in output-oriented form can be expressed as:

$$\max \frac{\sum_r u_r y_{rj} - u_0}{\sum_i v_i x_{ij}}$$

Subject to:

$$\frac{\sum_r u_r y_{rk} - u_0}{\sum_i v_i x_{ik}} \leq 1, \quad \text{for all } k$$

$$u_r, v_i \geq \varepsilon > 0$$

Where:

- u_0 = a variable reflecting convexity (not present in CCR)
- All other variables are as previously defined

A VRSTE score of 1.000 indicates pure technical efficiency—the period was operationally efficient regardless of its scale. The BCC model filters out scale inefficiency, allowing researchers to determine whether poor performance stems from managerial inefficiency or from operating at a non-optimal scale.

4.4.3. Scale Efficiency and Decomposition

Scale efficiency (SCALE) is derived as the ratio of the two models:

$$SCALE = (CRSTE)/(VRSTE)$$

This decomposition provides critical insights:

- If $SCALE < 1$, the inefficiency is attributable to scale issues (not operating at optimal size)
- If $SCALE \approx 1$, scale is not a constraint, inefficiency stems from pure technical performance

4.4.4. Output-Oriented Specification

Both models are implemented in output-oriented form in this study. This orientation is theoretically justified. This is because the research question addresses how efficiently given inputs (economic and demographic capacity) are transformed into outputs (tax collection performance). An output-oriented approach answers: “Given current resources, how much can tax collection performance be improved?” This aligns with the policy objective maximizing collection efficiency without proportionally increasing administrative burden.

4.5. Analysis Period

The data covers the period 1982–2021. For ease of analysis and to enable periodic comparisons, the data has been divided into four 10-year periods. The first period covers the years 1982–1991, the second period covers the years 1992–2001, the third period covers the years 2002–2011, and the fourth period covers the years 2012–2021. Inputs and outputs for each period were calculated by taking the average for that period.

4.6. Malmquist Index

The Malmquist Total Factor Productivity (TFP) Index was calculated to analyze the change in efficiency of decision-making units over time and technological developments. The Malmquist index is broken down into the following components:

- **EFFCH (Efficiency Change):** Change in the relative efficiency position of the decision-making unit,
- **TECHCH (Technical Change):** The movement of the production frontier over time,
- **PECH (Pure Efficiency Change):** Efficiency change based on the VRS model,
- **SECH (Scale Efficiency Change):** Changes in the scale structure,
- **TFPCH (Total Factor Productivity Change):** $EFFCH \times TECHCH$.

If the TFPCH value is greater than 1, there is an increase in productivity; if it is less than 1, there is a decrease in productivity.

The Malmquist Total Factor Productivity (TFP) index is calculated using output-oriented distance functions, defined by Caves et al. (1982) and framed within a data envelopment analysis-based framework by Färe et al. (1994). The index measuring the change in productivity between periods t and $t+1$ is expressed under the CRS assumption using the following geometric mean formula:

$$M_o(x^{t+1}, y^{t+1}, x^t, y^t) = \frac{[d_o^t(x^{t+1}, y^{t+1})/d_o^t(x^t, y^t)] \times [d_o^{t+1}(x^{t+1}, y^{t+1})/d_o^{t+1}(x^t, y^t)]^{1/2}}$$

Here, $d_o^t(x^{t+1}, y^{t+1})$ represents the distance of the data in period $t+1$ from the technology of period t . In the process of index decomposition, technical efficiency change (EFFCH) was calculated relative to the CRS frontier, while pure efficiency change (PECH) was obtained using the BCC model developed by Banker et al. (1984) over the Variable Returns to Scale (VRS) frontier. Scale efficiency change (SECH) is included in the model as the ratio of these two values ($SECH = EFFCH / PECH$). This technical distinction definitively reveals whether the increase in productivity stems from a technological leap (TECHCH) or a managerial improvement (PECH).

The temporal selection of benchmarking years in Table 4 is not arbitrary; it aligns with the inaugural year of each institutional sub-period defined in this study. For instance, the comparison between 1982 and 1992 captures the efficiency shift from the initial post—1980 stabilization phase to the era of early market liberalization. Similarly, the 2002–1992 comparison evaluates the transformative effect of the 2001 economic crisis and the subsequent institutional overhaul of the tax administration. By selecting the starting gates of each decade, it is possible to measure the decade-

long structural momentum that defines Türkiye's neoliberal tax evolution. Furthermore, to address the period-wide consistency, Table 5 presents the Malmquist decomposition based on decadal averages, ensuring that the results are robust against annual fluctuations and accurately reflect the structural transitions of the tax compliance frontier.

4.7. Bootstrap Analysis

Due to the small sample size, the Simar & Wilson (1998) bootstrap method was applied to correct for bias in data envelopment analysis efficiency scores. The analysis, conducted with 2000 bootstrap iterations, resulted in the calculation of original activity scores, bootstrap means and standard deviations, 95% confidence intervals, and bias-corrected activity scores.

4.8. Determining Returns to Scale

Using the CCR and BCC models, scale returns were determined for each decision-making unit:

- IRS (Increasing Returns to Scale): $CRSTE < VRSTE$
- CRS (Constant Returns to Scale): $CRSTE = VRSTE$
- DRS (Decreasing Returns to Scale): $CRSTE > VRSTE$

4.9. Software

Data envelopment analysis and Malmquist index calculations were performed using the Python 3.13.7 programming language. The linear programming problem was solved using the `scipy.optimize.linprog` library.

Python was selected for this study due to its transparency in computational procedures, reproducibility of results, and flexibility in implementing custom bootstrap procedures (Simar-Wilson methodology) that are not readily available in proprietary software packages such as DEAP or MaxDEA.

The `scipy.optimize.linprog` solver was configured with the following specifications:

- **Algorithm:** Interior-point method (default dual-feasible initialization)
- **Tolerance:** $1e-9$ (machine precision level for optimality conditions)
- **Presolve:** Enabled to reduce problem dimensionality
- **Iteration Limit:** 1000 iterations (sufficient convergence for all DMU evaluations)

4.9.1. Key Methodological Assumptions

The assumptions are as follows:

- **Technology assumptions:** Convexity of the production possibility set, weak disposability of inputs and outputs, no joint production, homogeneity of technology across periods.

- **Data assumptions:** Non-negativity of all inputs and outputs, no zero values present in the dataset, measurement accuracy.
- **Model-specific assumptions:** CCR (constant returns to scale, ray-based frontier), BCC (variable returns to scale, piecewise linear frontier).
- **Orientation assumption:** Output-oriented specification chosen to maximize tax collection performance given available administrative and demographic capacity.
- **Linear programming assumptions:** Linearity of the efficiency frontier, separability of input-output relationships, uniqueness of optimal weights.
- **Bootstrap assumptions:** Independence of bootstrap resamples, finite second moments of efficiency estimator, convergence with $B=2000$ iterations.

4.10. Limitations

This study measures the efficiency of tax compliance in Türkiye between 1982 and 2021 using the Malmquist index and data envelopment analysis. There are certain limitations due to the methodology and data set used.

Firstly, the data envelopment analysis method cannot identify direct causal relationships. This analysis focuses only on the efficiency of tax collection and does not show which factors are directly effective for tax compliance. Alternatively, supplementing it with panel data regressions or causality analyses may increase the power of the method.

Within the first limitation, it is also necessary to mention the risk of endogeneity and the interpretation boundary. The macro variables used in this study and tax collection performance can be determined to be in mutual interaction within the same period. For instance, improvements in collection performance can affect the public revenue structure and the level of economic activity through indirect channels. Similarly, growth and increases in public revenues can also strengthen collection capacity. For these reasons, the results should be interpreted not as inputs causally determining outputs, but rather in terms of the relative level of efficiency in converting a given macro-fiscal capacity into tax collection performance and in terms of cyclical patterns. This study does not estimate causal coefficients due to the nature of the DEA/Malmquist approach; it provides evidence for policy discussions by reporting efficiency scores and their changes over time.

Secondly, the study utilizes the total tax collection rate for Türkiye as a whole. However, performing a regional or sectoral analysis is impracticable without data specific to each province. If province-based or sector-based tax collection data were available, it would be possible to analyze how tax compliance varies in different regions.

Thirdly, the informal economy cannot be measured directly. Despite informal economic activities constituting a

substantial portion of Türkiye's economy, the absence of direct data necessitates the analysis of the informal sector via indirect indicators in this study. Future research should employ approaches that directly assess the informal economy.

Fourthly, this study does not take regional differences into account when analyzing tax compliance efficiency. This limitation represents a constraint of the study, as variations in economic structure, the size of the informal economy, and administrative capacity across regions may affect tax compliance. Future research may undertake a more comprehensive analysis of the influence of these disparities on tax compliance via regionally focused studies.

Finally, the data set used in the study is limited to the years 1982-2021. Upon the availability of more recent data, the analyses must be re-evaluated, and the long-term impacts of tax reforms should be investigated more thoroughly.

5. Results

5.1. Descriptive Statistics

Table 2 shows the descriptive statistics for the four periods analyzed. The data shows a consistent upward trend in both GDP and population across all periods. Real GDP (billion TL, 2021 prices) increased by 281.5%, from an average of 3,833.6 in Period 1 (1982-1991) to 14,625.2 in Period 4 (2012-2021). Similarly, the population also increased from 52.1 million to 80.3 million during the same period. However, tax collection performance exhibits a non-linear trend. After peaking at 88.9% in Period 3 (2002-2011), it showed a sharp decline to 80.0% in Period 4 (2012-2021), despite significant increases in both GDP and population.

Table 2. Descriptive Statistics by Period

Period	Years	GDP (Billion TL)	Population (million)	Tax Collection Performance (%)
1	1982- 1991	3,833.6	52.1	86.8
2	1992- 2001	5,808.8	62.4	86.1
3	2002- 2011	8,753.0	70.1	88.9
4	2012- 2021	14,625.2	80.3	80.0

Notes: GDP figures are in billion TL (2021 constant prices). Population figures are in millions. Tax collection performance represents the ratio of tax revenue to tax accrual.

5.2. Efficiency Analysis

Table 3 presents DEA efficiency scores under both CCR (constant returns to scale) and BCC (variable returns to scale) assumptions. The results reveal significant differences in tax compliance efficiency across the four

periods.

In the first period (1982-1991), CRSTE = 1.000, indicating full efficiency, which signifies optimal performance given the current inputs. During this period, increasing returns to scale (IRS) are observed, indicating that the tax administration can benefit from expansion. The second period (1992-2001) shows a decrease in productivity with CRSTE = 0.828 and VRSTE = 0.834. A scale efficiency of 0.992 indicates that the inefficiency during this period was primarily due to pure technical inefficiency rather than scale issues. The third period (2002-2011) presents an intriguing pattern: while VRSTE = 1.000 (fully efficient under VRS), CRSTE drops to 0.760. This inconsistency leads to a scale efficiency of 0.760 across all periods, which is the lowest with the IRS present. This suggests that despite achieving technical efficiency, the operation was at a suboptimal scale for the period.

The fourth period (2012-2021) shows the lowest levels of efficiency across all criteria: CRSTE = 0.598 and VRSTE = 0.649. The scale efficiency of 0.922 indicates that both pure technical inefficiency and scale inefficiency contributed to the overall poor performance, although technical inefficiency was the primary driver.

Table 3. DEA Efficiency Scores by Period

Period	Years	CRSTE	VRSTE	SCALE	RTS
1	1982- 1991	1	1	1	IRS
2	1992- 2001	0,828	0,834	0,992	-
3	2002- 2011	0,76	1	0,76	IRS
4	2012- 2021	0,598	0,649	0,922	-

Notes: CRSTE=Technical efficiency under constant returns to scale (CCR model). VRSTE=Technical efficiency under variable returns to scale (BCC model). SCALE=Scale efficiency (CRSTE/VRSTE). RTS>Returns to Scale. IRS= Increasing returns to scale.

5.3. Malmquist Productivity Index: Yearly Comparisons

Table 4 shows the Malmquist Productivity Index and its components for year-over-year comparisons across different time horizons. The results show a consistent pattern of productivity growth, indicating that this increase is entirely due to technological change (TECHCH > 1.000). On the other hand, the efficiency change (EFFCH) remains constant at 1.000 in all comparisons. The change in total factor productivity (TFPCH) ranged from 1.003 (1992 → 2002) to 1.136 (2005 → 2015), indicating productivity increases between 0.3% and 13.6%. The most significant productivity increase is seen in the 2005 → 2015 comparison (TFPCH = 1.136). This also reflects the impact of digitalization and technological infrastructure investments in the tax administration. The analysis reveals that all productivity gains stem from

technological advancements rather than productivity improvements. Pure efficiency change (PECH) and scale efficiency change (SECH) are equal to 1.000 in each period, indicating no catch-up effect. This indicates that while the production frontier shifts outward due to technological advancements, DMUs have not improved their relative efficiency positions.

Period-to-period comparisons (1982 → 1992, 1992 → 2002, 2002 → 2012, 2012 → 2021) show TFPCH values of 1.098, 1.003, 1.114, and 1.105, respectively. The lowest productivity increase occurred between 1992 and 2002 (TFPCH = 1.003), coinciding with the economic instability in Türkiye during the late 1990s and early 2000s.

Table 4. Malmquist Productivity Index-Yearly Comparisons

Period	Comparison	EFFCH	TECHCH	PECH	SECH	TFPCH
2	1982-1992	1.000	1.098	1.000	1.000	1.098
3	1992-2002	1.000	1.003	1.000	1.000	1.003
4	2002-2012	1.000	1.114	1.000	1.000	1.114
5	2012-2021	1.000	1.105	1.000	1.000	1.105
6	1990-2000	1.000	1.065	1.000	1.000	1.065
7	1995-2005	1.000	1.013	1.000	1.000	1.013
8	2000-2010	1.000	1.079	1.000	1.000	1.079
9	2005-2015	1.000	1.136	1.000	1.000	1.136
10	2011-2021	1.000	1.105	1.000	1.000	1.105

Notes: EFFCH = Efficiency change. TECHCH = Technological change. PECH = Pure efficiency change. SECH = Scale efficiency change. TFPCH = Total factor productivity change (EFFCH × TECHCH). Values > 1 indicate improvement, < 1 indicate deterioration.

5.4. Malmquist Productivity Index: Period Comparisons

Table 5 presents the Malmquist decomposition for the transition between three periods. The results complement the annual analysis by revealing the medium-term productivity dynamics over four 10-year periods. The transition from Period 1 to Period 2 (1982-1991 → 1992-2001) indicates a TFPCH of 1.099, representing a 9.9% increase in productivity solely due to technological advancements (TECHCH = 1.099). This period coincides with Türkiye's gradual market liberalization and initial reforms in tax administration.

technological progress, not to efficiency gains in existing operations.

Table 5. Malmquist Productivity Index-Period Comparisons

Period	Comparison	EFFCH	TECHCH	PECH	SECH	TFPCH
1	1982-1991 → 1992-2001	1.000	1.099	1.000	1.000	1.099
2	1992-2001 → 2002-2011	1.000	1.043	1.000	1.000	1.043
3	2002-2011 → 2012-2021	1.000	1.128	1.000	1.000	1.128

The transition from Period 2 to Period 3 (1992-2001 → 2002-2011) shows the lowest productivity growth among the transitions between the three periods (TFPCH = 1.043 or 4.3%). This modest improvement reflects the economic turmoil of the late 1990s and early 2000s, including the 1998-2001 financial crisis and its aftermath. Despite challenging economic conditions, technological progress (TECHCH= 1.043) continued at a moderate pace. The transition from the 3rd period to the 4th period (2002-2011 → 2012-2021) recorded the highest productivity increase (TFPCH = 1.128 or 12.8%), reflecting the impact of advanced digitalization initiatives, e-government services, and improved adaptation systems introduced during this period. This strong increase in productivity sharply contrasts with the decline in productivity observed in the DEA analysis for the 4th quarter. This also indicates that while the production frontier has advanced significantly, the real efficiency performance of the tax system has deteriorated relative to this frontier. Throughout the entire transition between the three periods, the change in productivity (EFFCH = 1.000) and its components (PECH and SECH) remain constant, supporting the finding that productivity gains can be attributed solely to

Notes: EFFCH = Efficiency change. TECHCH = Technological change. PECH = Pure efficiency change. SECH = Scale efficiency change. TFPCH = Total factor productivity change (EFFCH × TECHCH). Values > 1 indicate improvement, < 1 indicate deterioration.

5.5. Bootstrap Analysis and Robustness Checks

To assess the statistical reliability of DEA efficiency scores, we use the bootstrap analysis applied by Simar & Wilson (1998) with 2,000 repetitions. Table 6 reports the original efficiency scores, bootstrap means, standard deviations, and 95% confidence intervals for both CRSTE and VRSTE measurements. The bootstrap results reveal important statistical properties of the efficiency estimates. For the first period (1982-1991), the original CRSTE and VRSTE scores of 1,000 each yielded bootstrap means of 0.850 and 0.904,

respectively, with standard deviations of 0.144 and 0.130. The 95% confidence intervals range from [0.598, 1.000] for CRSTE and [0.649, 1.000] for VRSTE, indicating significant sampling variability. Period 2 (1992-2001) shows original scores of CRSTE = 0.828 and VRSTE = 0.834, while the bootstrap means are 0.849 and 0.906, respectively. The narrow gap between the original and bootstrap mean values (differences of 0.021 and 0.072) indicates a moderate level of bias in the original estimates. Confidence intervals reflect those from the first period, mirroring similar distribution characteristics. The third period (2002-2011) presents an interesting model: While the original CRSTE = 0.761 is close to the lower bound of the bootstrap confidence interval, VRSTE = 1.000 exhibits a bootstrap mean of 0.906. This difference suggests that a fully efficient VRSTE score may overestimate the period's actual productivity performance. The 4th period

(2012-2021) records the lowest original efficiency scores (CRSTE = 0.598, VRSTE = 0.649), which exactly corresponds to the lower bounds of the respective confidence intervals. The bootstrap averages (0.851 and 0.907) significantly exceed the original estimates, indicating a substantial upward bias correction. However, the position of the original scores at the lower bound of the confidence interval confirms that the poor performance in Period 4 is statistically robust and not a result of sampling variation. The bootstrap standard deviations remain remarkably consistent across periods (ranging from 0.144 to 0.146 for CRSTE and from 0.130 to 0.131 for VRSTE), indicating stable prediction accuracy. The uniform confidence intervals across all periods reflect the small sample size ($n=4$ periods) in the bootstrap procedure and the resulting limited degrees of freedom.

Table 6. Bootstrap Analysis Results (B=2,000 replications)

Period	Years	CRSTE _orig	CRSTE _mean	CRSTE _std	CRSTE _CI	VRSTE _orig	VRSTE _mean	VRSTE _std	VRSTE _CI
1	1982-1991	1.000	0.850	0.144	[0.598,1.000]	1.000	0.904	0.130	[0.649,1.000]
2	1992-2001	0.828	0.849	0.145	[0.598,1.000]	0.834	0.906	0.131	[0.649,1.000]
3	2002-2011	0.761	0.853	0.146	[0.598,1.000]	1.000	0.906	0.131	[0.649,1.000]
4	2012-2021	0.598	0.851	0.146	[0.598,1.000]	0.649	0.907	0.131	[0.649,1.000]

Notes: CRSTE_orig and VRSTE_orig = Original DEA efficiency scores. CRSTE_mean and VRSTE_mean = Bootstrap sample means. CRSTE_std and VRSTE_std = Bootstrap standard deviations. CRSTE_CI and VRSTE_CI = 95% bias-corrected confidence intervals. Bootstrap procedure follows Simar & Wilson (1998).

5.6. Bias-Corrected Efficiency Scores

Table 7 presents the bias-corrected efficiency estimates obtained from the bootstrap procedure. This bias-corrected score (BC) provides more reliable point estimates by subtracting the estimated bias from the original DEA scores, offering a conservative assessment of actual efficiency levels. For the first period (1982-1991), there is a large negative bias (CRSTE_bias = -0.150, VRSTE_bias = -0.096), indicating a significant downward bias in the bootstrap mean estimates. The bias-corrected CRSTE and VRSTE scores remain at 1.000, which confirms the full efficiency of the period despite the sampling variation introduced by the bootstrapping procedure. The second period (1992-2001) shows a modest positive bias (0.022) at CRSTE and a larger positive bias (0.071) at VRSTE. While the bias-corrected scores decreased to CRSTE_BC = 0.806 and VRSTE_BC = 0.763, this indicates that the original scores slightly overestimated the efficiency levels. Corrected scale efficiency remains close to unity (SCALE_BC = 1.000), indicating that scale effects do not explain the lack of efficiency.

The third period (2002-2011) exhibits the most striking bias pattern. Although the original VRSTE = 1.000, the bootstrap reveals a negative bias of -0.094, resulting in a bias-corrected VRSTE_BC of 1.000 (limited by the boundary). More importantly, the original CRSTE = 0.761 contains a significantly positive bias: 0.092. Bias-corrected

CRSTE_BC = 0.668, which is significantly lower than the original estimate, indicating that the period's productivity was significantly overestimated. The correction of the scale efficiency bias from 0.761 to 0.580 further emphasizes the severity of scale inefficiency during this period. The 4th period (2012-2021) shows the largest positive biases across all criteria: CRSTE_bias = 0.254 and VRSTE_bias = 0.258. The bias-corrected scores dropped dramatically to CRSTE_BC = 0.344 and VRSTE_BC = 0.390, which are significantly lower than their original estimates of 0.598 and 0.649, respectively. These significant downward revisions indicate that the weak performance in the fourth quarter was not a measurement error but rather reflected genuine operational shortcomings. The bias-corrected scale efficiency (SCALE_BC = 0.905) indicates that although scale effects contribute to inefficiency, they are secondary to pure technical inefficiency.

The pattern of bias observed over time reveals that the small sample bias in the DEA is most pronounced in the extreme periods (periods 1 and 4), where the observations are at or near the efficiency frontier. The positive biases in periods 3 and 4 indicate that the original DEA scores tend to overestimate efficiency, especially when actual efficiency is lower. This finding highlights the importance of bootstrap bias correction for robust inferences in DEA studies with limited observations.

Table 7. Bias-Corrected Efficiency Scores

Period	CRSTE _orig	CRSTE _bias	CRSTE _BC	VRSTE _orig	VRSTE _bias	VRSTE _BC	SCALE _orig	SCALE _bias	SCALE _BC
1	1.000	-0.150	1.000	1.000	-0.096	1.000	1.000	-0.060	1.000
2	0.828	0.022	0.806	0.834	0.071	0.763	0.992	-0.054	1.000
3	0.761	0.092	0.668	1.000	-0.094	1.000	0.761	0.181	0.580
4	0.598	0.254	0.344	0.649	0.258	0.390	0.922	0.017	0.905

Notes: CRSTE_orig, VRSTE_orig, SCALE_orig = Original DEA efficiency scores. CRSTE_bias, VRSTE_bias, SCALE_bias = Bootstrap-estimated biases. CRSTE_BC, VRSTE_BC, SCALE_BC = Bias-corrected efficiency scores (original score minus bias). Negative bias indicates the original score underestimated efficiency; positive bias indicates overestimation.

5.7. Hypothesis Testing Results

The empirical analysis provides clear evidence regarding each hypothesis. Firstly, H_0 is accepted. Despite GDP growth and population increase, tax efficiency declined in Period 4. Economic size alone does not determine tax compliance. Secondly, H_1 is rejected. Although GDP and population reached their maximum average levels in period 4, the tax efficiency scores and bias-corrected scores plummeted to their historical minimum. This negative correlation provides strong empirical evidence that macro-fiscal capacity indicators do not directly or linearly translate into tax compliance performance. Thirdly, H_2 is partially accepted. While TECHCH=1.128, operational efficiency (CRSTE=0.598) simultaneously declined, indicating technology alone is insufficient without institutional capacity. Fourthly, H_3 is strongly accepted. Scale efficiency deteriorated to 0.580 (bootstrap-corrected), showing organizational structure is more critical than economic size. Lastly, H_4 is partially accepted. While Malmquist index components (TECHCH>1) and theoretical discussions link compliance to structural factors like digitalization and institutional trust, the lack of province-based or sector-specific data limits our ability to empirically map these structural influences across different regional administrative capacities within Türkiye.

6. Discussion

According to the results of this study, the effectiveness of tax compliance in Türkiye is determined more by structural and institutional factors than by economic size. The analysis, covering the years 1982-2021 using non-parametric methods, indicates significant fluctuations in period-based efficiency scores. These fluctuations demonstrate that not only economic indicators such as GDP and population but also structural factors such as digitalization, institutional trust, and governance quality must be considered.

The deterrence theory, fiscal exchange theory, and slippery slope framework, which are at the core of the study, are useful for interpreting the results. Deterrence theory predicts that control and punishment mechanisms will improve tax compliance. The hypothesis in question appears to be partially supported by the findings obtained in the study.

Particularly during the 2012-2021 period, investments in digitalization and improvements in control mechanisms led to a significant increase in technological progress (TECHCH=1.128). However, the fact that efficiency scores decreased during the same period (CRSTE=0.598) shows that deterrence mechanisms alone are not sufficient.

The central argument of the theory of fiscal exchange is that the quality of public services received affects the willingness to pay taxes. The hypothesis in question is consistent with the analysis results obtained. For example, the decline in efficiency despite the reforms implemented between 2002 and 2011 shows that tax policies during this period were important not only for audit mechanisms but also for shaping perceptions of the quality of public services and the budget structure. The slippery slope framework, by emphasizing the interaction of power and trust, provides

The analysis conducted over four periods clearly shows that tax compliance effectiveness in Türkiye has not followed a uniform development. While the first period (1982-1991) provided full efficiency (CRSTE=1.000), this value should be approached with caution. Considering the low tax base and relatively simple tax structure during this period, the mentioned full efficiency may be relative. Bootstrap analysis revealed that the efficiency scores calculated for this period have a significant negative bias (CRSTE=-0.150), which proves that the actual efficiency is not as high as stated.

A significant decline in activity is observed in the second period (1992-2001) (CRSTE=0.828). This period covers the financial crises, inflation problems, and uncertain environment that Türkiye experienced at the end of the 1990s. Political instability and macroeconomic uncertainty may have prevented the tax administration from fully utilizing its capacity. The fact that the Malmquist index showed a low productivity increase for this period (TFPCH=1.003) indicates that the economic turmoil of the period also had a negative impact on technological progress.

The third period (2002-2011) reflects an intriguing paradox. While a decrease is observed at constant returns to scale (CRSTE=0.760), full efficiency is observed under variable returns to scale (VRSTE=1.000). This inconsistency indicates that despite significant reforms during the period, there were serious problems with scale efficiency. The

corrected scale efficiency score (SCALE_BC=0.580) from the bootstrap analysis confirms the presence of this issue.

The lowest efficiency scores are observed in the most recent period (2012-2021) (CRSTE=0.598, VRSTE=0.649). Although the highest technological progress for this period was recorded by the Malmquist index (TECHCH=1.128), the fact that the efficiency scores remained so low presents a significant inconsistency. This indicates that technological advancements cannot compensate for current operational deficiencies.

Decomposing the Malmquist index makes the conclusion that technological change (TECHCH) is positive when efficiency change (EFFCH) is zero even more apparent. This result indicates that the productivity gains experienced by the tax system in Türkiye are solely due to external technological advancements, while the tax administration's capacity to effectively utilize these technologies has remained constant. In other words, the developed digital tools and systems have not been used to their full capacity.

In tax administration, the concept of technology encompasses not only hardware infrastructure but also a broader information management ecosystem that includes the digitalization of tax processes, e-filing, e-invoicing, e-bookkeeping, and risk-focused data mining algorithms (Bassey et al., 2022). The analysis results, with TECHCH values exceeding 1, demonstrate that Türkiye's e-transformation initiatives over the past twenty years have raised the tax collection frontier. However, the decline of CRSTE, which was unable to keep pace with this progress, reflects a situation that can be described as the digital transformation paradox. Despite the development of technological infrastructure, the inability of managerial factors such as institutional capacity, staff expertise, and taxpayer compliance costs to keep pace with this speed has prevented the full potential of productivity gains from being realized. This result emphasizes that technology alone is not sufficient; it must be supported by institutional trust and governance quality.

These results indicate that tax administration needs serious improvements in organizational capacity, staff training, decision-making processes, and administrative structure. From the perspective of fiscal exchange theory, technological progress should be perceived by citizens as a better service and a fairer system in terms of taxation. However, deficiencies in institutional capacity may have hindered this perception.

The bias corrections calculated from the bootstrap analysis reveal a hidden truth for the fourth period. While the original efficiency score was CRSTE=0.598, after bias correction, they decreased to CRSTE_BC=0.344. This significant drop indicates that the true efficiency score is considerably lower. In other words, the operational failure of the tax administration during the 2012-2021 period was hidden in the initial calculations, and the bootstrap correction revealed

this fact. Consequently, despite investments in digital transformation, real operational efficiency has remained extremely low.

The rejection of H_1 and the partial confirmation of H_4 are intrinsically linked to the digitalization paradox observed in the fourth period. As deterrence theory predicts, technological power increases the ability to monitor, but fiscal exchange theory and the slippery slope framework remind us that without institutional legitimacy, this power leads to decreasing returns. The empirical fact that the most developed period in terms of GDP is the least efficient in terms of collection performance validates the necessity of looking beyond economic aggregates to structural and behavioral variables.

7. Conclusion & Recommendations

The results definitively show that the efficiency of tax compliance in Türkiye cannot be explained solely by economic indicators. Growth in economic size and population does not guarantee an automatic improvement in tax collection efficiency. Instead, institutional trust, the quality of governance, the organizational capacity of the tax administration, and the budget structure play a critical role in shaping the quality of services received by the public and thus in determining tax compliance. In other words, deterrence mechanisms and technological advancements do not lead to increased long-term tax compliance without being grounded in institutional legitimacy and public trust.

To resolve the technological progress-operational efficiency decline paradox, which is the main finding of the study, the following points need to be realized:

- **Human resource development:** The training of tax inspectors and accounting personnel should focus not only on technical knowledge but also on analytical decision-making skills. Training programs should be updated periodically and made mandatory to increase the efficiency of digital systems.
- **Restructuring the tax administration:** The significant decline in scale efficiency observed between 2002 and 2011 indicates that, despite the growth of the tax administration, coordination problems were experienced. Therefore, there is a need for the rapid improvement of central data systems that facilitate information sharing between regional and local tax offices.
- **Rationalizing decision-making systems:** Reducing implementation differences between tax offices and developing standardized procedures nationwide will increase scale efficiency.

Strengthening the relationship between paying taxes for

services, as predicted by the theory of fiscal exchange, is fundamental to the long-term increase in tax compliance. For this purpose, the following matters need to be realized:

- **Increasing budget transparency:** The allocation of tax revenues to public services should be communicated to citizens in a clearer and more accessible manner. Increasing citizen participation in the decision-making processes of central and local budgets will enhance the legitimacy of expenditures.
- **Tax-service link at the local level:** Local taxes should be directly linked to local services, and this link should be made visible to citizens. Doing so will demonstrate the tangible benefits of paying taxes.
- **Performance reporting:** Annual performance reports by the tax administration should show not only tax collection figures but also the efficiency of the resources spent.

However, in order for technological advancements to translate into operational efficiency, the following points need to be implemented:

- **Integrated tax information system:** A central system should be established to enable data integration across different tax types. Such a system will increase audit efficiency by preventing the same taxpayer information from being re-recorded in multiple systems.
- **Real-time monitoring mechanisms:** Early warning systems should be developed for taxpayers at risk of tax evasion using big data analytics and artificial intelligence techniques. Such systems will enable for a more efficient allocation of audit resources.
- **Taxpayer information platforms:** Mobile and web-based applications should be widely adopted, allowing taxpayers to track their tax demands and transactions in real time. Such platforms will increase trust in the tax administration.

As seen in the period 1992-2001, financial crises seriously negatively impact the efficiency of tax collection. To mitigate this risk, the following points need to be addressed:

- **Counter-cyclical tax policies:** Adjusting tax collection targets during economic crises according to macroeconomic conditions should not overstrain the tax administration's capacity.
- **Payment planning flexibility:** Offering taxpayers long-term payment plans and relaxing the penalty system during economic shocks will increase tax compliance.

- **Incentives for transitioning to the formal economy:** Tax exemptions and deductions that encourage the shift from the informal to the formal economy during periods of financial instability will broaden the tax base.

The balance of trust and power, which is the central projection of the slippery slope framework, needs to be improved in the following ways:

- **Accountability of the tax administration:** An independent appeal mechanism should be established against unnecessary interventions by tax inspectors, and transparency regarding tax refunds should be increased.
- **Strengthening taxpayer rights:** The independence of courts against the legal decisions of the tax administration should be strengthened, and taxpayers' access to courts should be facilitated.
- **Mechanisms for social dialogue:** Regular dialogue between the tax administration and business organizations, accounting chambers, and academic representatives will increase the legitimacy of tax policies.

As can be seen, the proposed reforms focus not only on economic factors but also on institutional, structural, and social factors. Technological investments will yield more effective results if trust and legitimacy, which are the foundation of these reforms, are established. This is because ensuring long-term tax compliance requires not only deterrence but also the systematic implementation of rewards and trust.

References

- Afonso, A., Montes, A. P., & Domínguez, J. M. (2025). A dynamic efficiency analysis for tax revenues in OECD countries. *Empirica*, 52, 377–411.
- Akdoğan, T., & Akdoğan, İ. (2023). Theoretical approach to tax compliance. *PressAcademia*, 16, 44–49.
- Alm, J., & Torgler, B. (2011). Do ethics matter? Tax compliance and morality. *Journal of Business Ethics*, 101, 635–651.
- Anastasiou, A., Kalligosfyris, C., & Kalamara, E. (2023). Measuring tax administrations efficiency using data envelopment analysis: Evidence from 26 European countries. *International Journal of Computational Economics and Econometrics*, 13(1), 61–109.
- Asmild, M., Balezentis, T., & Hougaard, J. L. (2016). Multi-directional productivity change: MEA-Malmquist. *Journal of Productivity Analysis*, 46, 109–119.
- Ay, H. M., & Ay, N. G. (2025). Evaluation of tax policies in terms theory of rational expectations and Türkiye. *Kronika Journal*, 25(6), 821–836.
- Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data development analysis. *Management Science*, 30(9), 1078–1092.
- Bassey, E., Mulligan, E., & Ojo, A. (2022). A conceptual framework for digital tax administration-a systematic review. *Government Information Quarterly*, 39(4), Article 101754.
- Batrancea, L., Nichita, A., Olsen, J., Kogler, C., Kirchler, E., Hoelzl, E., Weiss, A., Torgler, B., Fooker, J., Fuller, J., Schaffner, M., Banuri, S., Hassanein, M., Alarcon-Garcia, G., Aldemir, C., Apostol, O., Weinberg, D. B., Batrancea, I., Belianin, A., . . . Zukauskas, S. (2019). Trust and power as determinants of tax compliance across 44 nations. *Journal of Economic Psychology*, 74, Article 102191.
- Berejikan, J. D. (2002). A cognitive theory of deterrence. *Journal of Peace Research*, 39(2), 165–183.
- Bezerra, T. V. S., Oliveira, V. O., & De Souza, L. L. F. (2024). Tax compliance actions by tax authorities and voluntary compliance with tax obligations by individual taxpayers. *Advances in Scientific and Applied Accounting*, 17(1), 55–68.
- Boll, K. (2014). Mapping tax compliance: Assemblages, distributed action and practices: A new way of doing tax research. *Critical Perspectives on Accounting*, 25(4-5), 293–303.
- Bruttel, L., & Friche, T. (2014). On the path dependence of tax compliance. *European Economic Review*, 65, 90–107.
- Buchanan, J. M. (1976). Taxation in fiscal exchange. *Journal of Public Economics*, 6(1-2), 17–29.
- Casal, S., Kogler, C., Mittone, L., & Kirchler, E. (2016). Tax compliance depends on voice of taxpayers. *Journal of Economic Psychology*, 56, 141–150.
- Caves, W. D., Christensen, L. R., & Diewert, W. E. (1982). The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica*, 50(6), 1393–1414.
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429–444.
- De Neve, J. E., Imbert, C., Spinnewijn, J., Tsankova, T., & Luts, M. (2021). How to improve tax compliance? Evidence from population-wide experiments in Belgium. *Journal of Political Economy*, 129(5), 1425–1463.
- Dubin, J. A., & Wilde, L. L. (1988). An empirical analysis of federal income tax auditing and compliance. *National Tax Journal*, 41(1), 61–74.
- Dularif, M., & Rustiarini, N. W. (2021). Tax compliance and non-deterrence approach: A systematic review. *International Journal of Sociology and Social Policy*, 42(11/12), 1080–1108.
- Färe, R., Grosskopf, S., Norris, M., & Zhang, Z. (1994). Productivity growth, technical progress, and efficiency change in industrialized countries. *The American Economic Review*, 84(1), 66–83.
- Fenochietto, R., & Pessino, C. (2013). *Understanding countries' tax effort* (IMF Working Paper No. WP/13/244). International Monetary Fund.
- Førsund, F. R., Kittelsen, S. A. C., Lindseth, F., & Edvardsen, D. F. (2006). The tax man cometh-but is he efficient? *National Institute Economic Review*, 197, 106–119.
- Fuentes, R., & Lillo-Bañuls, A. (2015). Smoothed bootstrap Malmquist index based on DEA model to compute productivity of tax offices. *Expert Systems with Applications*, 42(5), 2442–2450.
- Gangl, K., Hoffman, E., & Kirchler, E. (2015). Tax authorities' interaction with taxpayers: A conception of compliance in social dilemmas by power and trust. *New Ideas in Psychology*, 37, 13–23.
- Graetz, M. J., & Wilde, L. L. (1985). The economics of tax compliance: Fact and fantasy. *National Tax Journal*,

- 38(3), 355–363.
- Gündem, P. (2022). Türkiye’de Vergi Uyumu ve Mekansal Farklılıklar, 2013-2020. *Fiscaoeconomia*, 6(3), 910–942.
- Hocaoğlu, H., Koca, G., Behdioğlu, S., & Demirhan, M. (2023). Vergi uyumunu etkileyen faktörlerin dematel yöntemi ile analizi: Kütahya ili örneği. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (58), 145–168.
- Katharakis, M., & Tsakas, M. (2010). Assessing the efficiency and managing the performance of Greek tax offices. *Journal of Advances in Management Research*, 7(1), 58–75.
- Kirchler, E., Hoelzl, E., & Wahl, I. (2008). Enforced versus voluntary tax compliance: The “slippery slope” framework. *Journal of Economic Psychology*, 29(2), 210–225.
- Kneip, A., Simar, L., & Wilson, P. W. (2020). Inference in dynamic, non-parametric models of production: Central limit theorems for Malmquist indices. *Econometric Theory*, 37(3), 537–572.
- Kogler, C., Olsen, J., Kirchler, E., Batrancea, L. M., & Nichita, A. (2023). Perceptions of trust and power are associated with tax compliance: A cross-cultural study. *Economic and Political Studies*, 11(3), 365–381.
- Kogler, C., Olsen, J., Müller, M., & Kirchler, E. (2022). Information processing in tax decisions. *Journal of Behavioral Decision Making*, 35(4), 1–15.
- Kornhauser, M. E. (2006). A tax morale approach to compliance: Recommendations for the IRS. *Florida Tax Review*, 8(6), 601–640.
- Lisi, G. (2019). *Slippery slope framework, tax morale and tax compliance: A theoretical integration and an empirical assessment* (Discussion Papers in Economic Behaviour No. 0219). University of Valencia, ERI-CES.
- Liu, X., Wu, Z., Xie, P., Hong, Q., Hu, X., & Antwi-Afari, M. F. (2022). Environmental efficiency evaluation of construction waste generation based on data envelopment analysis and Malmquist index. *Buildings*, 12(8), 1–12.
- Marshall, M. E., Farrar, J. M., Massey, D. W., Thorne, L., Wu, A., & Bui, T. (2024). In all fairness: A meta-analysis of the tax fairness-tax compliance literature. *Behavioral Research in Accounting*, 36(2), 105–130.
- Martínez, Y. U., Arzoz, P. P., & Arregui, I. Z. (2022). Tax collection efficiency in OECD countries improves via decentralization, simplification, digitalization and education. *Journal of Policy Modeling*, 44(2), 298–318.
- Moesen, W., & Persoon, A. (2002). Measuring and explaining the productive efficiency of tax offices: A non-parametric best practice frontier approach. *Review of Business and Economic Literature*, 47(3), 399–416.
- Muehlbacher, S., Kirchler, E., & Schwarzenberger, H. (2011). Voluntary versus enforced tax compliance: Empirical evidence for the “slippery slope” framework. *European Journal of Law and Economics*, 32, 89–97.
- Nartey, E. (2023). Tax compliance of small and medium sized enterprises in Ghana. *International Journal of Sociology and Social Policy*, 43(11/12), 1063–1083.
- OECD. (n.d.). *Digital transformation of tax administration*.
- Ogbolu, L. M., Ashibogwu, N. K., Chukwugeku, I. M., & Mark, T. (2024). Personal income tax and tax compliance on economic development in Nigeria. *ISAR Journal of Economics and Business Management*, 2(4), 28–36.
- Prinz, A., Muehlbacher, S., & Kirchler, E. (2014). The slippery slope framework on tax compliance: An attempt to formalization. *Journal of Economic Psychology*, 40, 20–34.
- Randlane, K. (2015). Tax compliance as a system: Mapping the field. *International Journal of Public Administration*, 39(7), 515–525.
- Sanchez, J. J. V. (2018). Malmquist index with time series to data envelopment analysis. In V. Salomon (Ed.), *Multi-criteria methods and techniques applied to supply chain management* (pp. 77–91). IntechOpen.
- Shojaie, S. E., Sadjadi, S. J., & Moghaddam, R. T. (2024). Malmquist productivity index for two-stage network systems under data uncertainty: A real world case study. *PLOS ONE*, 19(7), 1–35.
- Simar, L., & Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, 44(1), 49–61.
- Slemrod, J. (2019). Tax compliance and enforcement. *Journal of Economic Literature*, 57(4), 904–954.
- Syamsudin, S. H., Damayanti, R. A., & Fattah, S. (2023). Exploring of tax compliance: Tax practitioners interpretation. *Jurnal Multidisiplin Madani*, 3(2), 392–403.
- Tabbach, A. D. (2002). Criminal behavior, sanctions, and income taxation: An economic analysis. *The Journal of Legal Studies*, 32(2), 383–406.
- Tilahun, M. (2019). Determinants of tax compliance: A systematic review. *Economics*, 8(1), 1–7.
- Tohidi, G., & Razavyan, S. (2013). A circular global profit

Malmquist productivity index in data envelopment analysis. *Applied Mathematical Modelling*, 37(1-2), 216–227.

Tohidi, G., Razavyan, S., & Tohidnia, S. (2012). A global cost Malmquist productivity index using data envelopment analysis. *Journal of the Operational Research Society*, 63(1), 72–78.

Tsaregradskaya, Y. K. (2021). Tax compliance: Approaches to understanding. *Lex Russica*, 74(12), 66–73.

TÜİK. (2023). *Indicators of 100 years* [Dataset].

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