RIGA TECHNICAL UNIVERSITY
Faculty of Computer Science, Information Technology and Energy
Institute of Applied Mathematics

Elīna Elizabete Pabērza

Academic Master (Second Cycle) Studies program
Financial Engineering Mathematics
student, student ID No. 181RIB386

A DATA DRIVEN ASSESSMENT OF THE SHADOW ECONOMY

MASTER THESIS

Scientific supervisor, Dr. oec., associate professor
Māris Jurušs

RIGA 2025

RIGA TECHNICAL UNIVERSITY

Faculty of Computer Science, Information Technology and Energy
Institute of Applied Mathematics

Work Performance and Assessment Sheet of the Master Thesis

The author of the graduation thesis:

student Elīna Elizabete Pabērza ______________________

(signature, date)

The graduation thesis has been approved for the defense:

Scientific supervisor:

Dr. oec., associate professor Māris Jurušs ______________________

(signature, date)

ANOTĀCIJA

ĒNU EKONOMIKA, BALTIJAS VALSTIS, VAIRĀKU INDIKATORU UN VAIRĀKU CĒLOŅU MODELIS (MIMIC)

Šajā maģistra darbā tiek analizēts un pētīts ēnu ekonomikas apmērs Baltijas valstīs. Tas ir fenomens makroekonomikā, ko nav iespējams tieši aprēķināt. Dažādu iemeslu dēļ sabiedrības locekļi veic ekonomiskas darbības un aktivitātes, kuras nereģistrē vai nedeklarē, līdz ar ko valsts par tām nezina un nevar ieskaitīt nacionālā iekšzemes kopprodukta (IKP) aprēķinā. Tieši tāpēc tiek izvirzīts darba mērķis izstrādāt datos balstītu MIMIC modeli priekš ēnu ekonomikas izmēra aplēšanas Baltijas valstīs.

Iesākumā tiek analizētas jau pieejamās aplēses no esošiem pētījumiem par ēnu ekonomikas apmēru 3 Baltijas valstīs: Latvijā, Lietuvā, Igaunijā. Pēc tam tiek apskatīts teorētiskais ietvars par šo tēmu – gan ietekmējošie faktori, gan aprēķināšanas metodes, kas radušās gadu laikā. Pēc tam izmantojot publiski pieejamus makroekonomikas datus tiek izveidots MIMIC modelis katrai valstij. No izveidotajiem modeļiem tiek aplēsts ēnu ekonomikas apmērs Baltijas valstī no 2000. līdz 2023. gadam, kas tiek salīdzināts ar esošu pētījumu datiem.

Rezultātā tiek iegūts statistiski nozīmīgi, tomēr atšķirīgi modeļi katrai valstij, kā arī, izmantojot citos pētījumos iegūtos rezultātus, aprēķināts ēnu ekonomikas apmērs pētāmajā periodā visās 3 valstīs. Nobeigumā tiek veikti secinājumi par kopīgo un atšķirīgo, kā arī praktiskiem un teorētiskiem secinājumiem, kas var būt noderīgi turpmākos pētījumos.

Šis maģistra darbs sastāv no 70 lapām, 8 nodaļām, 13 tabulām un 26 attēliem.

ABSTRACT

SHADOW ECONOMY, BALTIC COUNTRIES, MULTIPLE INDICATORS AND MULTIPLE CAUSES MODEL (MIMIC)

This master's thesis analyzes and examines the size of the shadow economy in the Baltic States. The shadow economy is a macroeconomic phenomenon that cannot be directly measured. For various reasons, members of society engage in economic activities that are not registered or declared, meaning the state is unaware of them and thus cannot include them in the calculation of national gross domestic product (GDP). Therefore, the goal of this thesis is to develop a data-based multiple factors and multiple indicators (MIMIC) model for estimating the size of the shadow economy in the Baltic countries.

The thesis begins by reviewing existing estimates from previous studies on the shadow economy in the three Baltic States: Latvia, Lithuania, and Estonia. It then explores the theoretical framework of the topic – both the influencing factors and the calculation methods that have emerged over the years. Using publicly available macroeconomic data, a MIMIC model is developed for each country. Based on these models, the size of the shadow economy is estimated for the period from 2000 to 2023 and compared with data from existing research.

As a result, statistically significant but different models are created for each country. Using results from other studies, the size of the shadow economy is calculated for the reserach period in all 3 countries. The thesis concludes with a discussion of similarities and differences, as well as practical and theoretical implicatios that may be useful for future research.

This master thesis consists of 70 pages, 8 sections, 13 tables, and 26 figures.

TABLE OF CONTENTS

INTRODUCTION

Relevance of the Topic: The shadow economy in Latvia has always been a relevant and publicly debated topic. Since regaining independence in 1991, Latvia has experienced significant development and growth across various sectors, both politically and economically. However, in recent years, the public has increasingly sought to compare Latvia's economic situation with that of other European countries and Scandinavia, placing particular emphasis on comparisons with neighboring countries – Lithuania and Estonia.

In such discussions, the shadow economy is often mentioned, as its scale reduces the visible economy, hinders national development, and demotivates the honest part of society from continuing to act fairly and to trust in the state system. A particularly interesting aspect of this topic is the role of psychology and historical context in influencing decisions – why some members of society still choose to participate in the shadow economy.

Various studies yield differing results, but the shadow economy in Latvia is commonly estimated at around 25% of GDP (Schneider, 2022). There are also different methods for estimating it, but since these are market activities hidden from the state, determining the actual size is complex, and so estimates are necessary. It is common to use micro-methods – direct estimations based on surveys of various market participants. However, results obtained in this way have many shortcomings, which is why this master's thesis will focus on a macro approach, estimating the potential scale of the shadow economy using various market and economic data and indicators.

It is important to understand the main causes and drivers of the shadow economy and to calculate its size. This can help bring attention to the issue and encourage society to participate in the visible part of the economy, which is more beneficial to all members of society in the long term – although many have not yet realized this.

Objective of the Master's Thesis: To develop a data-based method for assessing the shadow economy in the Baltic countries and forecasting trends in its development.

Tasks of the Master's Thesis: Develop a method/model to estimate the shadow economy (as a share of GDP), assess the size of the shadow economy in the Baltics, compare the results with other available studies and analyze possible reasons for discrepancies, forecast shadow economy trends for the next five years, provide recommendations for reducing the shadow economy

Problem statement (hypothesis): The MIMIC model developed by the author, calibrated using known shadow economy estimates, will provide a consistent and scalable framework for estimating the size of the shadow economy in Latvia, Lithuania, and Estonia.

The thesis consists of 8 sections: introduction, analysis of current situation, review of theoretical framework, model description, raw data analysis, application of the model, discussion of results and conclusions.

1. THE EXTENT OF THE SHADOW ECONOMY IN THE BALTICS

This thesis will focus on the Baltics region; therefore, below is a summary of available data and data sources published by the Ministry of Finance Republic of Latvia.

According to the news article published on the ministry's webpage the ministry created a special project with the aim to develop plan to develop plans and policies on how to reduce the shadow economy in Latvia. As the project manager was selected Arnis Sauka, which together with his team developed various essential documents such as evaluation of current situation, detailed methodology, recommendations and other (*Pētījumi par ēnu ekonomikas apjomu | Finanšu ministrija*, n.d.).

The ministry also reference Schneider's most recent research paper where he continues his work in the field and updates the data with the results from 2021 and 2022 (Schneider, 2022). The MIMIC method is applied; however, during this research paper the Coronavirus pandemic occurred, which is one of the main factors for the shadow economy increase. According to the calculations the size of shadow economy in 2022 is 22.74, 19.94 and 22.36 percent of GDP in Estonia, Latvia, Lithuania respectively (Schneider, 2022). The trend is also seen in Figure 1 below.

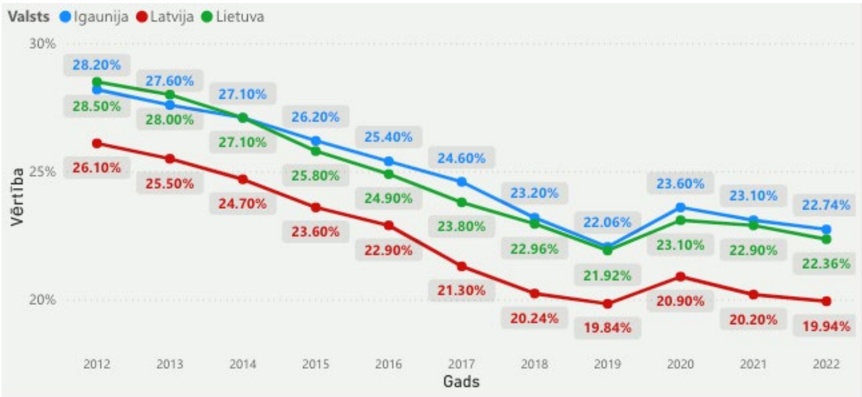

Figure 1. Recent research results in Baltic countries by Schneider based on MIMIC model (retrieved from (Finanšu Ministrija, 2022)

Note. Numbers indicated are size of shadow economy expressed as percentage of official GDP.

Alternatively, the ministry uses data from in the previous sections mentioned research by Arnis Sauka. Refer to Figure 2 for the results.

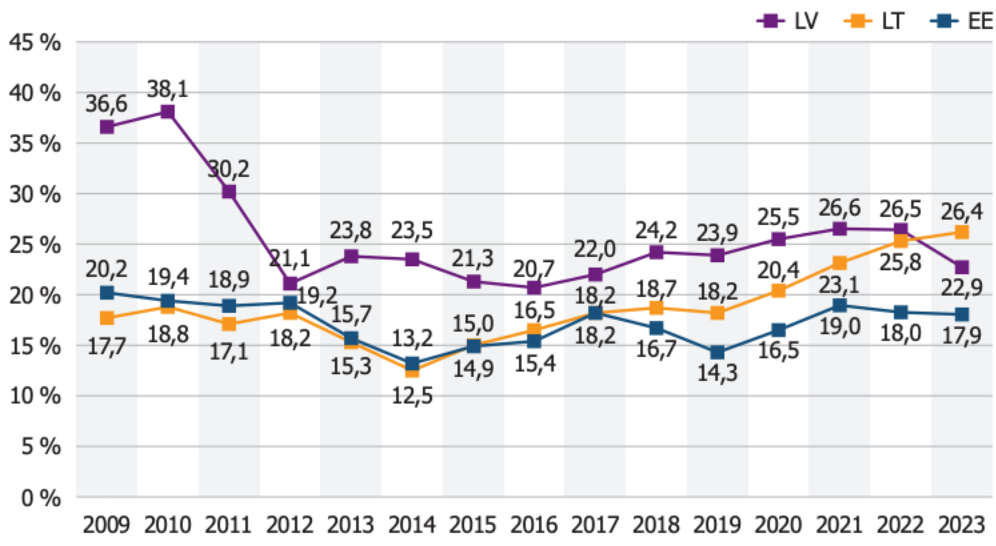

Figure 2. Research results by Sauka based on survey data (retrieved from ((Ēnu Ekonomikas Indekss 2023, n.d.))

Conversely, the results are different both in the absolute numbers and trends between countries. Clear pattern is seen that results from survey methodology used by Arnis Sauka shows smaller the estimated size of shadow economy if compared with the MIMIC model by Friedrich Schneider.

Nowadays there is much of political interest in the size of official GDP as various national interests depend on its size. Some examples include potential access to credit, it also determines whether or not the country will be part of country alliances such as G-5, G-7, G-10 or G-22 and other. The official GDP is extremely important measure in the context of European Union, for example, the fiscal deficit and public debt criteria set in Maastricht criteria if including unofficial economy would be easier to reach (Tanzi, 1999). In addition, there are various economists who have claimed if the underground economy would be fully accounted and included in the official output of GDP, the largest country club could change significantly.

These two figures clearly represent the essence of the complexity and contradiction within this topic. In addition, Fredirick Schneider and Arnis Sauka are well-known and recognized researchers of this topic, especially in the Baltics. Yet each of them uses very different estimation methods. Those are in detailed discussed in the next section.

After initial research, it has become clear that in the Baltics there is very limited number of researchers and available results and publications about the size of the shadow economy. This is a good indicator that additional research about this topic would help to broaden the perspective and offer alternative estimated results of shadow economy volumes in the Baltics.

2. ANALYTICAL REVIEW OF THEORETHICAL FRAMEWORK

This section of master's thesis will provide a detailed and analytical review of scientific literature and theoretical aspects that should be considered in the context of shadow economy. Firstly, it will discuss the history and various already available definitions of this term. Then dive deeper into the macro and microeconomic factors that influence and impact shadow economy. Taking into account that this thesis focuses on the Baltic states, then there are historical and psychological aspects that have impacted the development of shadow economy here. Finally, the paper will discuss the numerous methods and models for the determination of shadow economy size based on which a specific model will be selected for further calculations in the practical section of this thesis.

It is interesting to note that more than 45 years ago in 1979 this topic was topical and to be honest it is still to this day. Methods and assumptions have been developed, but problem stays. Feige (1979) wrote:

"So much economic activity is excluded from the national income statistics that the macroeconomic judgments currently made by economists must of necessity be wrong".

2.1. Definition of the shadow economy

Various authors and scientists have tried to define what is shadow economy. It is not as easy as it may seem since it is hard to define something that the society is trying to hide from the government institutions. In addition, some consider research in this area as passion for "knowing the unknown" (Schneider, 2011).

Feige (2015) described the shadow economy using various simple words that the society often uses, for example, grey, black, off the books, undeclared, unreported, illegal, informal. This helps to understand what it is using informal language. However, there are also more formal versions available.

There are still some disagreements about it. One definition that is well formulated is all market-based legal economic activity and production of goods and services that are hidden from the government due to the following reasons: to avoid tax or social security contributions, to avoid meeting labor market laws like minimum wage, maximum working hours or safety standards, to avoid filling various administrative forms (Schneider, 2011).

Some defines this sector as market-based production of goods and services, whether legal or illegal, which escapes detection in the official estimates of GDP ("The Underground Economy," 1997).

Bhattacharyya (1999) argues that the hidden economy is best described as unrecorded national income, calculated as the difference between the potential national income for the given currency in circulation and the recorded national income.

It must be noted, generally there is an agreement on and nor most of the offered definitions above, nor this master's thesis do not deal and cover illegal actions such as economic crime, burglary, robbery (Samuelson et al., 2000).

This master theses will further dive into this concept focusing on the Baltics region. Definition used by Arnis Sauka and Tālis Putniņš who are well-known professors and researchers in the Baltics whose main research object is shadow economy. They say that term shadow economy refers to all legal production of goods and services that is willingly and advisedly not disclosed to the public authorities. It also corresponds to what the Organization for Economic Co-operation and Development (OECD) in their comprehensive 2002 handbook "Measuring the Non-observed Economy" (OECD et al., 2002). It is consistent with what other researchers have employed (Schneider et al., 2010).

Most authors agree that it is a fuzzy concept and there is not one definition. Generally, it can be described as as legal economic, and productive activities that, if recorded, would contribute to national GDP.

Moreover, there are various ways and transactions how person can participate in the shadow economy. Often even not being aware of it. These components are similar across all three Baltic states. Those are underreporting of salaries, underreporting of employees and underreporting of income, see details in Figure 3 (Sauka & Putniņš, 2024).

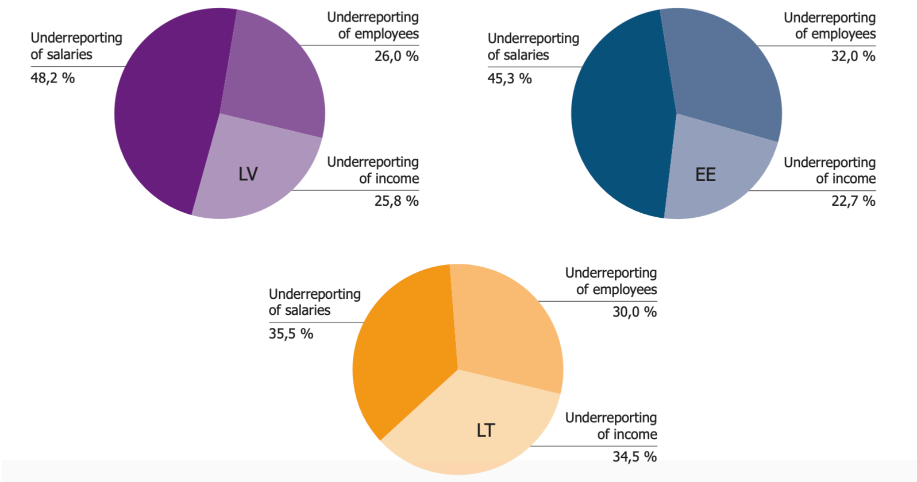

Figure 3. Components of the shadow economy in Baltics in 2023 (retrieved from (Sauka & Putniņš, 2024))

2.2. Factors affecting shadow economy

If one takes a step back and look to the big picture called macroeconomics, he or she understands that universal measure of the size of countries' economy is gross domestic product (GDP). Gross domestic product (GDP) was first introduced in the 1930s by Nobel laureate Simon Kuznets when the National Bureau of Economic Research was asked to develop a set of national economic accounts (Samuelson et al., 2000). GDP can be measured in various ways: the expenditures approach, income approach and value-added approach (Sahin, 2009). It can be summarized that calculations are different yet the idea is similar sum everything produced in the country in specific period, for example, the sum of expenditures, or purchases, by final users as per the expenditures approach. It is interesting that currently "GDP is a variable that has become the de facto universal metric for the standard of living" (European Commission & European Parliament, 2009). This is important because shadow economy's size is the value that is not included in this universal metric and, therefore, frequently expressed as percentage of GDP. Then why would members of society participate in the grey area of economy? There are various reasons and those will be further discussed in next section.

Most of the participants have their own reasoning; however, those can be categorized in three categories: historical background, low tax morale, certain industry specifics and way of survival. Although, all of these are closely related there are different aspects that this paper will look into.

2.2.1. Historical background

From the more than 100 years of history, Latvia has been occupied by the Soviet Union (USSR) from 1940s until 1990s. This have created significant impact on various generations and spheres of life that impact the macroeconomics and both public and private sector to this day (Kasekamp, 2017). Soviet Union implemented state socialism model which advocated for collective ownership of resources and industries and it was planned economy that the economy did not rely on market competition but rather on set rates and planned results by the government (Grudule & Kalnačs, 2024). However, as it is known this was not a successful model as it has various flaws. In general, the years spent under the Soviet flag have been known for high bureaucracy, inequality, unfairness from the government institutions. Large part of Latvian society has experienced these times and still have the intention to hide income or expenses from the government, do not register the company in official register to avoid tax or other inspections. Latvia restored independence in 1991. People aged 34 and below have not experienced the soviet times. As can be seen in Figure 4, the majority or 69% or 1.1 million have such experience.

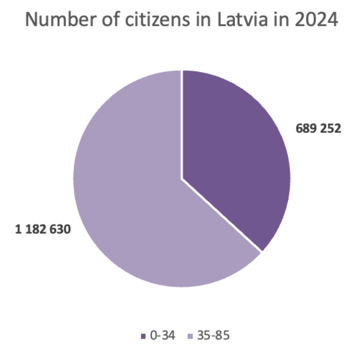

Figure 4. Number of citizens in Latvia in 2024 by age (modified from (Iedzīvotāju skaits pēc tautības un pa vecuma grupām gada sākumā 2011 - 2024, n.d.))

These people have experienced situations when official and declared business transactions did not result in a success or any benefit but rather a burden.

Alexeev (2001) in his research about the unofficial part of economy in the former Soviet republics concludes that the size of undeclared economy is more of a historical phenomenon and it has less impact from contemporary institutional factors. This is based on the fact that strong relationship between the size of unofficial economy and governance was discovered by Johnson in 1997 (Johnson et al., 1997). This research not only tried to estimate the size of unofficially economy as input using electricity consumption, but also examined the effects of political control on economic performance. The same author later found assurance on the initial conclusion since similar results were generated using amplified databases including OECD (Johnson & Kaufmann, 2001).

This leads to the next category. In the 1990s, various countries previously occupied by the Soviet Union, restored their independence. In the period of transition from socialism to capitalism, the very different political approaches and control mechanisms of economic growth between these two ideologies led to different results.

2.2.2. Tax morale in the Baltics

Many different authors frequently associate shadow economy with low tax morale and without a doubt it is highly connected with country's history which was discussed previously. It is the estimate of the symmetry between the laws and regulations of formal institutions and the unwritten socially shared rules within society (Williams & Horodnic, 2015). In simple terms the willingness to pay taxes and other duties to the government institutions.

How tax morale is connected with the shadow economy? As it was already discussed shadow economy is undeclared economic activity. Question is why the society would not declare it. And one of the reasons might be to avoid tax payments. Society commits to pay taxes, and the government and municipalities provides social security, takes care of surrounding environment, education, healthcare system and other (Rutkauskas, 2019). However, in real life the every taxpayer or member of the society in one way or the other makes a cost-benefit analysis before paying taxes (Rutkauskas, 2019).

The final conclusion in the research by Williams and Herodnic in 2015, was that the lower is the tax morality of individuals, the likelihood of participating in the shadow economy is greater. The conclusion made was based on the data from the

Baltic states. Also, Dean in 2006 obtained evidence that supports the hypothesis that tax morale has significant impact on the size of shadow or grey economy.

Based on this conclusion it can be assumed that the cost-benefit analysis performed by some of the taxpayers makes them conclude that it is not beneficial to pay taxes. One of the reasons for that is whether the tax payer beliefs that the tax payment is fair for all of the other taxpayers This is the "Kantian" morality approach (Dean, 2006).

Another reason is satisfaction with the political system and government, higher satisfaction leads to higher tax morale (Torgler, 2004).

In addition, it has been analyzed whether there exists relationship between corruption and tax morale. Various authors show strong evidence that relationship exists and tax compliance is positively related to environment in which the society is protected from corruption (Picur & Riahi-Belkaoui, 2006).

Let's look into the actual data about these three aspects in Latvia. Survey results from 2024 show evidence that 69 percent of citizens thinks that the tax system in Latvia is not fair to them, also the proportion of citizens who think that the system is unbiased has decreased from 26% to 21% (*Aptauja*, 2023). Results from surveys of Estonians show that also Estonians willingness to pay taxes is decreasing, for example, decrease in results for those who think that tax system is easy to understand and increase in those who think that they could buy goods and service in the black market (*Survey Shows a Decline in Estonian Residents' Willingness to Pay Taxes | Estonian Tax and Customs Board*, 2024). Not surprisingly, the results of fairness in Estonia can be assessed as better than in Latvia since 65% of Estonian respondents do not consider the tax system fair (*Survey Shows a Decline in Estonian Residents' Willingness to Pay Taxes | Estonian Tax and Customs Board*, 2024). Also, Lithuanians have various improvement points for their tax system. As noted by the European Commission in the country overview, there are 72 differentiated income tax treatments that make the system complex and harder to consider fair (European Commission. Directorate General for Economic and Financial Affairs., 2023).

As shows Figure 5, compared with other OECD countries Latvia (LVA) respondents have one of the highest proportions for the low or no trust answer (marked in yellow), and it is around 55%. In Estonia it if lower, around 50%. Answers were collected in 2023. Figure 6 is extract of answer to one of the questions that was asked to Latvians about corruption – has the person participated in any action that involve corruption in the last two years? Answer shows that 19% have. It also shows that results vary based on gender, region and age.

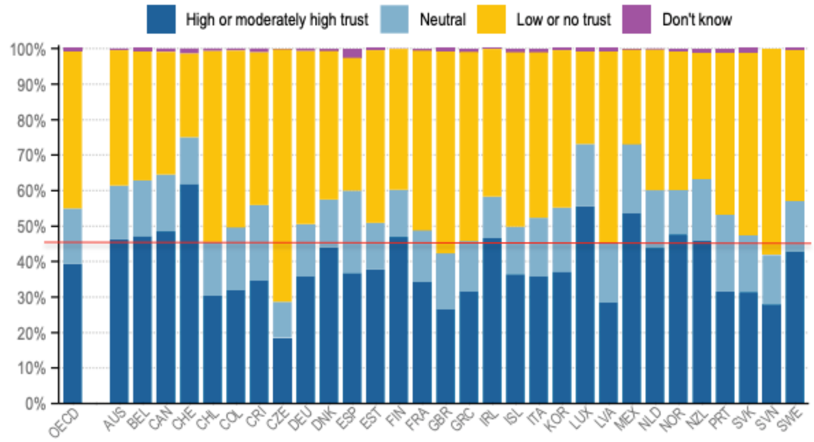

Figure 5. OECD country comparison on levels of trust in their national government (modified from (OECD, 2024))

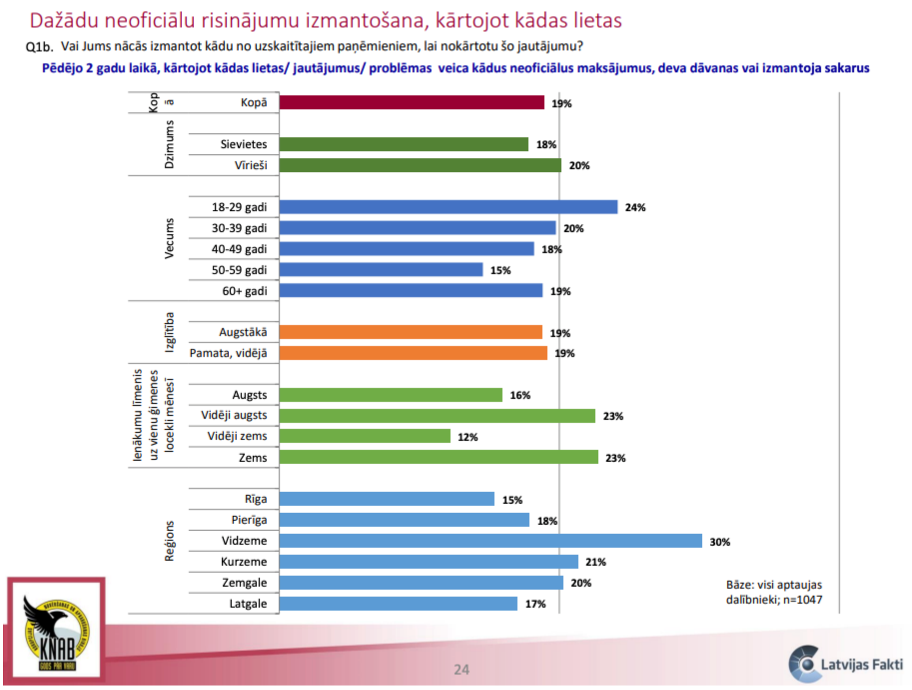

Figure 6. Corruption questionnaire results in Latvia (retrieved from (KNAB, 2023))

2.2.3. Industry specifics

Finally, exists evidence both statistically proven and just agreed in the society that some industries are more in the shadows than other. Research papers show that sectors which are more complex to tax, logically show larger shadow economy (Dean, 2006). Business transactions which are taxable can be distinguished in 2 categories: that can be audited directly by the government and that cannot. There is even a term for the later “hard-to-tax” (HTT) markets (Vázquez, 2003). Every year Putnins and Sauka updates their research with the actual data about shadow economy, including the sectors which are more in the shadows (Sauka & Putniņš, 2024). As can be seen in Figure 7, per 2023 data the top such sectors are construction, retail, services, manufacturing, wholesale.

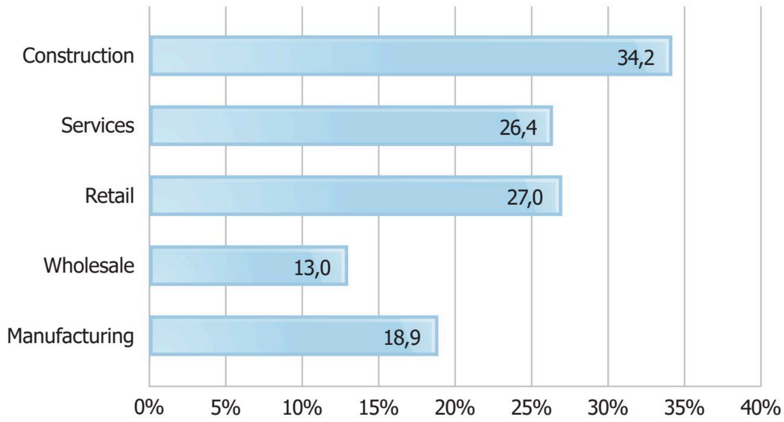

Figure 7. Size of shadow economy by sector in Latvia in 2023 (retrieved from (Sauka & Putniņš, 2024))

The next thing that needs to be discussed is why exactly these industries are the ones that have undeclared business transactions. Construction by far is one of the largest problems. In 2016 created research about the shadow economy in construction sector in Latvia, Arnis Sauka tried to find the causes (Sauka, 2016b). He identified some of the most frequently mentioned ones: significant competition in comparingly small market with low volumes and situation that in most procurements the main objective for selecting the winner is lower price (Sauka, 2016a).

Another interesting sector that is serious problem in Latvia is services. Two service provider categories that were in the spotlight in 2021 during covid crisis were beauty care, beaty salons and cafes, restaurants, hotels. These industries suffered very heavily during covid lockdowns since these services were not allowed to provide. These companies needed to survive months without income; however, there were instances where the government support was minor, due to declared part of business or employee official salaries being small consequently the support was proportionally small (lvportals.lv, 2021). Why this problem is even more worrying is that in the beauty industry it is not only the salaries from which taxes are not paid. In other words, not registering in state revenue service databases. Another significant aspect is that many beauty service providers are not officially registered in the respective institutions such as ministry of health and health inspection which are responsible for monitoring of cleanness, specialist licensing and other essential aspects (lvportals.lv, n.d.). This is another reason why people are in the shadow economy-to avoid the fulfillment of these requirements.

In order to fight with this problem, the government has developed a "Ēnu ekonomikas ierobežošanas plāns 2024-2027 gadam" (*Ēnu ekonomikas ierobežošanas plāns 2024.–2027. gadam*, 2024). It contains also measurable goals, for the beauty those are summarized in Figure 8.

Nozare	2023.gads	2024.gads	2025.gads	2026.gads	2027.gads	Piezīmes
Nodokļu maksātāju - juridisko personu bez būtiskiem nodokļu saistību izpildes riskiem ("A" klasē) īpatsvars kopējā segmentēto NM JP skaitā uz gada sākumu, %	4,9%	5,3%	5,7%	-*	-*	Paredzēts pieaugums ne vairāk par 0,5% katru gadu, pieauguma tempam nepārsniedzot 10%.
Nodokļu maksātāju - juridisko personu bez būtiskiem nodokļu saistību izpildes riskiem ("B" klasē) īpatsvars kopējā segmentēto NM JP skaitā uz gada sākumu, %	49,9%	49,5%	49,1%	-**	-**	Paredzēts samazinājums ne vairāk par 0,5% katru gadu, samazinājuma tempam nepārsniedzot 10%.
Darba ņēmēju ar ienākumiem virs 1000 *euro* skaita īpatsvars kopējā nozares darba ņēmēju skaitā (vidējais rādītājs mēnesī), %	6,4%	7,0%	7,6%	8,1%	8,4%	Prognozēts ar autoregresiju, pesimistiskais scenārijs.
Darba vietu ar ienākumiem virs 1000 *euro* skaita īpatsvars kopējā nozares darba vietu skaitā (vidējais rādītājs mēnesī), %	5,6%	6,1%	6,6%	7,0%	7,3%	Prognozēts ar autoregresiju, pesimistiskais scenārijs.

Figure 8. Goals for the beauty industry for 2024-2027 (retrieved from ((Ēnu ekonomikas ierobežošanas plāns 2024.–2027. gadam, 2024))

Note. * Value cannot be predicted, but positive tendency expected,
** value cannot be predicted, but decrease is expected.

The uniting factor that can be seen is that shadow economy covers the whole industry. Even if one company wanted to move out of the shadows it would lose competitive advantage and would not be able to offer lower prices.

To sum up, there are various factors that impact the size of the shadow economy. Without a doubt it is country's and industry's history. Various researchers have proven that for former Soviet Union countries, which have been in transition from socialism to capitalism the shadow economy proportion is larger and is impacting its development to this day. It is also important to mention tax morale which can be explained as taxpayer willingness to pay taxes. Data shows that in Latvia the tax morale is low. Finally, the shadow economy differs between industries. Therefore, also each industry specifics impacts the grey area in it. Due to various reasons, construction, retail and services have the higher risk.

2.3. How to estimate the size of shadow economy?

Now that we have discussed what is shadow economy and what impacts it, we could summarize the main estimation methods of its size. This is more of an estimation rather than calculation. Firstly, because this is something that the society wants to hide from institutions, therefore, there is no straightforward data about. Rather estimation formula based on multiple assumptions. Secondly, in these decisions various sociodemographic and psychological aspects are important, which may not always be rational or possible to calculate using mathematical formula. Now let’s look what methods scientists have developed so far. In general, there are two main estimation categories. One being the direct method, the other indirect method (Postea & Achim, 2023).

2.3.1. Direct or micro methods

Direct methods mean direct approach to estimate the shadow economy from micro-level observations. What can be more direct than surveys and tax audits? In the following paragraphs are examples of direct estimation methods.

There are several research papers written about conclusions derived from surveys about underground economy, suppliers, household activity, work forms. In 1994, research on micro data from a randomized survey carried out in Quebec, Canada, which focused on the work in the 'underground' economy and tax evasion (Lemieux et al., 1994). Similarly, Leonard (1994) performed survey in Belfast, United Kingdom on self-servicing household economic activity. In addition, scientific research paper about the usage popularity of informal suppliers in the United States from 1981 to 1985 (McCrohan et al., 1991). (Warde, 1990) used surveys to analyze different types and forms of work that members of society engage in, later deriving various conclusions on several alternative ways of delivering use-values that are connected with grey economy. However, these did not directly try to estimate the undeclared part of economic activity.

In contrast, there is a study based on surveys of households (Kazemier & van Eck, 1992). Here the questions were directed to households. Questions focused on home maintenance, unreported labor, tax evasion to later estimate part of the shadow economy.

In the Baltics, one of the most popular and frequently cited methods is surveys of company managers. It studies their attitude in order to measure the shadow economy (Putniņš & Sauka, 2015). Authors focused on developing a well-designed survey and used it to research satisfaction with the government rather than tax evasion or undeclared economic activity. The survey was divided into four main categories such as external influences, shadow activity, company and owner characteristics and company manager' s attitudes (Putniņš & Sauka, 2015). For the estimation they derived a formula for shadow economy index which is based on the GDP calculation income approach – which when simplified is made up of personal income (wages and salaries) and corporate income (profits generated by companies). Basically, it calculates the unreported proportion of employee remuneration using the answers in survey. Then estimate the shadow proportion of company's income from the unreported employee remuneration and operation income calculated in the previous step multiplying with ratio of employees' remuneration from Eurostat. Finally, take the weighted average of shadow proportion across sample of firms in selected country (Putniņš & Sauka, 2015).

Another method for the estimation of the shadow economy is tax audits. As it was previously discussed, tax evasion is one of the reasons why someone would participate in the shadow economy; therefore, tax audit results from local state revenue services are used to estimate the level of tax evasion (Feinstein, 1991). In addition, conclusions from scientific conference in Poland concludes that tax conditions have direct impact on existence of entities in the shadows (Shkarupa et al., 2019). They conclude that tax audits can be used to estimate the growth and scale of the shadow economy. In table 2.3.1 below is summarized the main advantages and disadvantages of direct methods.

2.3.1 table

Advantages and disadvantages of direct methods

Advantages	Disadvantages/limitations
Provide detailed information and analysis since can focus on specific sector, region, enterprises	Expensive and time consuming since many interviews must be performed
Can provide more detailed insights and guidance to policymakers	Higher underestimation risk of the total size of the shadow economy due to non-response or untruthful responses
Is precise about what parts, aspects are included in the estimates	Audits only detect a non-random fraction of all tax evasion
	Audits are also typically non-random, which bias estimates

2.3.2. Indirect or macro methods

This section will focus describing in more detail the indirect approach which in scientific papers are also known as macro methods. Why macro? Because the idea of these models is built on the macroeconomic data. There is data about all economic activity both reported and hidden activities. In essence, these models use this data to estimate the size of shadow economy by using formula: estimated size of all economic activity minus recorded economic activity (Schneider & Buehn, 2013). This leads to the macroeconomic data that contains all economic activity. These are transactions, currency and electricity, which are also the names of various indirect methods which will be discussed in the following sections.

Currency demand or monetary approach

This was one of the first methods and approaches on the estimation of shadow economy made public. For this approach the main idea is that undeclared or shadow transactions are made in petty cash in order to leave evidence of such transactions. Gutmann (1977) published article where he estimated the size of irregular economy in the United States of America in 1976 of 200 billion USD which was approximately 10 percent of the GNP by then. His approach which was later used by many is to examine the composition of the stock of money-currency and demand deposits. The calculation is straightforward:

- Currency/demand-deposit ratio in the period;
- Estimated the currency needed for legal transactions;
- The gap between the estimated result and actual currency outstanding is the cash used for shadow or irregular activities;
- Calculate that illegal cash generated irregular income which is the estimation result.

However, it must be noted that these calculations heavily rely on four main assumptions which were later widely criticized by Edgar L. Feige and others (Feige, 1979).

Transactions approach

This approach assumes that there is a constant relationship between the money supply for economic transactions and total value added or GDP (Feige, 1979). Feige in 1979 developed this method which was based on the Irving Fisher's argument that total macroeconomic activity can be calculated by adding all transactions paid for by cash and by check. Feige estimated it by:

- Calculating total transactions in an economy (checking transactions + currency transactions);
- Calculate the ratio of total transactions to observed income in a society multiplied by the physical volume of transactions and price index of new goods and real income;
- Adjust it my removing turnover rate from financial centers since there the variable is too volatile.

The results of this study showed irregular economy's size in the United States of America in 1976 of 19.1 % against 10 percent estimated by Gutman (Feige, 1979).

Electric consumption approach

Daniel Kaufman and Alexander Kaliberda where one of the first ones who identified variable that is closely related to the development of GDP everywhere in the world-electricity (Kaufmann & Kaliberda, 1996). They suggested that the increase in electricity consumption is a great indicator of increase in GDP in all over the world.

This is one of the most interesting approaches and the logic is similar to other indirect methods-electricity consumption is observable variable that could be a good measure of all economic activity both official and unofficial. Other authors focus their research on the former communist countries for which the available GDP figures do not capture unofficial activity and statistical offices provide some data that is too hard to compare across countries (Johnson et al., 1997). Therefore, data on total electricity consumption was used. They argue that it is a rough measure of the overall or total economic activity (Johnson et al., 1997). The model assumes that the short-run electricity to GDP elasticity is mostly close to one, there are also some adjustments for energy efficient, neutral and inefficient group of countries. The calculation is as follows:

- Run regressions using unit elasticity assumption;
- Difference between the change in electricity consumption and officially reported GDP is an estimate of the change in shadow economy size.

On the other hand, this method has received various critics and skepticism. For example, Lacko in points out:

- Several inconsistencies between Kaufmann and Kaliberda results and other estimations. The results contradict each other;
- Missing logic for results in some of the countries, for example, their results do not show growth of unofficial economy for Romania and Uzbekistan during the transition period although for other transitional post-soviet economies it increased, which is hypothesis and logics that seem more valid;
- The author also raises a question whether this approach can be used is market economies. For instance, for Finland in the research period GDP decreased, while electricity consumption increased. Therefore, based on the method logics, the hidden economy should have increased although results show the opposite (Lacko, 2000).

Therefore, some developments have been made to the model in order to avoid the deficiencies mentioned before. See the updated model to what can be seen in Figure 9 below. As it whows, the model is adjusted for purchasing power parity, share of agriculture, frequency of months with the need of heating in households, average temperature in January, ratio of other energy sources than electricity and other (Lacko, 2000).

$$\ln ER_{ij} = \alpha_1 \ln C_{ij} + \alpha_2 AG_{ij} + \alpha_3 G_{ij} + \alpha_4 Q_{ij} + \alpha_5 PR_{ij} + \alpha_6 H_{ij} + \alpha_7 \qquad (1)$$
$$\alpha_1 \geq 0,\ \alpha_2 \leq 0,\ \alpha_3 \leq 0, \alpha_4 \leq 0, \alpha_5 \leq 0, \alpha_6 \geq 0$$

where i: country; j: year; ER_{ij}: *per capita* household electricity consumption in country i in year j (kWh); C_{ij}: *per capita* real consumption of households (at purchasing power parity); AG_{ij}: the share of GDP produced in agriculture in total official GDP; G_{ij}: index for weather-differences = the relative frequency of months with the need for heating in houses (under 10°C) multiplied by the average temperature in January; Q_{ij}: the ratio of energy sources other than electric energy to all energy sources in household energy consumption; PR_{ij}: real price of consumption of 1 kWh residential electricity in US dollars (at exchange rate); H_{ij}: *per capita* output of the hidden economy.

Figure 9. Developed electric consumption approach by Lacko (retrieved from (Lacko, 2000))

In addition, in 2003 Feige and Urban tried to replicate the model. The main conclusion was that this model is highly sensitive to alternative initial assumptions (Feige & Urban, 2003). Moreover, they modified the simple electric consumption method (ECM) to include the changes in the electricity prices per unit and changes in the proportions of private and industrial sectors. This was called modified electric consumption method (MEC).

Later, in 2008, the same authors add more critics to this model and rise a question can GDP be trusted meassure (Feige & Urban, 2008). They remind that "that the MEC estimates of the unreported economy are derived as the difference between total economic activity as predicted by the adjusted electric consumption proxy and the actual recorded level of GDP" and if the estimated value of total economic activity is correct and also GDP is precise indicator of total economic activity, then the MEC of estimate should be zero. However, from their results it is not zero, but positive value. From such results it can be concluded that either it is due to overestimated total economic activity or underestimate of non-observed income (Feige & Urban, 2008). As such, it is not easy to determine which of the reasons it is. They suggest to use other methods.

In the previous sections various macro methods were analyzed that use only single indicator of economic activity. It is reasonable to assume that if more indicators were used, the results might be more accurate. Good news that someone has already developed such method. It takes into account several causes and indicators of the hidden economy. It is called the "multiple indicator multiple cause" (MIMIC) model.

MIMIC

Already in the 1980's economists were worried about the fact that all other methods for the determination of unobserved economy's size considers just one indicator whether it is currency, electricity consumption or other. In addition, it is clear that there are various factors that influence it and the other models do not include those, which may cause over or under estimation. Not only the number of indicators is the problem, but also the causes are not taken into account (Frey & Weck, 1983).

Therefore, in 1884 a research paper which simultaneously considers the multiple causes for the existence and increase of shadow economy and the implications of it (Frey & Weck-Hanneman, 1984). It builds on the statistical theory of unobserved variables. Frey and Weck-Hanneman were one of the first one who developed this method, of course, it was later improved by others. Let's look into more details about this model developed by Frey and Weck-Hanneman in 1984. The two main considerations and mathematical aspects:

1) LISREL (linear interdependent structural relationship) is the base of this calculation and it is generalization of the MIMIC model developed in 1975 (Jöreskog & Goldberger, 1975). However, the initial LISREL was first introduced by the same authors in 1972 (Jőreskog & van Thiilo, 1972). The model is described in highly complex mathematical research paper on 72 pages; however, in simple terms it tries to estimate unknow coefficients in a set of linear structural equations. The general form of the equation system consists of latent variables which are unmeasured and unknown and various measured variables (Jőreskog & van Thiilo, 1972). This is visually interpreted in Figure 10 below. In their case there is one unobserved variable which is the size of shadow economy and many determinants for it.

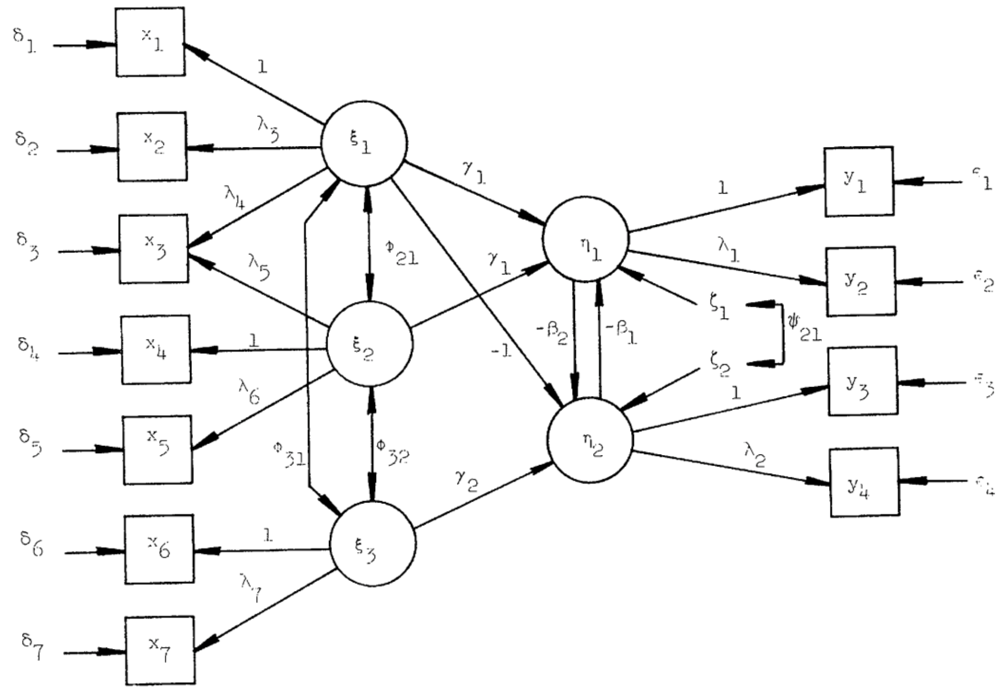

Figure 10. LISREL model graphical representation (retrieved from (Jőreskog & van Thiilo, 1972))

Note. Circles denote true variables; squares denote observed variables. the other variables are residuals or error variables, one way arrow denotes a direct causal influence, a double arrow denotes correlation or covariation without causal interpretation

2) The determinants are selected by the author and it is assumed that those are calculated/estimated without error. In the research paper from 1984, four were used: the burden from the government, tax morality, rate of unemployment and development of national economy. Later the statistically significant variables that are normalized to sum up to unity to interpret results in terms of wights are used to determine the hidden economy (Frey & Weck-Hanneman, 1984). Figure 11 shows the results graphically.

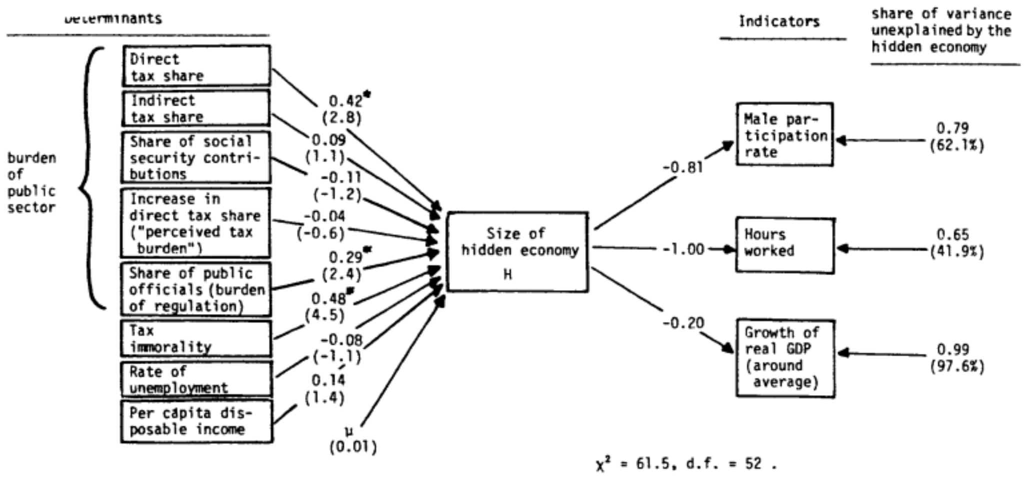

Figure 11. MIMIC model estimated results on the hidden economy estimation in OECD countries from 1960-1978 (retrieved from (Frey & Weck-Hanneman, 1984))

Note. Numbers on the arrows are parameter estimates, in parentheses is corresponfing t-statistics.

In 1999, the model was applied to data from New Zealand. In this research paper the author uses more causal variables than in the initial model, such as unemployment, income, economic regulation degree, applicable tax legislation, effect of inflation and tax burden (Giles, 1999). Furthermore, he added three indicators' variables or effects, for example, the real GDP growth rate, proportion of currency to M3 and male participation in the economy as percentage. According to Schneider (2000) Giles had created dynamic MIMIC model as he used time-series index of the hidden economy and then estimated "cash-demand-model" in order to be able to interpret the index into percentages. Significantly, Giles paid the needed attention to the raw data analysis and various statistical assumptions, such as stationarity and other (Giles, 1999).

Notably, Friedrich Schneider is a name that must be mentioned in the research of shadow economy since he is the author of various significant papers and books about this topic. Firstly, very interesting and all the methods and results gathering paper from 2000; however, no new methods are introduced there (Schneider & Enste, 2000). Secondly, in 2010 he published a paper "New Estimates for the Shadow Economies all over the World". As it was described in his previous research paper, there are various methods, data sets, assumptions, time periods researched which makes it impossible to compare the results. Therefore, this paper does it by trying to estimate the shadow economy in 162 countries in the period 1999 to 2006 using

MIMIC model in order to have comparable data points (Schneider et al., 2010). It is essential in the context of this master thesis to pay attention to a few important details that Schneider had implemented in his model in contrast to previous authors:

- Due to large sample of countries those are separated in categories since for each different set of causal and indicator variables are selected in the model. The country groups were as follows: developing countries, Eastern Europe and Central Asia (mostly also known as transitional economies or former Soviet union), high income OECD countries (Schneider et al., 2010).
- Causal variables used:
    - Government size;
    - Share of direct taxation;
    - Total tax burden;
    - Fiscal freedom;
    - Business freedom;
    - Economic freedom;
    - Unemployment rate;
    - GDP per capita;
    - Regulatory quality;
    - Openness;
    - Inflation.
- Author used various indicator variables such as
    - Tax burdens;
    - Regulation intensity;
    - Public sector services;
    - Official economy;
    - Monetary indicators;
    - Labor market characteristics-participation rate, growth rate of total labor force;
    - GDP per capita using purchasing power parity rates and growth of GDP per capita rate.

On the contrary, in the same year the same author published results on the shadow economy and undeclared earnings which were estimated using surveys which is from completely different methodology group that was discussed previously (Feld & Schneider, 2010). Nevertheless, the reason for doing it was the usually large difference between both methods estimated results give. The author further discussed the reasons behind it. In Figure 12 (1) to (4) those are listed.

Various kinds of shadow economy activities/values	Shadow economy in % of official GDP	Shadow economy in (billion €)	% share of the overall shadow economy
(1) Survey method: shadow economy activities (based on 'black' hours worked)	5.0–6.0	117–140	33–40
(2) Material (used)	3.0–4.0	70–90	20–25
(3) Illegal activities (goods and services)	4.0–5.0	90–117	25–33
(4) Already in the official GDP included illegal activities	1.0–2.0	23–45	7–13
Sum (1)–(4)	13.0–17.0	300–392	85–111
Overall (total) shadow economy (estimated by the MIMIC and calibrated by the currency demand procedure)	15.0	340	100

Figure 12. Comparison of the German shadow economy in 2006 using survey and MIMIC method (retrieved from (Feld & Schneider, 2010))

In 2014, the reasoning is supplemented with more potential causal variables (Schneider, 2014). This shows evidence that these factors indeed explain the variance and both methods can be compared. It is so fascinating that even the same author acknowledges advantages and disadvantages for both of the methods direct and indirect and to some extent proves that both, if properly used, can show similar and comparable results.

Now that the most notable indirect methods have been shortly discussed and described, the advantages and disadvantages are summarized in table 2.3.2.

2.3.2 table

Advantages and disadvantages of indirect methods

Advantages	Disadvantages/limitations
Usually is less time consuming and less resources are needed (people, money)	Some transactions are not related to income generations (e.g. savings), but are used in the model *(in the currency approach)*

Flexibility on the selection of causes, indicators (in MIMIC method), can also be disadvantage since can be subjective	Not all transactions in the shadow economy are conducted with cash, but method assumes it *(in the currency approach)*
Can be applied to compare the sizes of shadow economies across many countries or periods of time	Differences across countries, regions in the elasticity of electricity/GDP *(in the electric consumption method)*
	Very simplified and often unrealistic assumptions
	It is difficult to know what components of shadow activity are captured by the method (Buehn and Schneider, 2013)
	Not stable since changes in assumptions can cause significantly different results
	Higher electricity prices reduce consumption per unit of output and it does not consider the energy efficiency differences (*(in the electric consumption method)*
	MIMIC is sensitive to currency demand assumptions which may not always be correct

Now the methods and available data has been discussed one can reflect on the importance of this topic. In this thesis the author will try to estimate the size of shadow economy in the Baltcis using MIMIC model and compare the results and investigate reasons for differences between Baltic countries. The author will also provide suggestions to policymakers about the current effectiveness and future developments of the existing policies.

3. MODEL AND DATA USED

3.1. Model and variables

Based on the conclusions reached in the previous section the MIMIC model will be used in order to perform data-based estimation of the size of shadow economy in the three Baltic countries: Latvia, Lithuania and Estonia in the period from 2000-2023.

If simplified the model can be defined with formulae as follows:

$$y = \Lambda_y \eta + \varepsilon, \tag{3.1}$$

where $\Lambda_y -$ vector or matrix of results how strongly each indicator reflects the latent variable;
$y -$ vector of observed indicators;
$\eta -$latent variable, in this case the shadow economy;
$\varepsilon -$ measurement error.

Formula 3.1 is the measurement model which links the latent variable or unobserved variable (shadow economy) to indicators.

$$\eta = \Gamma x + \zeta, \tag{3.2}$$

where $\Gamma -$vector or matrix of regression coefficients how each cause (x) affects η;
$x -$cause variables;
$\zeta -$ the error term capturing influences on η not explained by x.

Formula 3.2 shows how the causes can be linked to latent variable.

However, it must be noted that the model provides standardized scores that can be used in comparative analysis of trends or changes, if absolute values are needed external anchor values should be used (Chaudhuri et al., 2016).

What's more to add, in 2008, Buehn and Schneider created an updated EMIMIC model which stands for error correction MIMIC model that is based on the equation 3.3:

$$\Delta y_t = \gamma \Delta x_t + \lambda u_{t-1} + w_t, \tag{3.3}$$

where $w_t -$white noise term;

$u_{t-1} -$one period lagged equilibrium error (error correction);

Δy_t *and* $\Delta x_t -$first differences.

The idea is the two cointegrated I (1) variables for which we know their first difference can be used as an error correcting variable. This allows to keep the long-run relationship in the model and corrects it for the deviations from their long-term trends (Buehn & Schneider, 2008). In further sections, it will be analyzed if this model could be applied.

Next, the determinant and indicator variables should be selected that will be used in the model. The author has selected those that have been used by other researchers in the past to avoid bias that might arise if introducing new variables. As the determinant variables tax burden, GDP per capita, unemployment rate, government size and regulatory quality will be used. On the other hand, as the indicators are selected growth rate of GDP per capita, the labor force participation rate.

3.2. Procedure

Below are listed the steps executed to estimate the size of undeclared economy:

1) Raw data collection from publicly available data sources.
    a. World Bank Open Data (https://data.worldbank.org/)
    b. OECD Data Explorer (https://data-explorer.oecd.org/?lc=en)
2) Data upload to R studio software.
3) Data formatting to applicable format for further steps.
4) Run descriptive statistics of the data.
5) Statistical tests for stationarity, homoscedasticity and other.
6) Raw data is of different measurement units. Therefore, the results are scaled by adjusting the values to a fixed interval.
7) Install and load "*laavan*" package.
8) Execute the model.
9) Calibrate the estimated results to absolute values (this is done because the initial results are standardized scores expressed as relative change).
10) Analyze results.
11) Make conclusions.

4. DATA ANALYSIS

4.1. Determinants

Figure 13 shows the stable growth and development of the GDP per capita in the Baltics over the 24-year period. GDP per capita is selected due to it representing the total magnitude and volume of official economic activities. The changes in these variable influences shadow economy. It is also used as determinant in almost every similar research. Similar, trend is observable in all three countries. They experienced stagnation in 2008-2008 due to World Economic crisis of 2008 which was caused by the housing crisis and bubble. The crisis very heavily influenced the overall economic situation, policy makers and daily life of citizens, significantly increased immigration. Another downfall can be seen around 2021, this was caused by the Covid-19 pandemic containment measures. During economic slowdown or stagnation, the magnitude of unrecorded economic activities increases (Nguyen et al., 2020). This proved especially during the covid crisis.

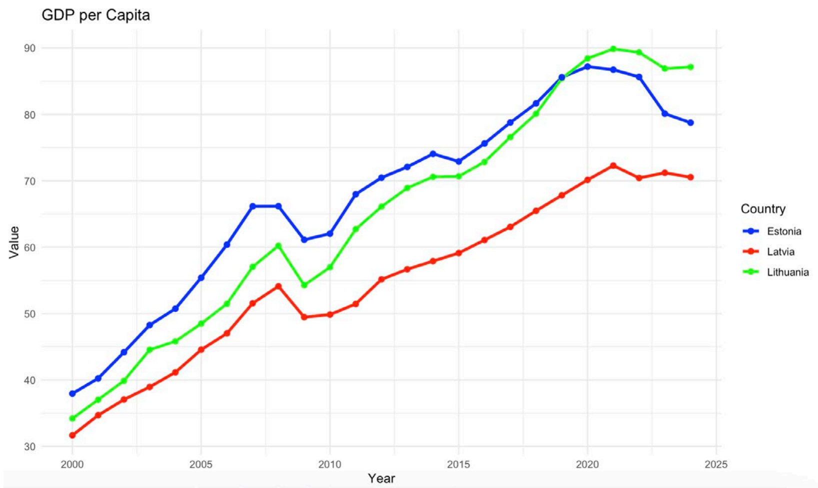

Figure 13. GDP per capita in the Baltics 2000-2024.

Dat retrieved from: World Bank Open Data. In the public domain. Visualized by author.

Another indicator variable that will be used in the model is government size in this case government final consumption expenditure expressed as percentage of GDP (*World Bank Open Data*, n.d.-a). This has been selected due to previous research showing that tax burdens are one of the main causes for shadow economy and size of government spending is a good and comparable representation of it (Schneider, 2011). The more tax income, the more government can afford to spend. Also, in Figure 14 similar trends for all countries are seen; however, larger gaps between the data points of each country. Significant, increase around 2010 is explained with the various types of government support needed to cure the consequences of housing crisis, similar increase in 2021 due to covid support to individuals and businesses.

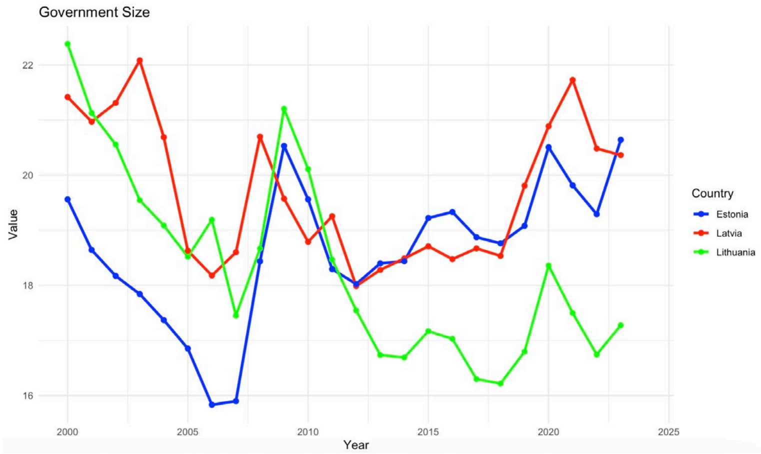

Figure 14. Government size of the Baltics 2000-2023.

Data retrieved from: World Bank Open Data. In the public domain. Visualized by author.

In contrast, Figure 15 visualizes the tax burden over years. Tax burden in this case is total tax revenue expressed as percentage of GDP (*OECD Data Explorer • Comparative Tables of Revenue Statistics in OECD Member Countries*, n.d.). This comes from similar group as the government size and higher tax burden might lead to higher shadow economy. For Latvia and Estonia increasing trend can be seen. On the contrary, for Lithuania 2012 shows very different results and significant decrease, later the ration managed to grow and outperform Latvia.

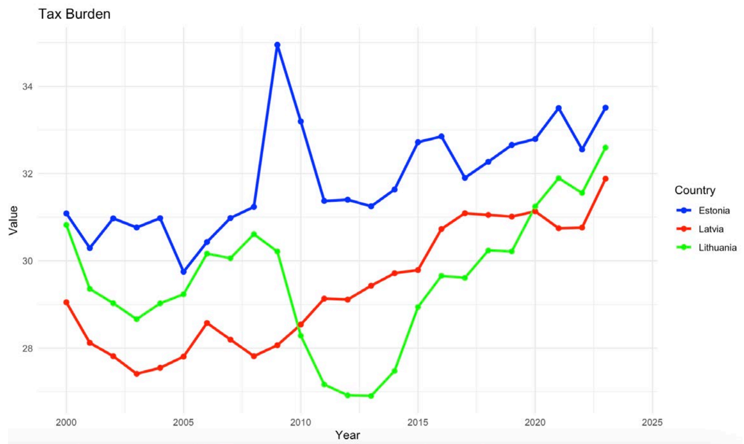

Figure 15. Tax burden in the Baltics 2000-2023.

Data retrieved from: OECD Data Explorer In the public domain. Visualized by author.

Moving on to the unemployment which is the proportion of people aged 15 or more expressed as a percentage of the labor force. This rate is used since it represent the size of official economy which is needed in order to estimate the unofficial economy. Similar trend for Latvia, Lithuania and Estonia. Most notable increase in 2010 due to world economic crisis of 2008.

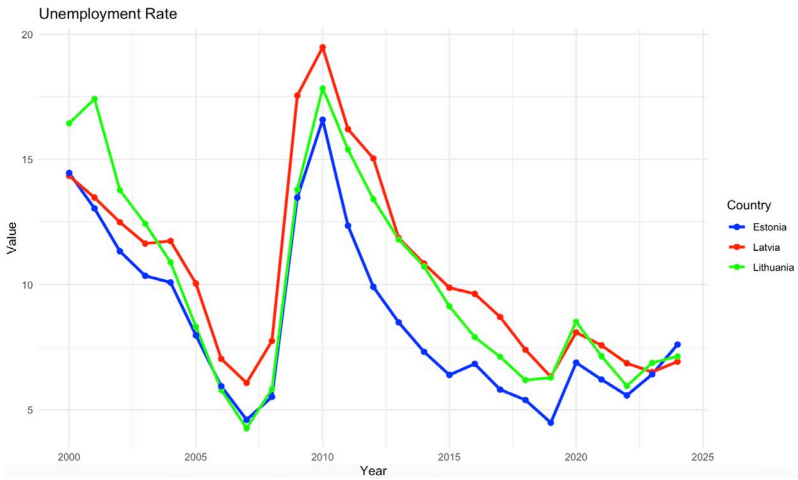

Figure 16. Unemployment rate in the Baltics 2000-2024.

Data retrieved from: OECD Data Explorer. In the public domain. Visualized by author.

Finally, regulatory quality is essential indicator that impacts the size of shadow economy. In various scenarios the quality of the public factor services and regulations is crucial factor for people's decision to work or not work in the shadows. This indicator is World Bank's estimated index which measure the market-unfriendly policies, it tries to capture the ability of the government to develop and implement sound policies that advocate private sector further development (*World Bank Open Data*, n.d.-b). It must be noted that this estimate gives each country a score in a range from -2.5 to +2.5. Conversely, regulatory quality indicator is the highest in Estonia and the worse in Latvia during the whole period, with only a few exemptions.

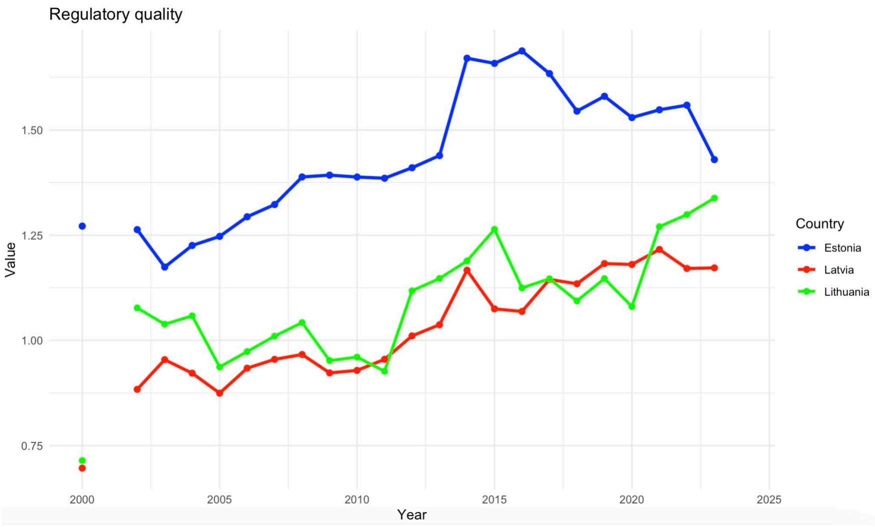

Figure 17. Regulatory quality index in the Baltics 2000-2023.

Data retrieved from: World Bank Open Data. In the public domain. Visualized by author.

4.2. Indicators or effects

Now that we are analyzing the data sets of indicator variables, in Figure 18 GDP per capita can be seen. Similar pattern to other indicators such that major downfall of the growth rate in 2008-2009 due to financial crisis. It reached negative growth rate of -15%. During the analyzed period the growth for countries differs and the standings of the countries also differ; however, at the end of 2023 Latvia has the highest growth rate of 1.8%.

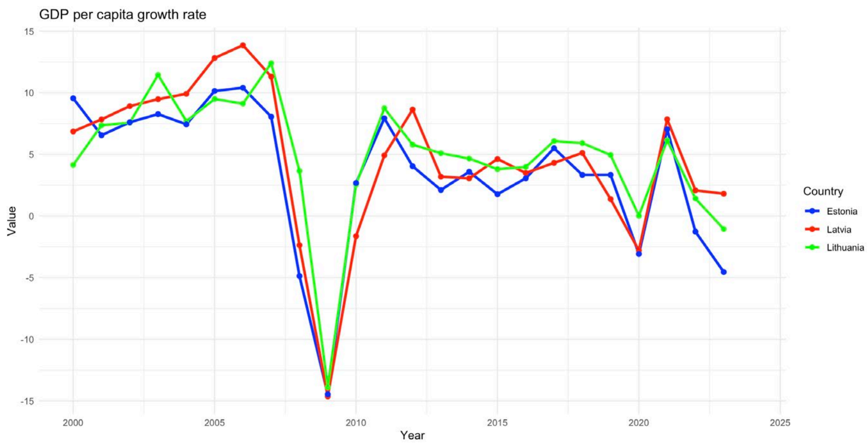

Figure 18. GDP per capita growth rate in the Baltics 2000-2023.

Data retrieved from: World Bank Open Data. In the public domain. Visualized by author.

In addition, the labor participation rate is essential indicator. It is the proportion of population in the age from 16 to 64 that is economically active. Similar growth trend observed in all three countries, reaching 76.4%, 78.3% and 81.2% in Latvia, Lithuania and Estonia respectively in 2024. It must be taken into consideration that the population size is decreasing in the Baltics, which have impact on the ratio.

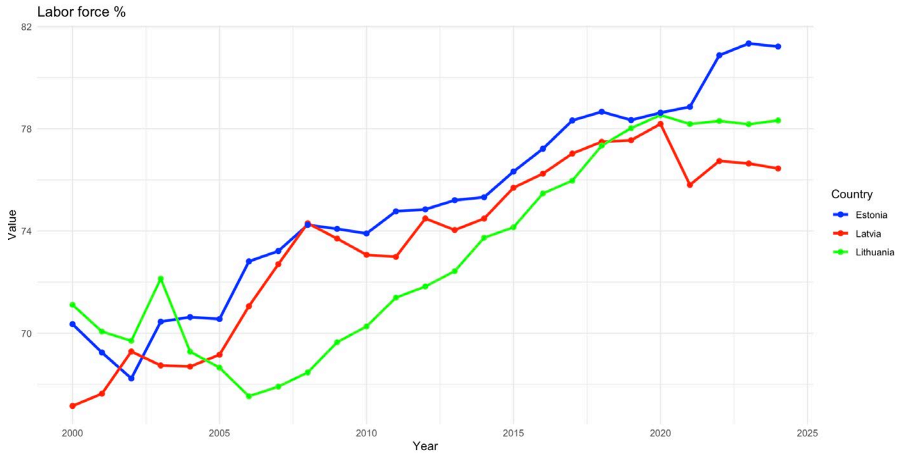

Figure 19. Labor force participation rate in the Baltics 2000-2024.

Data retrieved from: World Bank Open Data. In the public domain. Visualized by author.

4.3. Stationarity

Most of the research papers on macroeconomic data face the problem non-stationary time series. Logically, authors who previously worked with MIMIC model faced the problem. Already in 1983 was introduced Dynamic MIMIC model which introduced updated model that used first-differenced data (Aigner et al., 1986).

In the sections 4.1 and 4.2 the various data points or in this case time series were plotted. As expected for the macroeconomic data, increasing trend can be observed. It indicates non-stationarity since variable influence another with time lag and the mean and variance is increasing over time. Nevertheless, it must be checked with statistical tests. One of the most commonly used tests was developed by David Dickey and Wayne Fuller which is known as augmented Dickey-Fuller test (ADF) (Dickey & Fuller, 1979). The ADF test evaluates the null hypothesis that a unit root is present, indicating non-stationarity. A low p-value, typically less than 0.05, leads to the rejection of the null hypothesis, suggesting that the series is stationary. Conversely, a high p-value implies that the series is non-stationary and may require transformation, such as differencing, to achieve stationarity (Pemberton, 2011). Nowadays various programming applications have this test built in, in R studio “tseries” package and function “adf.test” is used. In the table 4.3.1. below results of the ADF test are summarized.

4.3.1 table

The Augmented Dickey-Fuller test p-value summary of raw data

Variable	p-value		
	Latvia	Lithuania	Estonia
GDPpc	0.3806	0.2894	0.8521
GM_size	0.7243	0.439	0.2664
tax_burden	0.1201	0.7849	0.1514
Unemployment	0.4547	0.426	0.4
regulatory	0.2955	0.06785	0.9284
GDP_growth_ind	0.264	0.2671	0.2904
labor_force_ind	0.8415	0.9001	0.3175

Note. Significance levels *p< .05, **p< .01, ***p< .001.

The statistical test results approve what was observed in the data plots. Since all of the p-values exceed the significance level of 0.05, the null-hypothesis is rejected and the alternative should be accepted, which indicates that selected time series are non-stationary.

Next, as was suggested by the test authors in order to avoid the problems that arises if non-stationary data is used in further analysis, it must be differenced (Dickey & Fuller, 1979). In simple terms, the changes between two values will be used rather than absolute values. However, by doing this in order to avoid non-stationarity problem the long-term relationship can be lost (Aigner et al., 1986). Nevertheless, first differentiation was performed, the results continued to show non-stationarity, therefore, second order differentiation was performed. The results of p-value are summarized in table 4.3.2 below. The results are unsatisfactory for further model implementation since the non-stationarity could not be eliminated even after various differentiations to reasonable confidence level. For example, for some the p-value varies around 0.4 which is high value, for some it is in the range of 0.11 to 0.16 and for Estonia GDP per capita and tax burden and labor force participation rate series it is below 0.09. It means that there is not sufficient evidence to conclude that these time series fulfill the stationarity requirement. It must be added that using the second order difference in the model would make very complex to interpret (Hyndman & Athanasopoulos, 2018).

4.3.2 table

The Augmented Dickey-Fuller test p-value summary of differenced data

Variable	First order difference			Second order difference		
	Latvia	**Lithuania**	**Estonia**	**Latvia**	**Lithuania**	**Estonia**
GDPpc	0.1485	0.198	0.2759	0.028*	0.0357*	0.0867
GM_size	0.01*	0.1662	0.2129	0.0212*	0.0336*	0.0334*
tax_burden	0.0562*	0.4129	0.1532	0.1688	0.3835	0.0869
Unemployment	0.1735	0.3975	0.2326	0.0589*	0.1199	0.0478*
regulatory	0.1959	0.2307	0.4495	0.0283*	0.1153	0.4078
GDP_growth_ind	0.0403*	0.0289*	0.0337*	0.0123*	0.01**	0.01**
labor_force_ind	0.0686	0.5841	0.2717	0.019*	0.01**	0.0679

Note. Significance levels *p< .05, **p< .01, ***p< .001.

It can be also concluded that most of the time series are not co-integrated I (1) or even I (2), since as can be seen from the p-values in table 4.3.2 the non-stationarity is not eliminated after various differentiation tries.

However, the fact that the time series are integrated to an order higher than one, do not fulfill the requirement of con-integrated at I (1) in order to apply EMIMIC (Buehn & Schneider, 2008).

These results indicate that there might be other problem with the time series data selected for testing. It might be that there is a trend or structural breaks. When once again looking at the graphs this cannot be excluded. Nevertheless, it must be tested with additional statistical tests.

4.4. Structural breaks and level shifts

4.4.1. Latvia

In previous section used ADF test is one of the simplest unit root tests. In the 1990s Zivot and Andrews argued the traditional opinion that current shocks or crisis have just a temporary effect on the macroeconomic events and generally they do not have long-term impact and influence. They challenged this view and developed more advanced test called Zivot-Andrews test (Andrews & Zivot, 1992). H_0: unit-root process without exogenous structural breaks and H_A: trend-stationary process with possible structural change occurring at an unknown point in time.

Also, this test is already built into R studio. Library “urca” and function ur.za is used, model specification “both” (Pfaff, 2004). See results for each country summarized in the tables below.

Table 4.4.1 shows that based on the test results it can be concluded that tax burden, unemployment and regulatory quality in Latvia are stationary with a level break that is significant since the test statistic is lower than critical values.

4.4.1 table

Zivot-Andrew’s test result summary for Latvia

Variable	Intercept	y.l1	Trend	DU	DT	Significance	Break Point
GDPpc	26.285 ***	0.1489	2.4840 ***	-5.8123 ***	-0.9733 **	Significant at 5%	8
GM_size	8.08 ·	0.54 **	0.83	-2.35 *	-0.77	Not Significant	3
tax_burden	19.97 ***	0.26 ·	0.15 ***	0.79 *	-0.10 ·	Significant at 5%	15
Unemployment	12.49 ***	0.14	-0.89 *	10.18 ***	0.1	Significant at 1%	8
regulatory	0.88 ***	-0.001	0.009 *	0.10 **	0.0004	Significant at 1%	13
GDP_growth_ind	6.84	0.14	0.61	-12.84 *	-0.23	Not Significant	7
labor_force_ind	40.78 **	0.42 *	-0.21	2.35 *	0.37	Not Significant	5

Significance levels: ·$p < 0.1$; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. DU-level break, DT-trend, y.l1-irst lag of the dependent variable.

Although Zivot-Andrew’s test provides the potential break point. It was decided to use “Fstats” function from “strucchange” package available on R studio. Plotting the F-statistic provides valuable insights that will help to determine the break point that will be introduced to the model. Syntax used in R studio to create the graphs:

```
install.packages("strucchange")
library(strucchange)
gdppc_lv<-ts(GDPpc$Latvia, start = 2000, end = 2023, frequency = 1)
plot(Fstats(gdppc_lv ~ 1), main ="GDP per capita LV")
```

Figure 20 shows how the F-statistic evolves over time, with the red horizontal line representing the critical threshold. Peaks in the black line suggest possible structural breaks, particularly notable around 2008 to 2010 for GDP per capita, tax burden, and unemployment, likely reflecting the global financial crisis. For regulatory indicators, the potential break appears slightly later, around 2013–2015. These plots visually suggest periods where significant shifts in underlying dynamics may have occurred.

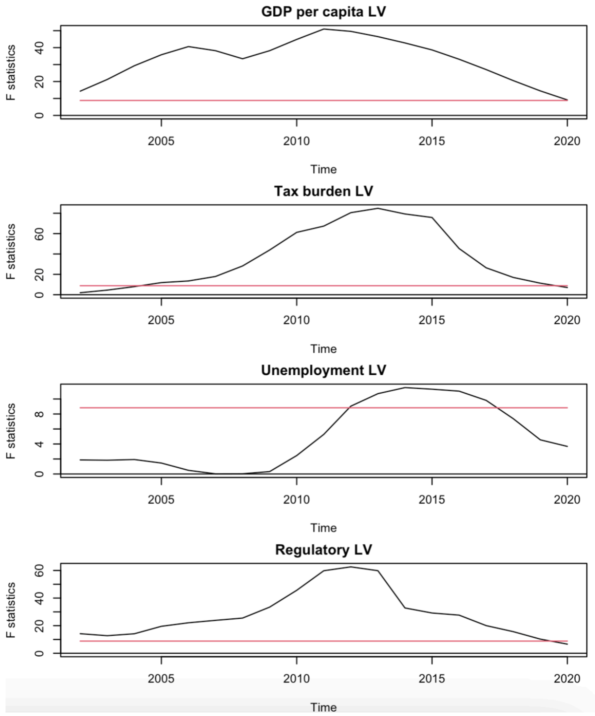

Figure 20. Plot of F-statistics for significant variables for Latvia.

In order to fight with the level break that is revealed a dummy variable will be introduced. This would solve the problem because the series are stationary around the broken trend or level not mean.

Now based on each figure seen in the previous section a break will be introduced and later introduced in the MIMIC model. In order to not overfit or overcomplify the model which is built on short time series only one break point will be introduced. For Latvia the most reasonable break point looks to be around 2010, at this time all of the used variable F-statistics reaches the highest point. Also, the Zivot-Andrews test showed the break point around the 10th period.

4.4.2. Lithuania

Table 4.4.2 summarizes the results for Lithuania. Results are similar to the results to Latvia. Here tax burden, unemployment and regulatory quality and labor force participation rate are stationary time series with significant level shift.

4.4.2 table

Zivot-Andrew's test result summary for Lithuania

Variable	Intercept	y.l1	Trend	DU	DT	Significance	Break Point
GDPpc	25.02 **	0.36 ·	1.46 **	6.75 *	-1.90 *	Not Significant	18
GM_size	13.22 **	0.3	0.01	-1.79 *	0.03	Not Significant	11
tax_burden	16.67 ***	0.39 ***	0.22 **	-3.10 ***	0.09	Significant at 1%	9
Unemployment	11.37 **	0.27 ·	-1.08 *	9.22 ***	0.48	Significant at 5%	8
regulatory	1.01 ***	0.07	-0.013 ·	0.16 *	0.024 *	Significant at 1%	11
GDP_growth_ind	15.70 **	-0.001	-1.69 *	8.78 ·	1.22	Not Significant	10
labor_force_ind	27.29 ***	0.51 ***	3.15 **	-4.25 ***	-2.80 **	Significant at 1%	3

Significance levels: $\cdot p < 0.1$; $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. DU-level shift, DT-trend or level break, y.l1-irst lag of the dependent variable.

Continuing with Figure 21 of the F-statistics for Lithuania. Tax burden and labor force participation show clear peaks around 2016–2018, suggesting significant shifts in those variables during that period. Unemployment shows a possible break in the mid-2000s, then stabilizes, reflecting recovery from early economic transitions. Regulatory indicators display a moderate peak around 2012–2014, indicating a potential policy shift or institutional reform. Together, these plots highlight periods of economic transformation in Lithuania, especially in the mid to late-2010s.

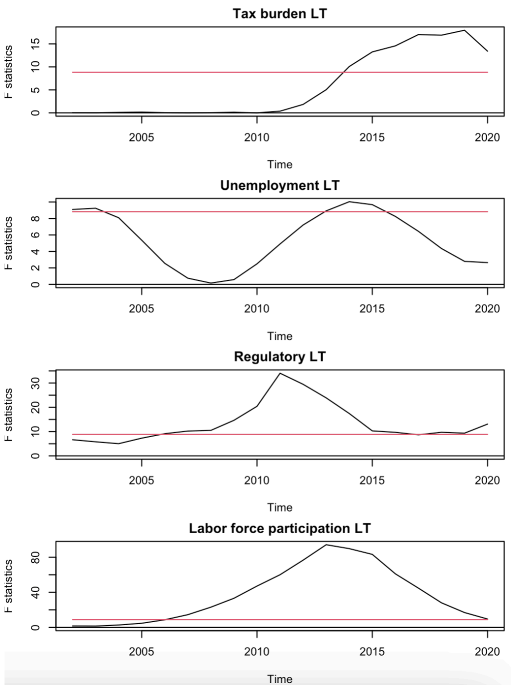

Figure 21. Plot of F-statistics for significant variables for Lithuania.

If we look at the figure above, then it can be seen that the patterns are very similar for all significant variables. In contrast to Latvia, the break appears to be later, so the break point is determined to be 2012.

4.4.3. Estonia

Unfortunately, as shows table 4.4.3 for Estonia data we can conclude on just one cause and one indicator time series that those are stationary with a structural break.

4.4.3 table

Zivot-Andrew's test result summary for Estonia

Variable	Intercept	y.l1	Trend	DU	DT	Significance	Break Point
GDPpc	21.75 **	0.56 **	0.83 *	2.62	-5.74 ·	Not Significant	21
GM_size	15.95 ***	0.16	-0.40 ·	2.53 **	0.46 ·	Not Significant	7
tax_burden	24.11 **	0.17	0.34 *	-1.74 *	-0.19	Not Significant	10
Unemployment	9.79 *	0.24	-0.81 ·	6.89 **	0.31	Not Significant	8
regulatory	1.02 ***	0.12	0.020 ***	0.24 ***	-0.042 ***	Significant at 1%	13
GDP_growth_ind	15.23 *	0.01	-1.89 *	9.95 ·	1.35	Not Significant	10
labor_force_ind	76.63 ***	-0.14	0.93 ***	-1.54 **	-0.32 **	Significant at 1%	9

Significance levels: ·p < 0.1; *p < 0.05, **p < 0.01, ***p < 0.001. DU-level break, DT-trend, y.l1-irst lag of the dependent variable.

This Figure 22 presents F-statistics for Estonia's tax burden and labor force participation. Both series show strong structural break signals peaking around 2008–2010, likely tied to the global financial crisis. The F-statistics decline afterward, suggesting relative stability in the following decade.

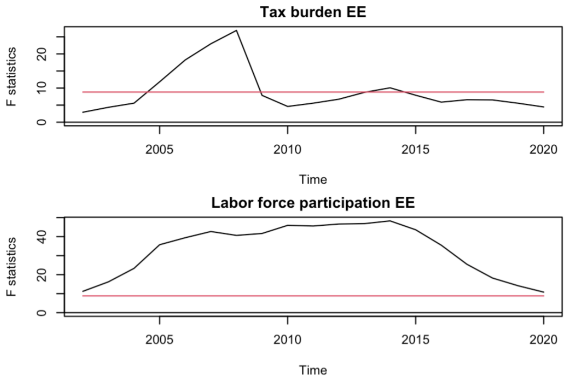

Figure 22. Plot of F-statistics for significant variables for Estonia.

For Estonia the break point is selected to be 2008.

5. MODEL APPLIED

In this section the most interesting part of the thesis will be described. Which variables will be determined to be significant and what will be the final model for the shadow economy size determination in the Baltics.

5.1. Latvia

As it was explained previously “lavaan” package in R studio is used. All variables which were assessed as valid for further implementation in the model are used. Therefore, the model specification for Latvia looks like this:

```
mimic_model_lv <- '
  # Measurement model: latent variable measured by indicators
  shadow =~ latvia_GDP_growth_1diff

  # Structural model: latent variable explained by causes
  shadow ~ GDPpc + tax_burden + Unemployment + break2010
'
## This variable is now 0 before 2010 and 1 for 2010 onward.
latvia_df_scaled$break2010 <- ifelse(latvia_df_scaled$Year >= 2010, 1, 0)

# Fit the model
fit_lv <- sem(mimic_model_lv, data = latvia_df_scaled)

# Summary of results
summary(fit_lv, standardized = TRUE, fit.measures = TRUE)
```

The summary of the model output is in the table below:

5.1.1 table

MIMIC model results for Latvia

Parameter	Estimate	Std. Error	z-value	p-value	Standardized Estimate
GDPpc	-6.041	3.051	-1.98	0.048 **	-0.983
tax_burden	0.555	3.648	0.152	0.879	0.09
Unemployment	-0.915	1.805	-0.507	0.612	-0.149
break2010	12.425	5.594	2.221	0.026 **	0.996
Latent Variance (shadow)	23.751	7.161	3.317	0.001 ***	0.659

Significance levels: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Based on the output it looks that only GDP per capita and dummy variable break are significant since their p-values are below 0.05. Also, the signs align with general understanding. Shadow economy in Latvia decreases as GDP rises. In addition, a clear structural increase happened after 2010 mostly due to 2008 financial crisis since the impact happen with a time lag. Overall, the model is significant since the p-value is below 0.001. This leads to conclusion that this latent construct is valid.

The Akaike (AIC) criteria of this model is 142.12, which is the higher if compared to other 2 models indicating that based on these criteria this model is worse fit than those.

Finally, now that significant model is created question arises how to interpret it and what is the size of shadow economy in Latvia in order to answer the research question of this thesis. As it was mentioned before the MIMIC model will not provide absolute values but rather relative changes and in order to convert the latent variable, in this case, the shadow economy size, into meaningful units' calibration must be performed. In other words, use external reference points of the variables size and calibrate using the built model.

These are the steps performed:

- Extracted the significant estimate coefficients from country's MIMIC model;
- Vouched the scaled variable values from previously created data frames;
- Added reference points from known research that is considered reliable and is takes as reference points in various Latvian government policies (Schneider, 2022);
- Calculated the latent scores for all of the years in the research period;
- Using package "forecast" function "forecast" forecast the value for next 5 periods.

Here are the results for Latvia.

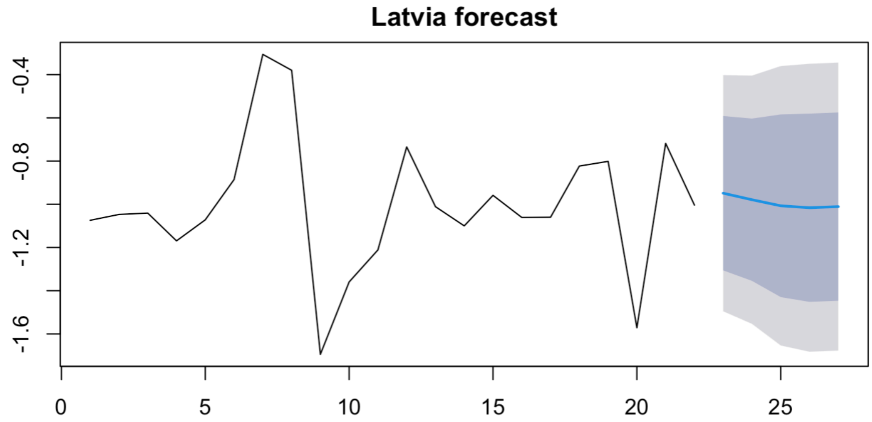

Figure 23. Shadow economy magnitude in Latvia estimated by model.

Note. X axis represent % change in percentage of GDP compared to previous period, Y axis represent time 0 is 2000.

Due to fact that for Latvia the indicator was used as differenced value also the reference was used as such. The result is also expressed as change in shadow economy as percentage of GDP compared to previous period. The reference point was used 2018 and the estimated value for 2023 is cumulative of the changes since then which translated to absolute values estimates that the size of shadow economy in 2023 in Latvia is 15.32%.

Using the “forecast” library a prediction for the next 5 years is made. It must be noted that in the predictions for all 3 countries the predicted values vary around the mean. Most importantly the function predicts that the decreasing trend will continue and after 5 years the cumulative effect will be -4.9% if calculated from the 2023. This looks like very optimistic prediction that should be considered cautiously. The 95% confidence interval estimated that the value of shadow economy in 2028 falls between -8% and -1.8% in 2028. To conclude, in the Hi95 case recalculated to absolute values it would be 13.46% of GDP in 2028 in Latvia.

5.2. Lithuania

Slightly different is the model for Lithuania:

```
mimic_model_lt <- '
  # Measurement model: latent variable measured by indicators
  shadowlt =~ labor_force_ind

  # Structural model: latent variable explained by causes
  shadowlt ~ tax_burden + Unemployment + regulatory + break2012
'
lithuania_df_scaled$break2012 <- ifelse(lithuania_df_scaled$Year >= 2012, 1, 0)
```

```
fit_lt <- sem(mimic_model_lt, data = lithuania_df_scaled)
summary(fit_lt, standardized = TRUE, fit.measures = TRUE)
```

5.2.1 table

MIMIC model results for Lithuania

Parameter	Estimate	Std. Error	z-value	p-value	Standardized Estimate
tax_burden	0.443	0.093	4.75	0.000 ***	0.443
Unemployment	0.259	0.105	2.458	0.014*	0.259
regulatory	0.042	0.13	0.323	0.747	0.042
break2012	1.692	0.237	7.124	0.000 ***	0.864
Latent Variance (shadowlt)	0.132	0.039	3.391	0.001 **	0.138

Significance levels: *p < 0.05, **p < 0.01, ***p < 0.001.

The MIMIC model for Lithuania tries to estimate the shadow economy in Lithuania as the indicator variable using labor force participation proportion. Tax burden, unemployment and break in 2012 are statistically significant variables, while regulatory quality is not. The model shows relationship that increases in tax burden and unemployment are associated with expansion of the shadow economy in Lithuania. Additionally, a significant structural increase in observed after 2012. The latent construct demonstrates significant variability, validating the model's ability to capture meaningful changes in informal economic behavior. In other words, the model is significant. The AIC criteria of this model is 28.74, which is much lower than for Latvia, indicating that this is the best model fit if compared to other models.

Similarly, the same steps are performed on Lithuania data points and model in order to arrive at the size of the shadow economy. The estimated shadow economy size in 2023 in Lithuania is 22.22% of its GDP. Refer discussion section for detailed implications of the results.

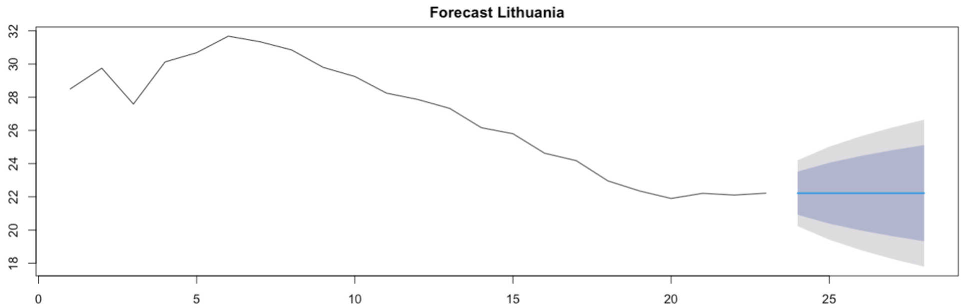

Figure 24. Shadow economy magnitude in Lithuania estimated by model.

Note. X axis represent percentage of GDP, Y axis represent time 0 is 2000.

The prediction was made also for Lithuania. Similar observations ad for Latvia, the prediction is mean model which predicts the same average value going forward. This is mostly due to short and comparingly flat time series. It must be noted that the prediction was made of the calculated shadow economy size which indeed does not show significant trends in long period that the function could use in the prediction. This might be a research limitation and/or future research opportunity; more details are commented in the discussion section of the thesis.

5.2.2 table

Prediction for Lithuania

Year	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2024	22.22	20.92	23.51	20.24	24.20
2025	22.22	20.39	24.05	19.42	25.02
2026	22.22	19.97	24.46	18.79	25.65
2027	22.22	19.63	24.81	18.26	26.18
2028	22.22	19.32	25.11	17.79	26.64

Note. Results from the auto.arima function used on estimated shadow economy percentages for Lithuania.

The 80% confidence interval says that the value will fall between 19.32 and 25.11% in 2028. The 95% confidence interval predicts the value to be between 17.79% and 26.64%.

5.3. Estonia

Finally, the very simple model for Estonia:

```
mimic_model_ee <- '
  # Measurement model: latent variable measured by indicators
  shadowee =~ labor_force_ind

  # Structural model: latent variable explained by causes
  shadowee ~ regulatory + break2008
'
estonia_df_scaled$break2008 <- ifelse(estonia_df_scaled$Year >= 2008, 1, 0)
fit_ee <- sem(mimic_model_ee, data = estonia_df_scaled)
summary(fit_ee, standardized = TRUE, fit.measures = TRUE)
```

5.3.1 table

MIMIC model results for Estonia

Parameter	Estimate	Std. Error	z-value	p-value	Standardized Estimate
regulatory	0.402	0.188	2.134	0.033*	0.402
break2008	1.012	0.401	2.527	0.011*	0.476
Latent Variance (shadowee)	0.296	0.087	3.391	0.001**	0.31

Significance levels: *p < 0.05, **p < 0.01, ***p < 0.001.

Due to non-stationarity and other problems with the raw data of Estonia very limited number of variables were included in the model. However, also this model is statistically valid. The AIC criteria of this model is 43.3, it is higher than for Lithuania (28.74), but lower than for Latvia (142.2). If looking into more detail the model identified regulatory quality a significant driver for the growth of shadow economy. Meaning that worse regulatory quality increases shadow economy. Moreover, significant structural increase has followed after 2008.

Now if we look on Estonia. The estimated shadow economy size in 2023 in Estonia is 19.77% of its GDP

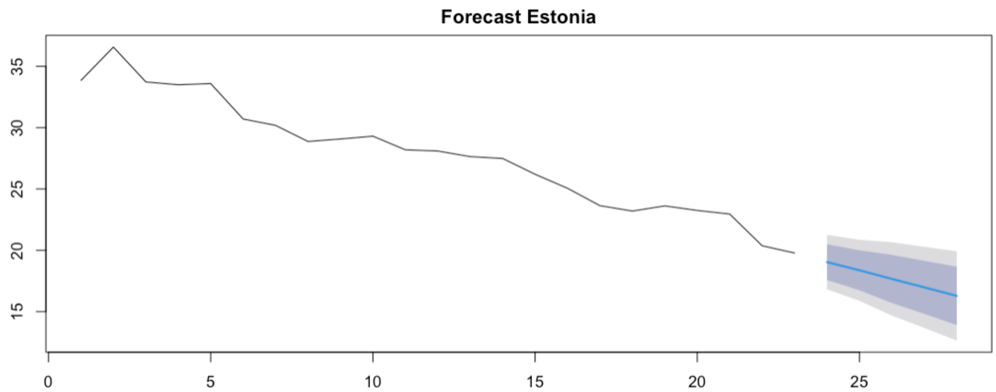

Figure 25. Shadow economy magnitude in Estonia estimated by model.

Note. X axis represent percentage of GDP, Y axis represent time 0 is 2000.

The table presented below 5.3.2 shows the estimated results for Estonia.

5.3.2 table

Prediction for Estonia

Year	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2024	19.04	17.59	20.49	16.82	21.25
2025	18.37	16.74	19.99	15.88	20.86

2026	17.66	15.71	19.62	14.67	20.66
2027	16.98	14.82	19.14	13.67	20.28
2028	16.28	13.90	18.66	12.64	19.92

Note. Results from the auto.arima function used on estimated shadow economy percentages for Estonia.

Interestingly, for Estonia the point forecast works better since the point forecast is different than just the average values. The 80% confidence interval estimate the value to between 13.90% and 18.66%, while the 90% interval prediction is it will be between 12.64% and 19.92% in 2028.

In summary, the forecasted values for 2028 for all of the counties would continue to decrease and in the negative scenario with 95% confidence will be 13.46% of GDP in Latvia, 26.64% of GDP in Estonia and 19.92% of GDP in Estonia.

6. RESULT DISCUSION

In previous sections results for each country were shown separately. Nevertheless, it is essential to see the results in context of each other.

6.1 table

Result overview

Year	Latvia	Lithuania	Estonia
2002	36.33	29.75	36.57
2003	35.28	27.59	33.73
2004	34.24	30.13	33.50
2005	33.07	30.68	33.59
2006	32.00	31.68	30.71
2007	31.12	31.34	30.19
2008	30.81	30.85	28.87
2009	30.43	29.80	29.08
2010	28.74	29.25	29.31
2011	27.38	28.25	28.19
2012	26.17	27.86	28.10
2013	25.43	27.33	27.64
2014	24.42	26.16	27.49
2015	23.32	25.80	26.20
2016	22.36	24.62	25.06
2017	21.30	24.18	23.64
2018	20.24	22.96	23.20
2019	19.42	22.35	23.62
2020	18.62	21.90	23.25
2021	17.04	22.21	22.96
2022	16.33	22.10	20.37
2023	15.32	22.22	19.78

Note. Results of the model. Represent size of the shadow economy in each country expressed as percentages of their GDPs.

For better understanding the results are plotted in Figure 26. Results estimated by this model show decreasing trend in the period from 2000-2023 for all countries. In the 2000s the estimated size of shadow economy varied around 35% from the GDP in Latvia and Estonia and around 30% for Lithuania. Nowadays, the ranking of countries has changed. The model has estimated that the undeclared economy in Latvia is the lowest and varies around 15 percent, in the second place is Estonia with 20 percent and Lithuania with 22 percent. Notably, very similar trend was estimated by Schneider in 2022, refer to Figure 1 on page 8. However, this is logical since this research was used as reference point for the model developed in this thesis. In contrast, Schneider's model calculated that there was an increase in the unofficial economy after 2020 which is probably due to Covid-19 pandemic crisis. My model does not predict it. This is due to fact, that as discussed in the previous sections only one level break of 2008 crisis was included in the model.

No significant spikes or drops can be observed in the graph for Latvia, it looks like linear decrease. For Lithuania the starting point is estimated to be lower yet around 2006 it reaches the other country level. For Estonia the values are most changing in the end finishing in the second place.

There are various reasons for such clear pattern. One of those being the fact that GDP has significantly increased during the research period, the undeclared economy part is hard to conclude on since it can be either decreased or remained on the same level and the decrease is due to growth of official economy. Another reason is the fact that Baltic countries are comparingly new countries and in the 2000s just started to adapt to capitalism after the exit from socialism due to collapse of the Soviet Union in the 1990s. As was discussed in the analytical review of available information the historical background has strong impact on both the macroeconomic indicators and society mindset, habits and attitudes which do not change fast. It takes time to adapt and understand the individual and collective benefits of official and declared economic activities.

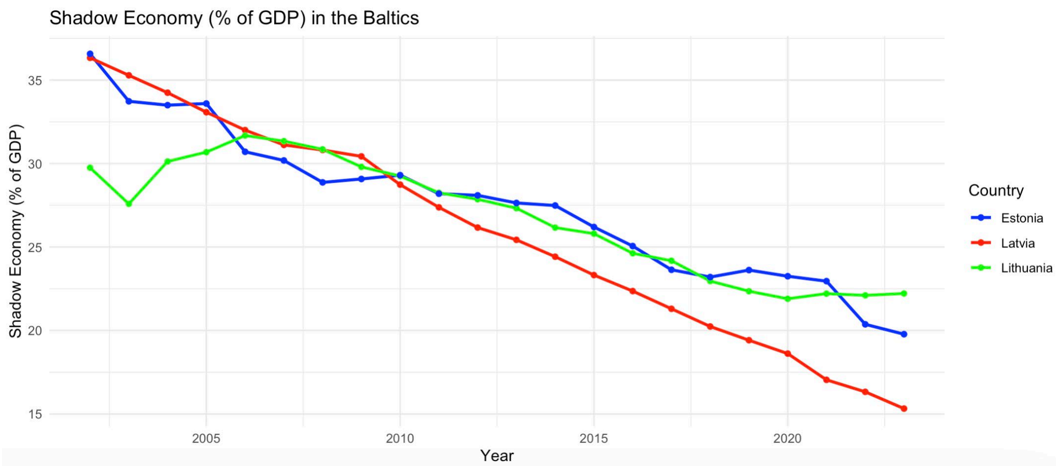

Figure 26. Shadow economy in the Baltics estimated by the model.

6.1. Theoretical and practical implications

The results reconcile with the already existing theoretical framework. Framework of MIMIC model already applied by other scientists' states that various macroeconomic indicators are used as causes and effects and the model estimates the size of shadow economy. In addition, this paper used selection of those and tried to estimate the shadow economy in the Baltics. There are a few theoretical implications that might be used in further research in connection with the estimation model. For example:

- Nowadays both the Dynamic MIMIC and error correcting MIMIC models are becoming more complex to apply to real life raw time series data due to non-stationarity and various structural breaks or level shifts in the period. This means that either the model needs to be updated or not limited number of macroeconomic indicators can be used. This paper in detailed described the steps and measures taken in order to create a mode that is statistically significant.
- It can be observed that more the time series data contain significant breaks in the data due to various crisis. If previously it was just the 2008 financial crisis, now there is history of the 2020 covid-19 pandemic which significantly influenced the financial markets and macroeconomic environment. Both of these shocks were in the research period of this thesis. Additionally, as time goes there are new crisis and extraordinary events that influence macroeconomics that the model should account for. The war in Ukraine has significant impact on the government spending limits especially in Eastern Europe which are Baltic countries.

Nevertheless, there are implications from the results of the estimated shadow economy size. One can say that decreasing trend means that the situation is good and with time its size will reach zero level. This is not the case since shadow economy leaves long-term impact on private lives and on the country as whole. Some of the impacts being unearned tax revenue by the government, unearned personal income and social taxes that decrease the pension fund for current seniors and decrease the official future pension for the current employees that work in the shadows currently. Finally, it may seem that the personal benefit exceeds the one if economic activity would have been declared; however, this influences the collective benefits that the people receive from the government. The quality of roads, educational system, healthcare system and availability, support when on sick or maternity/paternity leaves and other support. Therefore, the governments of the Baltic counties should take stricter steps and measures in order to decrease the shadow economy.

6.2. Recommendations

Some of the recommendations would be:

- The developed models for 3 countries showed that the break variable was significant. Therefore, the governments should build policies and resilience to address and buffer the potential future economic shocks since model show that during those the shadow economy increases.
- Model for Latvia indicated that GDP per capita is significant, therefore, the logical recommendation is to work on increasing it which would lead to decreased unofficial economy. My recommendation would be to focus on productivity increase and encourage digitalization since these are actions that might not require new resources just more efficient usage of existing ones.
- Lithuanian model presented significant variables tax burden and unemployment. For both of these variables there is a strong positive relationship with the shadow economy. In other words, higher tax burden and unemployment increase the unofficial economic activities in Lithuania. Consequently, the suggestion is to revisit tax politics and rates and assess its reasonableness. In addition, develop and improve programs for unemployed people to reduce the unemployment.
- Increase regulatory quality. Although this World Bank score was included in the Lithuanian and Estonia model and it was statistically significant only in Estonia, it is essential factor that influences why the members of society would participate in the shadow economy.

6.3. Research limitations

It must be noted that this research paper has various limitations such as small sample size due to limited macroeconomic data availability on more frequent basis than yearly and historical background of the Baltic countries which is why macroeconomic data is available only starting from the 1990s.

Building statistically significant model on macroeconomic time series data is very complex. In addition, the research period from 2000 to 2023 is very specific since during this period two crisis happened – the financial crisis of 2008 and covid-19 pandemic in 2020. Due to this the time series contain at least one or more shocks, shifts and breaks which can be negative effect when trying to build model. In connection to this, only one dummy variable for level shifts was introduced for each country although potentially there were more of those. The decision was made in order to not overcomlify the model.

When applying MIMIC model, it is extremely complex to avoid subjectivity since due to time and other resource limitations only a few indicator and causal variables could be selected. The author had tried to justify the choice yet complete random selection was not performed.

Another aspect is that the results are strongly dependent on other researcher estimated results in order to determine my model's results. The results could change if different reference point would be taken. In this thesis paper relevant and trustworthy author's work is used yet errors or imprecisions could exist. Another question is how acceptable it is to use reference point that was determined using the same method, but different model.

Limitations in this research lies in the forecasting component, specifically the use of the auto.arima function for predicting future shadow economy trends based on the MIMIC model outputs. Although ARIMA models are widely used for time series forecasting, their effectiveness depends heavily on the stationarity and linear characteristics of the underlying data. In this case, the predicted values – derived from standardized latent variables and first-differenced indicators may not capture complex nonlinear trends or structural breaks in the data. Additionally, the automatic model selection in auto.arima occasionally led to overly simplistic models, for example, ARIMA(0,0,0) with non-zero mean, resulting in flat or unrealistic forecasts.

These limitations reduce the reliability of long-term projections and suggest the need for more robust forecasting techniques or enriched datasets to improve predictive accuracy. Maybe prediction should have been performed for the raw data and then these values could have been used in the model.

Finally, the results from each country's developed model have been presented. The results of all countries together have been discussed and some recommendations and research limitations have been noted. The final conclusions if and to what extent the research tasks were fulfilled and hypothesis approved or rejected can be found in the conclusions section on the next page.

CONCLUSIONS

The current situation assessment and existing theoretical framework with regards to shadow economy has been analyzed and summarized by the author. This interesting and detailed assessment has been focused on the Baltics states. The region is especially interesting due to its historical background and other specific factors such as trust to the government and willingness to pay taxes.

At the beginning of the thesis various tasks were identified:

- to develop a method/model to estimate the shadow economy (as a share of GDP);
- assess the size of the shadow economy in Latvia, compare the results with other available studies and analyze possible reasons for discrepancies;
- forecast shadow economy trends for the next five years;
- provide recommendations for reducing the shadow economy.

The steps were executed and the thesis hypothesis is confirmed since the developed MIMIC model for each of the Baltic countries is consistent and scalable and can be used in other or further research to estimate the size of shadow economy is specific countries or regions using no complex programing software R studio and publicly available macroeconomic data.

In short, for each of the analyzed country the structural break has been determined as significant variable indicating that serious economic shocks have influenced the macroeconomic data in the research period.

The model for Latvia indicates that GDP per capita has significant negative influence on the shadow economy in Latvia-if GDP per capita decreases, shadow economy activity increases. The estimated size of it is 15% of the GDP in 2023. If compared to Lithuania and Estonia, this is the lowest percentage.

In contrast, the Lithuanian model estimated that tax burden, unemployment has significant positive impact on the unofficial economy. If these indicators increase, also the shadows increase. Model estimated the magnitude of shadow economy in Lithuania to be 22%.

For Estonia due fact that for various time series the author was not able to avoid stationarity problem, a single cause variable of regulatory quality was used. It is significant and has positive influence on the researched topic. Surprisingly, the shadow economy in Estonia was estimated to be around 26% in 2023 which is the highest among remaining other countries.

Nevertheless, there are various future research opportunities to expand this research by selecting more causal and indicator variables, creating various models and selecting those with higher significance or other criteria that is determined to be applicable.

REFERENCES

Aigner, D., Schneider, F., & Ghosh, D. (1986). *ME AND MY SHADOW: ESTIMATING THE SIZE OF THE U.S. UNDERGROUND ECONOMY FROM TIME SERIES DATA* (M8615). https://doi.org/10.22004/ag.econ.296050

Alexeev, M. (2001). *A Note on Measuring the Unofficial Economy in the Former Soviet Republics*.

Andrews, D., & Zivot, E. (1992). *Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis*. https://doi.org/10.2307/1391541

Aptauja: Latvijā 69% iedzīvotāju uzskata, ka nodokļu sistēma pret viņiem nav godīga. (2023, August 30). https://jauns.lv. https://jauns.lv/raksts/zinas/574474-aptauja-latvija-69-iedzivotaju-uzskata-ka-nodoklu-sistema-pret-viniem-nav-godiga

Bhattacharyya, D. K. (1999). On the Economic Rationale of Estimating the Hidden Economy. *The Economic Journal*, *109*(456), 348–359. https://doi.org/10.1111/1468-0297.00438

Buehn, A., & Schneider, F. G. (2008). MIMIC Models, Cointegration and Error Correction: An Application to the French Shadow Economy. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1135882

Chaudhuri, K., Kumbhakar, S. C., & Sundaram, L. (2016). Estimation of firm performance from a MIMIC model. *European Journal of Operational Research*, *255*(1), 298–307. https://doi.org/10.1016/j.ejor.2016.05.005

Dean, R. (2006). *The Value of Humanity in Kant's Moral Theory*. Clarendon Press.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American*

Statistical Association, *74*(366a), 427–431. https://doi.org/10.1080/01621459.1979.10482531

Ēnu ekonomikas ierobežošanas plāns 2024.–2027. Gadam. (2024). LIKUMI.LV. https://likumi.lv/doc.php?id=349352

Ēnu ekonomikas indekss 2023: Latvija starp plāniem un realitāti | Stockholm School of Economics in Riga. (n.d.). Retrieved May 20, 2025, from https://www.sseriga.edu/enu-ekonomikas-indekss-2023-latvija-starp-planiem-un-realitati

European Commission. Directorate General for Economic and Financial Affairs. (2023). *2023 country report: Lithuania.* Publications Office. https://data.europa.eu/doi/10.2765/986341

European Commission & European Parliament. (2009). *Beyond GDP: Measuring progress, true wealth and the well-being of nations*. Publications Office of the European Union. https://data.europa.eu/doi/10.2779/54600

Feige, E. L. (1979). How Big Is the Irregular Economy? *Challenge*. https://www.tandfonline.com/doi/abs/10.1080/05775132.1979.11470559

Feige, E. L. (2015). Reflections on the Meaning and Measurement of Unobserved Economies: What Do We Really Know About the "Shadow Economy"? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2705879

Feige, E. L., & Urban, I. (2003). *Estimating the Size and Growth of Unrecorded Economic Activity in Transition Countries: A Re-evaluation of Electric Consumption Method Estimates and their Implications* [Working Paper]. http://deepblue.lib.umich.edu/handle/2027.42/40022

Feige, E. L., & Urban, I. (2008). Measuring underground (unobserved, non-observed, unrecorded) economies in transition countries: Can we trust GDP? *Journal of*

Comparative Economics, *36*(2), 287–306. https://doi.org/10.1016/j.jce.2008.02.003

Feinstein, J. S. (1991). An Econometric Analysis of Income Tax Evasion and Its Detection. *The RAND Journal of Economics*, *22*(1), 14. https://doi.org/10.2307/2601005

Feld, L. P., & Schneider, F. (2010). Survey on the Shadow Economy and Undeclared Earnings in OECD Countries. *German Economic Review*, *11*(2), 109–149. https://doi.org/10.1111/j.1468-0475.2010.00509.x

Finanšu Ministrija. (2022). *Pētījumi par ēnu ekonomikas apjomu | Finanšu ministrija*. https://www.fm.gov.lv/lv/petijumi-par-enu-ekonomikas-apjomu

Frey, B. S., & Weck, H. (1983). What Produces a Hidden Economy? An International Cross Section Analysis. *Southern Economic Journal*, *49*(3), 822. https://doi.org/10.2307/1058719

Frey, B. S., & Weck-Hanneman, H. (1984). The hidden economy as an 'unobserved' variable. *European Economic Review*, *26*(1), 33–53. https://doi.org/10.1016/0014-2921(84)90020-5

Giles, D. E. A. (1999). Modelling the Hidden Economy in the Tax-Gap in New Zealand. *ResearchGate*. https://doi.org/10.1007/s001810050076

Grudule, M., & Kalnačs, B. (2024). Soviet Latvia and Exile: Political Changes in the Aftermath of WWII and Their Impact on Latvian Literary Histories. In L. Steinby, B. Kalnačs, M. Oshukov, & V. Parente-Čapková (Eds.), *The Politics of Literary History: Literary Historiography in Russia, Latvia, the Czech Republic and Finland after 1990* (pp. 173–195). Springer International Publishing. https://doi.org/10.1007/978-3-031-18724-7_8

Gutmann, P. M. (1977). The Subterranean Economy. *Financial Analysts Journal*, *33*(6), 26–34. https://doi.org/10.2469/faj.v33.n6.26

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice*. OTexts.

Iedzīvotāju skaits pēc tautības un pa vecuma grupām gada sākumā 2011—2024. (n.d.). Oficiālās statistikas portāls. Retrieved May 20, 2025, from https://data.stat.gov.lv:443/pxweb/lv/OSP_PUB/START__POP__IR__IRE/IRE040/

Johnson, S., & Kaufmann, D. (2001). 10 Institutions and the Underground Economy. In *A Decade of Transition*. International Monetary Fund. https://www.elibrary.imf.org/display/book/9781589060135/ch10.xml

Johnson, S., Kaufmann, D., Shleifer, A., Goldman, M. I., & Weitzman, M. L. (1997). The Unofficial Economy in Transition. *Brookings Papers on Economic Activity*, *1997*(2), 159. https://doi.org/10.2307/2534688

Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable. *Journal of the American Statistical Association*, *70*(351a), 631–639. https://doi.org/10.1080/01621459.1975.10482485

Jőreskog, K. G., & van Thiilo, M. (1972). Lisrel a General Computer Program for Estimating a Linear Structural Equation System Involving Multiple Indicators of Unmeasured Variables. *ETS Research Bulletin Series*, *1972*(2), i–71. https://doi.org/10.1002/j.2333-8504.1972.tb00827.x

Kasekamp, A. (2017). *A History of the Baltic States*. Bloomsbury Publishing.

Kaufmann, D., & Kaliberda, A. (1996). Integrating the Unofficial Economy into the Dynamics of Post-Socialist Economies: A Framework of Analysis and

Evidence. In *The International Politics of Eurasia: V. 8: Economic Transition in Russia and the New States of Eurasia*. Routledge.

Kazemier, B., & van Eck, R. (1992). Survey investigations of the hidden economy: Some methodological results. *Journal of Economic Psychology*, *13*(4), 569–587. https://doi.org/10.1016/0167-4870(92)90012-V

KNAB. (2023). *Attieksme pret korupciju Latvijā*. https://www.knab.gov.lv/lv/media/4803/download?attachment

Lacko, M. (2000). Hidden economy—An unknown quantity? *Economics of Transition*, *8*(1), 117. https://doi.org/10.1111/1468-0351.00038

Lemieux, T., Fortin, B., & Fréchette, P. (1994). The effect of taxes on labor supply in the underground economy. *The American Economic Review*. https://www.semanticscholar.org/paper/The-effect-of-taxes-on-labor-supply-in-the-economy-Lemieux-Fortin/f82ee44579c8bb30cf5d14d85c9d15b5d801a146

Leonard, M. (1994). *Informal Economic Activity in Belfast*. https://www.semanticscholar.org/paper/Informal-Economic-Activity-in-Belfast-Leonard/039c78f29ced37d05eb66b0944a856a419b9c043

lvportals.lv. (n.d.). *Reģistrēšanās skaistumkopšanas pakalpojumu sniegšanai, prasības un nodokļu pienākumi (I)—LV portāls*. Retrieved March 1, 2025, from https://lvportals.lv/skaidrojumi/367348-registresanas-skaistumkopsanas-pakalpojumu-sniegsanai-prasibas-un-nodoklu-pienakumi-i-2024

lvportals.lv. (2021). *Paplašina atbalstu par dīkstāvi skaistumkopšanas nozarē strādājošajiem—LV portāls*. https://lvportals.lv/skaidrojumi/324156-paplasina-atbalstu-par-dikstavi-skaistumkopsanas-nozare-stradajosajiem-2021

McCrohan, K. F., Smith, J. D., & Adams, T. K. (1991). The Cyclical Nature of Household Purchases in Informal Markets. *European Journal of Marketing*, *25*(7), 22–40. https://doi.org/10.1108/EUM0000000000620

Measuring the economy: GDP and NIPAs. (n.d.). ResearchGate. Retrieved February 16, 2025, from https://www.researchgate.net/publication/326190805_Measuring_the_economy_GDP_and_NIPAs

Nguyen, C. P., Schinckus, C., & Thanh, D. S. (2020). Economic fluctuations and the shadow economy: A global study. *Global Economy Journal*, *20*(03), 2050015. https://doi.org/10.1142/S2194565920500153

OECD. (2024). *OECD Survey on Drivers of Trust in Public Institutions – 2024 Results*. https://www.oecd.org/en/publications/oecd-survey-on-drivers-of-trust-in-public-institutions-2024-results_9a20554b-en.html

OECD Data Explorer • Comparative tables of Revenue Statistics in OECD member countries. (n.d.). Retrieved May 3, 2025, from https://data-explorer.oecd.org/vis?lc=en&fs[0]=Topic%2C1%7CTaxation%23TAX%23%7CGlobal%20tax%20revenues%23TAX_GTR%23&fs[1]=Measure%2C0%7CTax%20revenue%23TAX_REV%23&fs[2]=Unit%20of%20measure%2C0%7CPercentage%20of%20GDP%23PT_B1GQ%23&pg=0&fc=Unit%20of%20measure&snb=6&vw=tb&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_REV_COMP_OECD%40DF_RSOECD&df[ag]=OECD.CTP.TPS&df[vs]=1.1&dq=LTU%2BLVA%2BEST..S13._T..PT_B1GQ.A&pd=2000%2C2023&to[TIME_PERIOD]=false

OECD et al. (2002). *Measuring the Non-Observed Economy: A Handbook.* https://doi.org/10.1787/9789264175358-en.

Pemberton, J. (2011). Time Series Analysis with Applications in R, Second edition. *Journal of Applied Statistics*, *38*(6), 1311–1312. https://doi.org/10.1080/02664760903075663

Pētījumi par ēnu ekonomikas apjomu | Finanšu ministrija. (n.d.). Retrieved March 22, 2025, from https://www.fm.gov.lv/lv/petijumi-par-enu-ekonomikas-apjomu

Pfaff, B. (2004). *urca: Unit Root and Cointegration Tests for Time Series Data* (p. 1.3-4) [Dataset]. https://doi.org/10.32614/CRAN.package.urca

Picur, R. D., & Riahi-Belkaoui, A. (2006). The impact of bureaucracy, corruption and tax compliance. *Review of Accounting and Finance*, *5*(2), 174–180. https://doi.org/10.1108/14757700610668985

Postea, M.-M., & Achim, M.-V. (2023). Estimation methods for the shadow economy. A systematic literature review: Métodos de estimativa para a economia subterrânea. Uma revisão sistemática da literatura. *Brazilian Journal of Business*, *5*(3), 1574–1594. https://doi.org/10.34140/bjbv5n3-010

Putniņš, T. J., & Sauka, A. (2015). Measuring the shadow economy using company managers. *Journal of Comparative Economics*, *43*(2), 471–490. https://doi.org/10.1016/j.jce.2014.04.001

Rutkauskas, V. (2019). Tax Morale and Public Sector Revenues in Baltic Countries. *Socialinė Teorija, Empirija, Politika Ir Praktika*, *19*, 26–43. https://doi.org/10.15388/STEPP.2019.9

Sahin, K. (2009). *Measuring the economy: GDP and NIPAs*.

Samuelson, P. A., Nordhaus, W. D., & Froyen, R. T. (2000). *The Gross Domestic Product and the National Income and Product Accounts*.

Sauka, A. (2016a). *ĒNU EKONOMIKA LATVIJAS BŪVNIECĪBAS NOZARĒ*.

Sauka, A. (2016b). *ĒNU EKONOMIKU IETEKMĒJOŠO FAKTORU ANALĪZE LATVIJAS BŪVNIECĪBAS NOZARĒ*.

Sauka, & Putniņš. (2024). *Shadow Economy Index for the Baltic Countries | Stockholm School of Economics in Riga*. https://www.sseriga.edu/shadow-economy-index-baltic-countries

Schneider, F. (2011). *Handbook on the Shadow Economy*. Edward Elgar Publishing.

Schneider, F. (2014). *The Shadow Economy: An Essay*.

Schneider, F. (2022). New COVID-related results for estimating the shadow economy in the global economy in 2021 and 2022. *International Economics and Economic Policy*, *19*(2), 299–313. https://doi.org/10.1007/s10368-022-00537-6

Schneider, F., & Buehn, A. (2013). Estimating the Size of the Shadow Economy: Methods, Problems and Open Questions. *Turkish Economic Review*. https://www.semanticscholar.org/paper/Estimating-the-Size-of-the-Shadow-Economy%3A-Methods%2C-Schneider-Schneider/5717aa6ac1db41a03d3d2b53e056d6433fa73d5e

Schneider, F., Buehn, A., & Montenegro, C. E. (2010). New Estimates for the Shadow Economies all over the World. *International Economic Journal*. https://www.tandfonline.com/doi/abs/10.1080/10168737.2010.525974

Schneider, F., & Enste, D. H. (2000). Shadow Economies: Size, Causes, and Consequences. *Journal of Economic Literature*, *38*(1), 77–114. https://doi.org/10.1257/jel.38.1.77

Shkarupa, E. A., Mytareva, L. A., & Yalmaev, R. A. (2019). Tax Factor of Developing and Intensive Growth of the Shadow Economy. *Proceedings of the International Scientific Conference "Far East Con" (ISCFEC 2018)*.

Proceedings of the International Scientific Conference “Far East Con” (ISCFEC 2018). https://doi.org/10.2991/iscfec-18.2019.296

Survey shows a decline in Estonian residents’ willingness to pay taxes | Estonian Tax and Customs Board. (2024). https://www.emta.ee/en/news/survey-shows-decline-estonian-residents-willingness-pay-taxes

Tanzi, V. (1999). Uses and Abuses of Estimates of the Underground Economy. *The Economic Journal*, *109*(456), F338–F347.

Torgler, B. (2004, February 1). *Tax Morale, Trust and Corruption: Empirical Evidence from Transition Countries*. https://www.semanticscholar.org/paper/Tax-Morale%2C-Trust-and-Corruption%3A-Empirical-from-Torgler/01cf486c2743234b64116b91dc53535b281665df

Vázquez, J. A. (2003, December 1). *“Sizing” the problem of the Hard-to-Tax*. https://www.semanticscholar.org/paper/%22Sizing%22-the-problem-of-the-Hard-to-Tax-V%C3%A1zquez/484c4a5180b2ef2b5029e882e4d8574c58ef5e5e

Warde, A. (1990). Household Work Strategies and Forms of Labour: Conceptual and Empirical Issues. *Work, Employment and Society*, *4*(4), 495–515. https://doi.org/10.1177/0950017090004004002

Williams, C. C., & Horodnic, I. A. (2015). Explaining and tackling the shadow economy in Estonia, Latvia and Lithuania: A tax morale approach. *Baltic Journal of Economics*. https://www.tandfonline.com/doi/abs/10.1080/1406099X.2015.1114714

World Bank Open Data. (n.d.-a). World Bank Open Data. Retrieved May 3, 2025, from https://data.worldbank.org

World Bank Open Data. (n.d.-b). World Bank Open Data. Retrieved May 3, 2025, from https://data.worldbank.org