Master Thesis U.S.E.

Shadow economy determinants of the United States[1]

Universiteit Utrecht

09-08-2021

ABSTRACT

The thesis provides an interdisciplinary literature and empirical research on which determinants have an impact on the size of the shadow economy of the United States in the period of 1991 till 2015 and the size of the impact of these determinants. This is done by applying the Pearson correlation coefficient and a multivariate regression analysis on the data acquired on the variables tax, unemployment rate, GDP per capita, import, export, participation rate, overall government score, regulatory quality, cashless payments, social security contribution and tax morality. It can be concluded that GDP per capita, unemployment rate and tax burden have a significant impact on the size of the shadow economy of the United States. Where tax burden has the largest impact, followed by the unemployment rate and lastly the GDP per capita.

Name: Elise van Rooden

Student number: 5864151

Email: e.vanrooden@students.uu.nl

Supervisors: A. Kaakeh

Master: Financial Management

Acknowledgement

Table of Contents

Introduction

"If it looks like a duck, quacks like a duck, and acts like a duck, then it is a duck—or so the saying goes. But what about an institution that looks like a bank and acts like a bank? Often it is not a bank—it is a shadow bank" (Kodres, 2013).

In 2008, when the financial crisis was largely impacting the whole economy, a large amount of non-bank financial institution dealt with impaired assets as their values dropped sharply. Some examples are NY Fed, which bailed out and rescued the world's largest insurer AIG, to prevent major catastrophes. Another example is the fifth largest investment bank, Bear Stearns, was sold to JP Morgan Chase in 2008. Later that year in 2008, a well-known example of the Lehman Brothers, which were filed for bankruptcy protection. The Lehman brothers were the fourth largest investment bank in the world. After the financial crisis in 2008, the importance of monitoring the financial flows of non-bank financial institutions became a priority to prevent risks and thus prevent activities in the shadow banking world (Konno *et al,* 2013).

Shadow banking is only a small part of the total shadow economy. Enste & Schneider (2000) point out there are many more, which pertain to the shadow economy. An example is criminal activities, which are categorized to the illegal activities, like dealing drugs, trade in stolen goods, prostitution, smuggling, and fraud. Examples of legal activities are tax avoidance and tax evasion. All these different types of activities make it harder to define the shadow economy. Furthermore, the shadow economy is like running water, it adjusts and changes easily as factors influencing the shadow economy change. Therefore, it can be difficult to research and estimate the size of the shadow economy and which determinants are impactful. For example, when taxes change due to adjusted and new regulations set by the government, this changes the size of the shadow economy. It also can cause some shadow activities to grow more compared to other (Enste & Schneider, 2000).

The difficulty of estimating the size of the shadow economy lies with the concept as it cannot be perfectly defined. This causes numerous ways academics can go with researching the shadow economy. Different approaches, data and determinants are used. Kelmanson *et al* (2019) therefore states that the model used in a research depends on the data availability and the research objectives. The research of this thesis is structured and based on previous literature work to expand visions and create proof for statements done on the determinants for the shadow economy. The literature work of Gasparênienè *et al* (2016) provides the basis for the methodology used. However, their research is extended as this thesis provides a combination of determinants used by the research of numerous literature work. The determinants used in this thesis are not only purely economic factors. Social-economic factors as behaviour must also be considered. When taxes increase due to new and stricter regulations set by the government, individuals are within a trade-off. Humans are rational beings and decide if the cost of breaking the law is worth the benefits of entering the shadow economy. For some it is more beneficial, for some it is not. The trade-off exists when entering. When the shadow activities executed by this individual are discovered, they will experience losses and they will be penalized. When not discovered, it might be more beneficial to stay in the shadow economy to create more profits (Medina & Schneider, 2018). It can therefore be stated that strict regulations have a positive relation with the size of the shadow economy. Then why would the government set these regulations when more people get attracted to the shadow economy? Government tries to decrease the amount of these activities with prosecution and punishment; however, some governments do not consider the fact that individuals enter the shadow economy due to overburdening regulations. They should fight the shadow economy by changing tax and social security systems (Enste & Schneider, 2000). It can be seen as a wake-up call for governments.

For example, in California august 2012, the California state decided to introduce legislation against the shadow economy in the United States. In California only, as one of the fifty states belonging to the United States, has a shadow economy ranging from $60 billion till $140 billion annually (Wiseman, 2013). The government should watch out for applying too overburdening regulations. However, introducing of legislation against the shadow economy is needed. Not only in California, but also worldwide. According to the research done by the Financial Stability Board, shadow banking is 80% of the global GDP, has an 90% influence on the outstanding financial assets and influences the industries of the G20 countries (Lemma, 2016). The shadow economy plays a large role in the global economy. Even when the size of the shadow economy compared to the world average is a small fraction, it is important to take it into account, as even this small portion influences the whole economy. For example, in the United States, the shadow economy was estimated at 8.34 as percent of GDP for the period of 1991 till 2015 (Medina & Schneider, 2018). It is still almost 1/10 of the total economic activity. Even though the United States has a small portion of shadow economy, it might therefore be even more interesting to know which determinants are the cause for this small amount. As for every type of category (developing, transition and developed country) has different causes and characteristics which define the shadow economy in that specific type of country.

Building on the previous literature work done, the data availability and the research gap concluded from the literature; the following research question is stated: *Which determinants influence the size of the shadow economy and what is the impact of these determinants for the United States in period of 1991 till 2015?*

This thesis contributes to the contradictions between what determinants influence the size of the shadow economy in an interdisciplinary approach. As seen in research conducted by Gasparênienè *et al* (2016), Kanniainen *et al* (2004), Aigner *et al* (1986), Ion & Alexandru (2009), Medina & Schneider (2018) and Enste & Schneider (2000); many different determinants are used for the empirical analysis, ranging from economic to non-economic. Which determinants needs to be used, mostly depends on the research objectives that must be achieved and therefore also on which method applied. Furthermore, it also depends on what type of country it is, as other factors influence the size of the shadow economy in developing, transitioning and developed countries. For example, in developing countries the main factor for the existence of the shadow economy is the lack for opportunities in the formal sector. As the shadow economy is a broad concept, many types of research objectives, methodologies and determinants can be found. However, from literature research, the determinants taxes, tax morality, social security contribution and regulations have the strongest influence on the size of the shadow economy.

In this thesis the Pearson correlation coefficient and a multivariate regression analysis will be applied to measure which determinants have the highest impact in the United States on the size of the shadow economy. These methods are chosen as these fulfil the characteristics for reaching the research objectives. The variables that will be used conducted from literature research are: Tax, unemployment rate, GDP per capita, import, export, participation rate, overall government score, regulatory quality, cashless payments, social security contribution and tax morality. As stressed by Enste & Schneider (2000), it is also important to take determinants as tax morality, regulatory quality, and the overall government score. These are seen as social-economic variables which can explain the social decision process behind entering the shadow economy of an individual.

For this research, first, the Pearson correlation will be used to identify the independent variables which have a significant correlation with the dependent variable (size of the shadow economy of the United States in percentage of GDP). These variables which have a correlation will then be used in

the multivariate regression analysis. The multivariate regression analysis will be checked for multicollinearity to prevent biased results (Studenmund, 2017).

From the results provided by the empirical research in this thesis can be concluded that the determinants influencing the size of the shadow economy of the United States for the period 1991 till 2015 are GDP per capita, the number of individuals which are unemployed and the taxes. The variable tax has the largest impact on the size of the shadow economy of the United States, followed by the number of unemployed individuals and lastly the GDP per capita. The focus of this research was an interdisciplinary view; however, the social-economic variables resulted in not being significant. This can be explained by the limitation in this thesis in the lack of data of these variables. As well, there were levels of multicollinearity found, which is possibly also caused due to the lack of data. Despite the limitations and implications, the concluding results are in line with the literature found on previous research.

The structure of the thesis is as follows. First, the literature review. In this chapter, first, the concept of the shadow economy is defined to determine the direction of the thesis. It is important to state the definition correctly which applies on the research conducted. In the second section, the areas where the shadow economy has impact on are presented. In the third section, the difference between the developing, transition and developed countries is shown to determine the reasons behind measuring the shadow economy of the United States. Furthermore, this section provides a description of every type of country, where different types of shadow activity are active. This is important to know, to determine which variables play a larger role. Following the literature review is the chapter of the theoretical framework where the determinants of the shadow economy are described which are used for the empirical research. These determinants are based on the section where the areas are described where the shadow economy has impact on. Furthermore, an explanation of the MIMIC model is provided which is used by Medina & Schneider (2018) for estimating the size of the shadow economy. This data will be used for the dependent variable in this thesis. Also, in this section will be further emphasized on why it is beneficial to apply the multivariate regression analysis and Pearson correlation coefficient to meet the research objective of this thesis. The theoretical framework is followed by the methodology, which further emphasizes and explains which type of data is used. Following up are the results, where all findings will be discussed. Lastly the discussion and conclusion of the thesis are presented.

Literature review

Defining the shadow economy

Before starting to analyse the shadow economy of the United States, it is necessary to state a definition of the concept. When reading published papers, it becomes clear that stating a definition for the shadow economy is not as easy as it seems. The subject is very controversial and there are many different ideas on the definition of the shadow economy, empirical methods for estimation and which policies should be used for minimalizing the shadow economy (Enste & Schneider, 2000).

Most literature states that the name 'shadow economy' was created by Paul McCulley. McCulley was a managing director at PIMCO and an economist. He used this term in a financial symposium hosted by the Kansas City Federal Bank (GSP Working group, 2018; Nesvetailova, 2018).

The name, shadow economy, denotes a negative connotation. In literature, the shadow economy is stated as a negative and undesirable event (Fleming *et al*, 2000). Other names used in the literature are the hidden economy, the gray economy, the black economy, the lack economy, the cash

economy, and the informal economy (Medina & Schneider, 2018). All these names have a negative tone. In this thesis, the term shadow economy will be used. With defining it, many research papers depend their definition on which area of the shadow economy they investigate. The shadow economy is known for a broad range of events. Examples are drugs dealing on the streets, an employee who has a second job which is not known by the government and a timber who fixes a chair for a customer who is paid in cash and does not let his earnings be known to the tax collector. These examples show activities which are "off the books" and thus belong under the definition of the shadow economy (Schneider & Enste, 2002).

According to Enste & Schneider (2000), there are two types of activities, legal and illegal activities. These are represented in table 1. As shown in this table, the shadow economy includes many different activities.

Type of Activity	Monetary Transactions		Non Monetary Transactions	
ILLEGAL ACTIVITIES	Trade in stolen goods; drug dealing and manufacturing; prostitution; gambling; smuggling and fraud		Barter: drugs, stolen goods, smuggling etc. Produce or growing drugs for own use. Theft for own use.	
	Tax Evasion	**Tax Avoidance**	**Tax Evasion**	**Tax Avoidance**
LEGAL ACTIVITIES	Unreported income from self-employment; Wages, salaries and assets from unreported work related to legal services and goods	Employee discounts, fringe benefits	Barter of legal services and goods	All do-it-yourself work and neighbor help

Table 1: Categorization of the activities belonging to the shadow economy (Source: Enste & Schneider, 2000).

However, Fleming *et al* (2000) has a different view on defining the shadow economy, as the definition of the shadow economy is not only limited to the activities. First, Fleming states the definition of the shadow economy as the economic activity which is not under the government's view of accounting. This is a very general definition with small details. This definition is based on the principle that the shadow economy is based on economic activity which is not known to the government (definitional approach) and the behavioural reasoning behind entering the shadow economy as employee or employer (behavioural approach) (Fleming *et al,* 2000). These approaches are mostly used separately while both approaches have a large impact on the shadow economy. Therefore, for this thesis a combination of these two approaches will be used to state a definition. Other definitions stated by Fleming *et al* (2000) are the activities which are not considered while measuring the national income accounts. Another definition described for the shadow economy is the unrecorded national income. The methodology belonging to this definition is subtracting the possible national income for a currency with the recorded national income. The possible methodologies which are used on measuring the shadow economy and its usefulness will be discussed later in the thesis.

Another definition found by Fleming *et al* (2000) is that the shadow economy is a market of goods and services (illegal or legal), which are not measured in the official GDP of the country. There are different types of markets within the shadow economy where a definition can be based on, the informal sector, the household sector, the irregular sector, and the criminal sector. The informal

sector are economic activities which avoid costs and are thus not included with the benefits of the law. The household sector is for example production done from home, which is not viewed by the government. The irregular sector is the market where legal products are sold but avoid legal requirements (tax evasion). The criminal sector are products which are produced illegally and avoid interaction with the government. Examples of this can be the production and trade with drugs (Fleming *et al,* 2000).

Referencing back to table 1, this type of differentiating between activities within the shadow economy is also done by Fleming *et al* (2000), however, there is a different focus. Fleming *et al* (2000) focusses on different sectors where shadow activity can arise. Schneider & Enste (2002), first differentiate between legal and illegal activities, then divide it in monetary a non-monetary transactions and lastly, differentiate between tax evasion and avoidance in both monetary and non-monetary situations. The European System of National Accounts and The System of National Accounts both also use a classification to make defining the shadow activities easier. The three categories are: Illegal production (prohibited by the law), Informal production (jobs which are done occasionally) and underground production (not directly observed production by the government) (Ion & Alexandru, 2009). According to Ion & Alexandru (2009), no definition can describe the shadow economy perfectly, as it consists of many different aspects. Therefore, expanding the definition with categorization makes the definition more accurate.

Other way of defining the shadow economy is to define it based on the institution which are active in the shadow economy. According to the International Monetary Fund (IMF), more recent studies use the type of entity which is active in the shadow economy as basis of the definition of the shadow economy. Also, in 2011, the FSB (Financial Stability Board) used the following definition for the shadow economy: "the occurrence of credit intermediation of entities which are not part of the regular banking system" (GSP Working group, 2018). Nesvetailova (2018) shares this view as she stated that the definitions used are different between an academic and politicians. Furthermore, it also depends on the institution. For example, for banks they see shadow entities as hedge funds. Hedge funds see the shadow banks as money funds, money funds see the shadow banks as too-big-too-fail banks. This confusion between banks exists as no clear definitions is/can be stated (Nesvetailova, 2018). Other definitions, like the ones mentioned before, are more focused on markets or instruments used (GSP Working group, 2018). Some examples of markets are the labour market, or the output market as stated by Kelmanson *et al* (2019).

Other definitions stated by literature highlight another important aspect of the shadow economy. The government does not view the activities executed in the shadow economy. For example, Gasparênienè *et al* (2016) defines the shadow economy as: "Shadow economy covers all goods and/or services the income received for which is deliberately hidden from authorities with a view to evading income, VAT or other taxes, social insurance contributions, avoiding compliance to particular legal labour market regulations such as minimal wages, maximal duration of working hours, safety standards, etc.". Enste & Schneider (2000) have a comparable vision as definition for the shadow economy, as they also focus on the income produced by goods and services. They state: "Shadow economy includes unreported income from the production of legal goods and services, either from monetary or barter transactions - hence all economic activities which would generally be taxable were they reported to the tax authorities.".

Other definition may take a broader perspective on the shadow economy. For example, Medina & Schneider (2018) state the following definition: "The shadow economy includes all economic activities which are hidden from official authorities for monetary (avoiding taxes and social security contributions), regulatory (avoiding governmental bureaucracy or burden of regulatory framework),

and institutional reasons (corruption law, weak rule of law and quality institution).". This view is shared by Awasthi & Engelschalk (2018) as they define it very broadly, however they state it more defined compared to Medina & Schneider (2018): "The shadow economy includes all market-based legal production of goods and services that are deliberately concealed from public authorities to avoid payment of income, value added or other taxes; to avoid payment of social security contributions; having to meet certain legal labour market standards, such as minimum wages, maximum working hours, safety standards, etc.; and complying with certain administrative procedures, such as completing statistical questionnaires or administrative forms."

There are thus many ways to define the shadow economy. This is realistic as many factors influence the size and the reasons why firms, employees or employers enter the shadow economy. Fleming *et al* (2000) states that one of the most important factors are the institutional rules. This view is shared with Kelmanson *et al* (2019) as labour and product market regulation make workers and firms enter the shadow economy as these regulations are of a too large scale. Furthermore, other reasons for entering are the amount of taxes and the size of the amount of pension and social security contribution. Some might enter the shadow economy as it provides employment and thus financial safety.

Concluding, many definitions can be found on the shadow economy. According to Ion & Alexandru (2009), no definition can correctly scientifically define the shadow economy due to the many aspects connected to the term. For this thesis, the definition needs to propose a broad definition as the research considers many aspects of the shadow economy. Not only the economic factors are of importance but also the behavioural aspect behind entering the shadow economy. This is important as this research has an interdisciplinary view as mentioned before in the introduction. Some literature work copied the definition of other work without analysing the broader perspective of the shadow economy. As own research is done, it is of importance to construct a definition fitting for the research done by yourself which is fitting for the constructed research question. Therefore, the definition for the shadow economy I used for the thesis will be based on these characteristics which will match the research and view of this thesis. This thesis is a broad research and tries to take all the aspects in to account. No specific definitions on institution, activities, categories, or sectors will be used. This resulted in the following definition**:** *The shadow economy covers all economic activity, legal and illegal, which individuals try to conceal from the government. This behaviour is caused due economic and social-economic factors like a high tax burden, high social security contribution, overburdening regulations, and weak institutional quality.*

Why study the shadow economy?
Now that a definition is stated, it is important to know why you should study the shadow economy. As the shadow economy creates risks for banks, institutions, and the overall economy, it is necessary to monitor and regulate the shadow economy (GSP Working group, 2018). Some academics get attracted by the fact that it is hard to measure and investigate. This can be substantiated by the following quote of Enste & Schneider (2000): "scientific passion for knowing the unknown.". This view is shared by myself as I have the same interest in analysing the unknown. Some might say it is a way to make it very difficult for yourself, however, I disagree because the academics world is about challenging the challenges. It is thus a combination of the eagerness to investigate something very difficult and to prevent large damage to the overall economy, by enlarging the knowledge about the subject.

The shadow economy is actually a product of inefficiencies of the overall economy (Kelmanson *et al,* 2019). Often, these inefficiencies are connected to each other which creates a negative cycle or spiral and thus damage the overall economy. An example of why individuals enter the shadow economy can be the large amount of regulatory burden. They prefer to exit the formal market and enter the shadow economy (Enste & Schneider, 2000). As more people enter the shadow economy, less taxes will be paid and thus the public revenues are lower. These revenues are used for public goods as social security contribution, education, and training programmes. As these public goods and services are lower and of less quality, more individuals doubt the effectiveness of the government. Individuals are less willing to pay taxes and enter the shadow economy. This is a spiral which continues (Kelmanson *et al,* 2019; Schneider & Enste, 2002).

Another type of spiral is with focus on taxes. To maintain the public goods as mentioned before, an amount of taxes is needed to keep the welfare state intact. However, as taxes or government regulations increase, more individuals enter the shadow economy. This results in more pressure on the government revenues which creates less investing and lower quality of public goods. Resulting from this, taxes will be raised, and more individuals will enter the shadow economy. This cycle will continue (Enste & Schneider, 2000).

These cycles promote the shadow economy. The costs of the shadow economy can be found in wages, output, labour market, public goods (as mentioned before), state revenues, innovation and production of firms and access to financing from banks. Banks are more sceptical to whom they lend. Unregistered firms and individuals without a job in the formal economy will have low chances of getting a loan (Kelmanson *et al,* 2019). The effects of the existence of the shadow economy can thus also be found in the banking world, as mentioned before in the introduction. The risks for banks can come from having a connection to shadow banks (knowing or not knowing it) or shadow banks which provide liquidity or credit to the economy (GSP Working group, 2018). The shadow banks are non-bank financial institutions which raise short-term funds and use these funds to buy long-term maturity assets. However, they can't borrow during emergencies from U.S. Central Bank and the depositors' funds are not covered by insurance (Kodres, 2013).

The connections to the shadow banks can be divided in direct and indirect connection. The direct connection is for example deposits. The financial risk is issuing or holding debt securities and lending to shadow banks. Indirect exposure is for example the holding of assets. When there is a period of financial distress shadow banks dispose their assets and this will influence the valuation of assets held compared to other financial institutions. This situation can influence the overall banking market (GSP Working group, 2018). Due to these risks, it is therefore important to have Non-bank Financial Intermediaries (NBFI's) to provide funds and drive innovation to make products and services available for all individuals and promote economic growth. This is especially important in transition economies and developing countries. This was stated during a conference held in Moscow hosted by the Alliance for Financial Inclusion (AFI) and the Bank of Russia on "Financial Inclusion and Shadow Banking: Innovation and Proportionate Regulation for Balanced Growth" (GSP Working group, 2018).

Continuing this paragraph, the important areas are highlighted where the shadow economy has a negative impact. Some were already shortly mentioned in the explanation of the cycles.

REGULATIONS

To prevent these cycles, regulations are set by the government. It is therefore also very important to study the shadow economy to get a better and more accurate view of the total economic activity (formal and informal). This overview is very important to design policies which respond correctly to economic development and fluctuations (Medina & Schneider, 2018). For example, during a crisis,

when the wrong fiscal of monetary policy will be applied, this will only worsen the current economic situation. This went wrong in the beginning of the 80's in the United States, where a large part of the employment shift to the shadow economy was not recognized, which caused high unemployment and a large recession. The national statistic for unemployment missed a large part of the real employment due to the lack of knowledge of the shift of individuals to the shadow economy (Fleming *et al,* 2000). This view is also shared by Enste & Schneider (2000) as a total overview of the economic activity can help national statistics to be more accurate. These are used to create applicable regulations for the economic market. Not only the regulations are needed for the markets itself, furthermore, by researching the size of the shadow economy, the amount of tax evasion can be traced. The government can then decide on how to act on these tax evasions and how to regulate them (Medina & Schneider, 2018).

However, the increasing size can also say something about the amount, and possible too high amount, of regulations set by the government. The regulations and policies might be too much of a burden to the participants of the market (Fleming *et al,* 2000). However, the individuals have a decision which they can make on how to act on the oppressive number of regulations. According to Enste (2000), there is a two-pillar strategy. This two-pillar strategy consist of an exit option and a voice option. The exit option is when individuals decide to enter the shadow economy instead of letting their voices be heard about the too burdening number of regulations set by the government, and thus not entering the shadow economy. Examples of regulations which can be overburdening for the individuals are labour market regulations, labour restrictions for immigrants, license requirements, overall laws, and trade barriers. From previous research it can be concluded that these regulations limit the choice for participants and increase the labour costs which is shifted from the employers to the employees. This results in the employees taking the exit option and thus entering the shadow economy (Johnson *et al,* 1998).

A solution for individuals choosing the exit option instead of the voice option is to give them the ability to participate in the decision making, thus, to give them a voice and opportunity to let them be heard. This can be done through legal initiatives, referendums, and other democratic instruments. This gives individuals the opportunity to participate in the design of the tax system. The feeling of restriction and overburdening regulations will diminish due to this. This will increase tax morality and individuals will prefer the voice option above the exit option (Enste, 2018). This is exactly Enste (2018) his opinion as these strict and oppressive regulations only makes it worse. Enste (2018) states: "The shadow economy should not be seen as solely an economic problem, to be resolved by attacking the symptoms through higher fines and tougher controls". Other illegal activities as corruption and organized crime should however be fought stricter with laws, enforcements, and controls.

NATIONAL ACCOUNTS

As mentioned before, one of the main reasons to measure is to correctly calculate the national account statistics, and thus have an overview of the shadow economy. Based on these statistics, fiscal and monetary policies and regulations are determined. If these statistics are not correctly measured the polices will not apply correctly to the economic situation the market is in and will not present the true state of the national economy. This will negatively impact the current economic situation and possibly make it even worse (Fleming, 2000; Kelmanson *et al,* 2019; Enste & Schneider, 2000). An Example is stated by Fleming (2000) on the social welfare programmes. Due to the shadow economy, the national account statistic for unemployment is not correctly measured as it is overestimated. This will cause inefficient allocation of the social welfare program, where individuals who receive income due to the shadow economy will also receive the welfare benefits. The

knowledge on the size and amount of the shadow economy is therefore necessary for also allocating the public resources of the government (Kanniainen *et al,* 2004).

An interesting national account statistics is the trade of a country. This determinant was used in the empirical research for determining the size of the shadow economy by Kelmanson *et al* (2019), Medina & Schneider (2018) and Gasparênienè *et al* (2016). The link between trade and the size of the shadow economy is based on the trade openness. This relationship is a negative relation as trade is more transparent and thus easier to tax. These taxes are harder to avoid (Kelmanson *et al,* 2019).

PUBLIC GOODS

Public goods are essential for the society as they provide stability, security, basic needs, fundamental rules, and behaviour regulations. These public goods are non-excludable and non-rival goods. This means that it cannot be prevented by individuals that others of the society make use of this public good. And the use of this good by one person does not reduce quality and quantity of the public good (Moon *et al,* 2017). These public goods are funded by the tax revenues. As mentioned before, individuals and firms can have many reasons for entering the shadow economy. One of the main reasons is the too large amount of taxes (as mentioned earlier). As firms and individuals decide to avoid these taxes, the government has lower tax revenues and thus less revenues to spend on the public goods (Fleming *et al,* 2000). The spiral associated with this phenomenon is explained earlier.

LABOUR MARKETS

As people choose the exit option, thus entering the shadow economy, there will be an incorrect image of the total labour market. This will make it harder to effectively apply labour market policies (Kelmanson *et al*, 2019). In the labour market perspective, there are multiple reason why individuals prefer the informal labour market before the formal labour market. The first reason is the restriction on the number of hours of work per individual. The idea behind a restriction is to redistribute the hours and evenly divide it among employees. However, some prefer to work more hours. These individuals will enter the informal labour market to get the extra hours of work. The second reason is the obligated early retirement. Some individuals want to continue working, however if this is not possible, this individual will enter the informal labour market to be able to continue their work untaxed (Enste & Schneider, 2000). However, what these individuals need to consider is the serious negative side effects of the informal labour market. When working in the formal labour market there are safety regulations, minimal wages, social protection, and training programmes. These do not apply in the informal labour worked. Nevertheless, it provides the individual of basic needs, but individuals need to decide if they would take the risk. The chances of working below the poverty line, with possible unsafe working conditions which might also be oppressive, no social protection from the government or training programmes. Often in the shadow economy, the workers' rights are exploited (Fleming, 2000; Williams, 2015). For the firms is this also very unpleasant as employees enter the informal labour market. This causes less workers available for firms and increases competition in the official labour market (Enste & Schneider, 2000).

Not only the employees can decide to enter the shadow economy, but firms can also have this same decision. The reason for firms to enter the shadow economy depends on the amount of taxes they need to pay and the amount of regulatory pressure. What firms need to consider when enter the shadow economy is that as a firm stay smaller. This is due to the lower possibility to enter research and development programmes to maintain innovative. Also, the access to physical capital and human capital is lower as everything is done under the radar. This will weaken productivity and reduces output generated by the firm (Kelmanson *et al,* 2019).

Concluding, both employees and firms have a behavioural trade-off where they have to decide if it is worth to enter the shadow economy. For employees, is it worth the unsafe working conditions and working below poverty line instead of working in the safe official economy? For firms the trade-off between paying the taxes and dealing with the regulations while probably having issues with growing as a firm and allocating physical and human capital?

POLITICS

Another aspect influenced by the increasing shadow economy is the world of politics. Referencing back to the national account statistics, these are of importance for allocating and getting access to credit from the IMF and the World Bank, to see if countries meet EU criteria in the European Union and to determine the contribution of that specific country to EU budget (Fleming *et al,* 2000). As the rise in shadow economy causes taxes to be higher (as stated before) and thus lowers the revenues. This will increase the budget deficit of that specific country which influences the economic policy measures negatively (Aigner *et al,* 1986).

CURRENCY

Stated by Kanniainen *et al* (2004), Fleming *et al* (2000), Schneider & Enste (2002) and Aigner *et al* (1986), the amount of cash transaction or thus demand for currency is an indicator for an increase in the size of the shadow economy. The transactions are mostly done in cash, as this type of transaction is hard to observe and regulate. The level of transparency is low. Kanniainen *et al* (2004) even mentioned that half of the currency hold by the OECD countries are from the shadow economy. In the section where the shadow economy is defined, were multiple names mentioned for the shadow economy. One of these names was the cash economy (Medina & Schneider*,* 2018). One approach for measuring the shadow economy, which will be mentioned later in the thesis, is the currency approach. This is based on the amount of cash payment and thus the increase of currency (Aigner *et al,* 1986). One way to avoid the usage cash by individuals as payment method in the shadow economy is to add additional costs on cash withdrawals. However, such a type of tax might also negatively impact the official economy as overall less cash transactions will be done. This will also have a negative impact on trade, as less cash is available for purchases. Another approach taken by countries is reporting obligations by sectors which use high amounts of cash or sector where control is difficult (Awasthi & Engelschalk, 2018).

Concluding, this section stated some parts of the total economy which are negatively influenced by the shadow economy and what these negative influences are. This paragraph gives more a vision on what is already found in the literature. When reading this part, it is noticeable none of the authors had a different opinion or vision on these effects. The visions used by these authors will be used in the thesis to determine which variables will be used in the empirical method. These variables will be further explained in the methodology with some additional literature background.

Positive factors relating to the shadow economy

The shadow economy does however not only consist of negative side effects. The first positive effect, as mentioned before, is that an increase in the shadow economy implies inefficiencies in economic policies. Government can see this as a warning that their regulations and policies are too strict and the amount of taxes are too high (Fleming *et al,* 2000). The second positive effect of the shadow economy is that it provides the basic needs of income for individuals. The shadow economy is therefore even more attractive for immigrants as they have less education, live in rural areas, and are trained in more intensive labour activities in comparison to the workers which are in the formal

economy. The shadow economy can be a safety net for immigrants to maintain their basic needs as entering the formal market might be difficult sometimes (Kelmanson *et al,* 2019; Fleming *et al,* 2000). The third positive side effect is the extra value created in the shadow economy is directly spend in the official economy which stimulates the official economy positively (Ion & Alexandru, 2009). According to Enste & Schneider (2000), two-third of this income is immediately spent in the official economy.

Enste & Schneider (2000) also state that the shadow economy is also efficient as it responds directly to the demand of individuals in the shadow economy. It is ideal for small manufacturing firms and urban services. The two author's state: "the informal sector provides the economy with a dynamic and entrepreneurial spirit and can lead to more competition, higher efficiency and strong boundaries and limits for government activities".

However, considering what is stated in the paragraph on the negative sides of the shadow economy, this is debatable. More competition in the labour market and official economy can have positive effects and some serious negative effects, as mentioned before. It can bring the markets out of balance, which in turn make it harder for firms to acquire employees. Furthermore, regulations do not apply correctly to the real state of the market. Another positive effect mentioned above which is doubtful is the shadow economy creates higher efficiency. Firms which participate in the shadow economy spend less on R&D and innovate less. As they must stay from the surveillance of the government, they must be careful how to invest their profits. It is better for these firms to stay small and hidden to maintain their position in the shadow economy. Furthermore, have less access to good physical or non-physical capital.

Developed, developing and transitioning: where can it be found in the world?

Now the definition is stated for the shadow economy, the negative aspects of the shadow economy are shown, and the few positive side effect are stated; it is interesting to know how the shadow economy is allocated in global economy. This section is divided between developing, transition and developed countries, as these types of countries are marked with their own characteristics.

The summarize of what is stated before, the most important reasons for an increase, also according to Enste (2018), are: high taxes, high social security contributions, low participation in the design of regulations, respect for institution and law, intensity of regulations (restriction of hours of work an early retirement), and low tax morale. These factors make how the current amount of shadow economy in different countries is. According to Schneider & Enste (2002), the largest percentage of shadow economy can be found in developing countries. In the second place are the transition countries and the developed/OECD countries are the countries with the lowest rate of shadow economy. This is shown in table 2.

Country group	Shadow economy in % of GDP, 1988 - 2000
Developing	35 – 44%
Transition	21 – 30%
OECD	14 – 16%

Table 2: Size of the shadow economy in % of GDP, 1988 - 2000 (Source: Fleming *et al,* 2000).

The Schneider & Enste (2002) also state that the smallest amount of shadow economy is found in countries with a small public sector. Examples of these countries are Switzerland, Japan, and the

United States. Another factor is the high tax morale. This can especially be found in Switzerland and the United States.

DEVELOPING

A developing country can be defined as a country with poor government quality, underdeveloped industrial based and has low human living standards (low Human Development Index) (Ray *et al,* 2016). The developing countries are mainly based in Latin America, Asia, Africa, and the Caribbean (United Nations, 2014).

The large size of the shadow economy in developing countries is due to a lack of opportunities in the formal sector. Individuals will then enter the informal sector (Kelmanson *et al,* 2019). In table 3 is the size of the shadow economy given in a few developing countries for the period 1990 till 1993 (Fleming *et al,* 2000).

Country (developing)	Size of the shadow economy in % of GDP, average over 1990-1993
Africa: Nigeria and Egypt	68 – 76%
Tunisia and Morocco	39 – 45%
Central and South America: Guatemala, Mexico, Peru and Panama	40 – 60%
Chile, Costa Rica, Venezuela, Brazil, Paraguay and Colombia	25 – 35%
Asia: Thailand	70%
Philippines, Sri Lanka, Malaysia and South Korea	38 – 50%
Hong Kong and Singapore	13%

Table 3: Size of the shadow economy in % of GDP, average over 1990-1993 (Source: Fleming *et al,* 2000).

Table 3 shows that the size of the shadow economy in developing countries ranges between 13% and 76% of GDP. For the continent Africa, Nigeria and Egypt have the largest amount of shadow economy with a range of 68-76%. For Central and South America, Guatemala, Mexico, Peru, and Panama have the largest amount of shadow economy with a range of 40-60%. Lastly, in the continent Asia has Thailand with ease the largest amount of shadow economy with 70% (Fleming *et al,* 2000). According to a later research conducted by Schneider & Enste (2002), the size of the African shadow economy increased by 1%. In the period of 1998-1999, the size of the shadow economy of Thailand stayed the same at 70% of GDP.

According to the ILO (International Labour Organization) is the informal employment 61% of the total urban labour force of Africa. In Asia, this amount is 40-50%. Latin America has the largest amount with 80% in the informal sector. However, according to Fleming *et al* (2000), not all shadow economy is due to the informal employment market. The criminal sector is not included in the informal sector according to Fleming *et al* (2000), but it can be seen as a separate sector. The distribution of drugs, prostitution, human trafficking, and cultivation are examples of the criminal sector. The criminal sector also plays a large part in the size of the shadow economy in the developing countries, especially in Latin America ("Stopping Crime and Violence", 2018).

TRANSITION

A transitional country can be defined as a country of the third world which is conversing based on liberal market structures and liberal democracy (Fernando, 2009). Examples of this type of country are countries based in South-Eastern Europe, Moldavia, Georgia, Kazakhstan, Belarus, and Armenia (United Nations, 2014).

For transition countries, two aspects are of importance for the size of the shadow economy. These are the level of macroeconomic stability and the amount of political and economically liberalization. When these are not present or have a small presence, mainly the criminal and irregular sector play a large role in the shadow economy of transitioning countries. The transition countries have shadow economy which ranges between 7% and 43% of GDP (Fleming *et al,* 2000). Estimates determined by Fleming *et al* (2000) are shown in table 4 of the size of the shadow economy in percentage of GDP in the period of 1990-1993.

Country (transition)	Size of the shadow economy in % of GDP, average over 1990-1993
Central Europe: Hungary, Bulgaria, and Poland	20 – 28%
Romania, Slovakia, and Czech Republic	7 – 16%
Former Soviet Union: Georgia, Azerbaijan, Ukraine, and Belarus	28 – 43%
Russia, Lithuania, Latvia and Estonia	20 – 27%

Table 4: Size of the shadow economy in % of GDP, average over 1990-1993 (Source: Fleming *et al,* 2000).

For Central Europe, Hungary, Bulgaria, and Poland have the highest amount of shadow economy of 20-28%. For the former Soviet Union countries, Georgia, Azerbaijan, Ukraine, and Belarus have the highest amount of shadow economy with 28-43% of GDP (Fleming *et al,* 2000). According to research conducted by Schneider & Enste (2002), these numbers increased largely compared to the period of 1998-1999. For example, Bulgaria's shadow economy is now stated at 34% of GDP. The size of the shadow economy of Georgia increased to 64%, which it almost doubled in size.

DEVELOPED

A developed country can be defined as a country with high income per capita, high industrial developed with own production and economy, good infrastructure, technology advanced and high Human Development Index (Golic, 2019). Examples of developed countries are European Union countries, other Europe states like Switzerland, Norway and Iceland and countries outside Europe like Canada, Japan, Australia, New Zealand, United Kingdom, and the United States. The focus of this thesis, as mentioned before, is on the United States.

The developed countries have a small size of shadow economy compared to the transition and developing countries due to good laws, low presence of corruption and decent regulatory and tax burden (Enste & Schneider, 2000). However, the small amount that is present in the developed countries are mainly undeclared labour and tax evasion (Kelmanson *et al*, 2019). As mentioned before, in transition and developing countries, criminal activity is also included in the shadow economy. According to Enste & Schneider (2000), the higher amount of shadow economy in transition countries is mostly due to the high number of overburdening regulations which causes incentives for bribery and corruption.

In table 5 are the estimations presented provided by Fleming *et al* (2000), which show the size of the shadow economy in percentage of GDP in the period of 1990-1993 for developed countries. The size ranges from 8% till 30% of GDP. These percentages are significantly lower than the percentages present with the transition and developing countries (Fleming *et al,* 2000; Kelmanson *et al,* 2019).

Country (developed)	Size of the shadow economy in % of GDP, average over 1990-1993
Greece, Italy, Spain, Portugal, and Belgium	24 – 30%
Sweden, Norway, Denmark, Ireland, France, The Netherlands, Germany, and Great Britain	13 – 23%
Japan, United States, Austria, and Switzerland	8 – 10%

Table 5: Size of the shadow economy in % of GDP, average over 1990-1993 (Source: Fleming *et al,* 2000).

As shown in table 5, the Southern European countries and Belgium have the largest amount of shadow economy ranging between 24-30%. In second place are the Scandinavian countries, Ireland, France, The Netherlands, Germany, and The UK with a size of 13-23%. The developed countries with the least amount of shadow economy are Austria, Switzerland, Japan, and the United States. According to research done by Schneider & Enste (2002) on these countries with data of five years later, Greece and Italy still have the largest amount of shadow economy with 30 and 27%. The Scandinavian countries are still at second place. Austria and the United States are at 10%. Switzerland is at 9%. However, it is worth mentioning that the shadow economy of the United States has doubled compared to the year 1970.

As the focus of the thesis is on analysing the factors influencing the shadow economy in the United States, it broadens the view and idea of the shadow economy to know how the state of the shadow economy is in the rest of the world. When comparing the size of the developing, transitioning, and developed countries, the United States (as developed country) has a very low amount of shadow economy compared to the rest of the world. Even if this is a small amount compared to other countries, this small amount still has an impact on the overall economy of the United States. It is therefore necessary to don't ignore the shadow economy in the developed countries. This is part of the reason for analysing the United States in this thesis. The other reason was the large amount of available data on the United States. As shadow economy is known for being hard to research and analyse with empirical models, the availability of data is an important factor for determining the focus of the thesis. Furthermore, this section provides an indication on which determinants are more impactful for a developed country like the United States. As tax evasion and undeclared labour are the main shadow activities for developed countries, the focus should be on these two activities for deciding which determinants should be added to the empirical research.

Theoretical Framework: introduction to the method applied
The determinants influencing the size of the shadow economy
In the section "Why study the shadow economy?", some determinant for measuring the shadow economy were already mentioned. In this section the focus will be more specified on the determinants within these areas. These determinants are conducted from analysing literature of the shadow economy which are in this section shown and explained. I found that these determinants were the best fit for the empirical research applied in this thesis. The determinants are the independent variables for the empirical research.

As mentioned earlier in the introduction, the thesis will be using the multivariate regression analysis and Pearson correlation coefficient to measure the effects of these determinants of the shadow economy. This is based on the work of L. Gasparênienè, R. Remeikienè, & M. Heikkila (2016). They measured the impact of the shadow economy determinants with a case study on Ukraine. This study

will be extended in this thesis with more determinants. The determinants used by Gasparênienè *et al* (2016) are GDP per capita, total tax rate, total payments, export and import, freedom from corruption, total labour force, business freedom, unemployment rate, domestic credits to private sector and government expenditure.

Based on literature, Enste (2018), Aigner *et al* (1986) and Enste & Schneider (2000) agree on the high taxes and overburdening regulations as the main drivers of the shadow economy. Aigner *et al* (1986) based this on the surveys taken in the period from 1955 till 1982, concluded another driver: tax morality. Both authors Enste & Schneider (2018 & 2000) see another main factor, the high payments for the social security contribution. Nesvetailova (2018) mainly sees as main driver, and states it as the origin of the shadow banking world, the regulatory avoidance and arbitrage. Regulatory arbitrage is defined as individuals, or economic agents, work their way around different levels of regulations because the different levels of regulations have both costs and benefits (Coendet, 2021). Concluding, the main drivers, Regulatory burden, taxes, tax morality and payments for socials security contribution are included in the empirical research conducted in this thesis. The variables tax morality and social security contribution were not added to the research of Gasparênienè *et al* (2016).

In this thesis the following determinants are used: Size of the shadow economy in percentage of GDP (dependent variable), tax burden, unemployment rate in percentage of total labour force, participation, GDP per capita, imports, exports, regulatory quality, cashless payments, social security contribution and tax morality. These determinants will be further emphasized and explained in this section.

TAX BURDEN (IN % OF GDP)

This variable is stated as one of the most important determinants for the growth in the shadow economy. For the empirical research in this thesis, this variable is defined as tax revenue in percentage of GDP. This measure shows how much of the taxes generated by the government are part of the country's output. This is used to measure the tax burden of a country ("Tax revenue", 2020). This variable is influenced by the difference between the after-tax earnings from work and the labour cost of the official economy. When this difference is large, individuals have a stronger incentive to enter the shadow economy (Enste & Schneider, 2000). Between taxes and the size of the shadow economy is a negative relation expected. This negative relation can be explained with the Laffer curve.
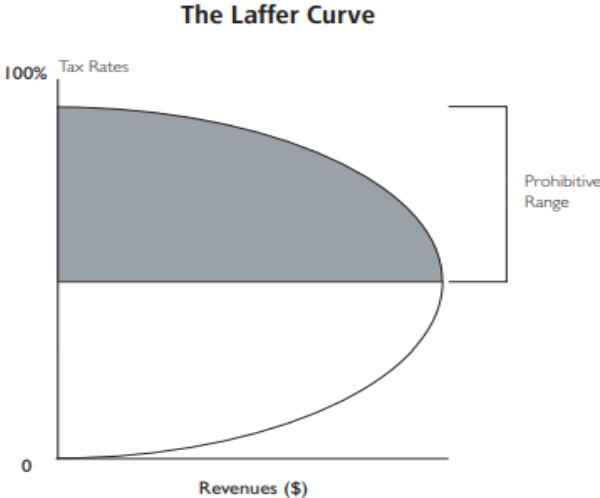

Figure 1: The Laffer curve (Source: Laffer, 2004).

The Laffer Curve, created by the economist Arthur Laffer, provides an overview on the effects of the change in tax rate on the tax revenues of the government. This is shown in figure 1. The prohibitive area represents the arithmetic effect, where the tax revenue increases while the tax rate increases. After this area, the economic effect is of relevance (Laffer, 2004). For this thesis the focus will be on the economic effect, which explains the reasoning for the negative effect between tax burden and the size of the shadow economy. The economic effect of the Laffer curve is applied on the size of the shadow economy in Figure 2.

Figure 2: Economic effect applied on the size of the shadow economy with relations between the variables.

The variable used in the thesis is thus the tax-to-GDP ratio. In the thesis is a negative relationship expected between the tax revenue and the size of the shadow economy. As the size of the shadow economy increases, more individuals exit the official economy, thus less tax revenues are acquired by the government. This negative relationship is influenced by the positive relation between tax rate and the size of the shadow economy. As the tax rate increases, more individuals are less willing to work in the official economy. More will enter the shadow economy (Laffer, 2004; Kelmanson *et al,* 2019 & Schneider & Enste, 2002).

Another interesting aspect, which has an influence on individuals entering the shadow economy due to taxes, is the complexity of the tax system. According to Schneider & Neck (1993), not much attention has been paid to this matter. They state that when the tax system is complex, more possibilities arise for legal tax avoidance. The introduction of reductions of taxes and tax exemptions makes the risk of getting caught, while being active in the shadow economy, less attractive (Enste & Schneider, 2000). The complexity of the tax system on the size of the shadow economy, might be interesting for future research.

To conclude, for the empirical research conducted in this thesis there is a negative relation expected based on multiple results of other literature work as tax revenue is used.

SOCIAL SECURITY CONTRIBUTION (IN % OF GDP)

The variable social security contribution is defined as the amount paid for the social security contribution in percentage of GDP. This variable also depends on the difference on the after-tax earnings and the total cost of labour in the official economy. When this difference is larger, the incentive for individuals (and firms) to enter the shadow economy also increases (Enste & Schneider, 2000; Schneider & Enste, 2000). Another interesting aspect found by Enste & Schneider (2000) is when individuals receive the social welfare payments due to not having a job, they still have an incentive to enter the shadow economy. Their total income received from the welfare payments and due to working in the shadow economy, increases this amount.

To conclude, there is a positive relation between the social security contribution and the size of the shadow economy.

REGULATORY QUALITY (INDEX)

The variable regulatory quality is based on the regulatory quality index from the World Bank. The index shows how the government can implement good quality regulations and policies which promote private sector development. The index ranges from -2.5 to 2.5. A high index implies high regulatory quality ("Regulatory quality", 2019). Regulations are also one of the main drivers of the growing shadow economy. This is due to governments implementing too overburdening government regulations which suppress entrepreneurial freedom (Kanniainen *et al,* 2004; Kelmanson *et al,* 2019). Examples of regulations are trade barriers, labour market regulations, license requirements and entry costs (Schneider & Enste, 2002; Kelmanson *et al,* 2019). These regulations increase the cost of labour. Firms can easily shift this cost on their employees, which gives incentives for the employees to enter the shadow economy (Enste & Schneider, 2000; Schneider & Enste, 2002). According to Gasparênienè *et al* (2016), Kelmanson *et al* (2019) and Schneider & Enste (2002), it is often a combination of weak institutional quality and overburdening regulations which causes individuals to enter the shadow economy. Furthermore, these are perfect conditions for corruption to arise.

To conclude, overburdening regulations do not only harm the individuals in the official economy, but also the firms. Therefore, there can be a negative relation be found. As the index increases (higher regulatory quality), the size of the shadow economy will decrease.

OVERALL GOVERNMENT SCORE (INDEX)

This variable represents how to government performs on the following 12 area's: property rights, judicial effectiveness, government integrity, tax burden, government spending, fiscal health, business freedom, labour freedom, monetary freedom, trade freedom, investment freedom and financial freedom. This index is also known as the 2021 Index of Economic Freedom ("Overall score", 2021). This variable gives a broader perspective compared to the variable regulatory quality and therefore interesting to include.

TAX MORALITY (IN %)

Tax morality is a variable which is more of a social-economic variable compared to the others (Enste & Schneider, 2000). This variable shows if morals and social standards influence the activities executed in the shadow economy (Kanniainen *et al*, 2004). As this research implies an interdisciplinary view, considering these types of variables are necessary. The tax morality is described as the willingness to pay taxes ("Tax morale", 2019). For this research, this variable is seen as the percentage "too high". So, from a survey conducted, how many individuals found the tax burden too high and thus have a lower morality for paying taxes. This variable is seen as hard to measure due to the social aspect of it. Only way to measure it, is to apply a micro perspective approach by conducting surveys (Enste & Schneider, 2000 & Kanniainen *et al,* 2004). Tax morality is influenced by taxes, social security contribution and regulations set by the government. When the government implements too many regulations, high taxes and high payments which are used for social security contribution, individuals feel less loyalty to the institutions and enter the shadow economy (Enste & Schneider, 2000). However, an interesting concept mentioned by Kanniainen *et al* (2004), is that individuals can keep others from entering the shadow economy. Individuals in the official economy show social disapproval to free riders in the shadow economy. The free riders have risks of being socially stigmatized (Kanniainen *et al,* 2004).

To conclude, when there is high tax morality, there is a smaller shadow economy. Therefore normally, there would be found a negative relation between tax morality and the size of the shadow economy. However, because this research makes use of a "percentage too high" survey data, this

implies the higher the percentage, the lower number of individuals willing to pay their taxes and thus more individuals enter the shadow economy. As the percentage goes up, which implies that individuals find the tax burden too high, the shadow economy increases. So, a positive relationship can be found.

GDP PER CAPITA (IN CURRENT $)

GDP per capita is defined as the value created by production of goods and services in a certain period for a country ("Gross domestic product (GDP)", 2021). This measure is known for an indicator to show the overall state of the economy. In the following graph (Figure 3), the relationship between GDP and the size of the shadow economy is shown.
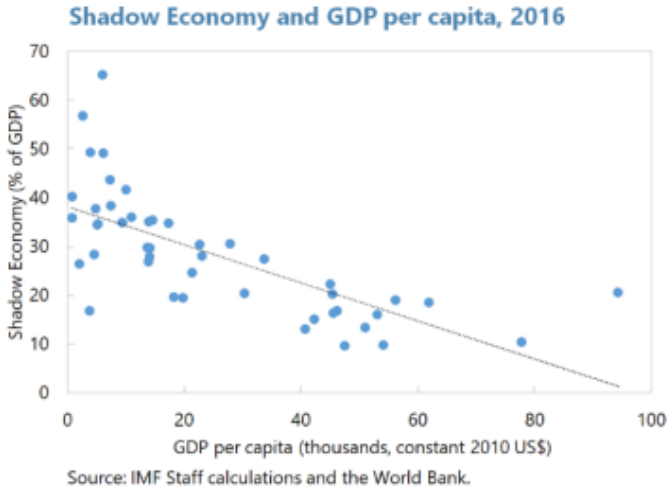
Figure 3: relation between the size of the shadow economy and GDP per capita (Source: Kelmanson *et al,* 2019).

From figure 3, it can be concluded there is a negative relationship between the size of the shadow economy and GDP per capita. When GDP increases, the shadow economy decreases in size. As more economic activity is in the shadow economy and not in the official economy, the size of the official GDP is lower, this can be interpreted from figure 3. Furthermore, from earlier research conducted, can be stated that countries with more regulation tend to have a larger share of shadow economy in the total GDP (Kelmanson *et al,* 2019; Enste & Schneider, 2000).

To conclude, there is a negative relationship between the GDP per capita and size of the shadow economy.

IMPORTS & EXPORTS (IN BILLION $)

The variables imports and exports are measured in billions of dollars. According to Kelmanson *et al* (2019), when individuals trade, it is harder for them to conceal shadow economy trade. This is since trade is relatively transparent and thus easier to tax. Avoiding taxes is therefore harder. Therefore, there is a negative relationship expected between trade and the size of the shadow economy. This view is also shared by Medina & Schneider (2018) and Kelmanson *et al* (2019).

CASHLESS PAYMENTS (IN BILLIONS OF TRANSACTIONS)

The variable cashless payments is defined as the amount of online transactions. This is an important variable as the shadow economy and cash economy are closely linked to each other. Most transactions are done with cash as this type of transaction can be easily hidden (Awasthi &

Engelschalk, 2018; Kanniainen *et al*, 2004). According to Kanniainen *et al* (2004), most shadow economy transactions are done with cash, however this is challenged with new payment methods. As their research is conducted in the year 2004, in the meantime, a lot might be changed due to, for example, the emerging of cryptocurrencies. However, for this thesis is assumed that most shadow economy transaction are done with currency. This might be interesting as a point for future research.

Awasthi & Engelschalk (2018) however did not focus on the variable cash payment, instead, they looked at the amount of cashless payments. According to them, multiple studies showed that a larger number of electronic payments is associated with a smaller size of shadow economy. Furthermore, for electronic payments it is also easier to find data, as more is known compared to cash payments. Electronic payments are easier registered and thus easier traceable. Therefore, promotion of the government for using more electronic payments and less cash payments might be a solution. An option might be to introduce additional costs to cash withdrawals; however, this solution is possibly unpopular with the business community. These taxes will reduce the amount of cash available for purchases and this will have a negative effect on trade. Another option might be tax incentives for non-cash payments (Awasthi & Engelschalk, 2018).

The US has also taken measures against unreported cash transactions from the shadow economy. They introduced the requirement of reporting transactions with an amount of $600 or more with independent contractors. These transactions must be reported to the IRS (International Revenue Service) (Awasthi & Engelschalk, 2018).

To conclude, as cashless payments increase, the size of the shadow economy will decrease. Therefore, there is a negative relation expected.

UNEMPLOYMENT RATE (IN % OF TOTAL LABOUR FORCE)

Which is interesting to read in the literature is that some make use of the variable total labour force and some the unemployment rate as percentage of total labour force. As stated by Gasparênienè *et al* (2016), the individuals which are unemployed can easier hide their illegal income compared to the hired employees. However, in their research, they made use of the variable total labour force. They used this variable since total labour force consists of both employed and unemployed and used this measure for the unemployment part. However, I would prefer to only take the unemployment part as percentage of total labour force as these are the most active group in the shadow economy. Enste & Schneider (2000) share this view as they see the rise in unemployment rate as one of the main causes in the increase in size of the shadow economy. So, they state a positive relation between the unemployment rate and the size of the shadow economy.

When analysing the literature, Ion & Alexandru (2009), state that it is unsure whether the relationship between the unemployment rate and the size of the shadow economy is a positive or negative. It could be negative as the shadow economy activity could be positively related to a growth in GDP and then the unemployment rate is negatively correlated to the size of the shadow economy. However, the positive relation is due to unemployed working part of their time in the shadow economy, as this is easier for them than the employed, as mentioned before. Therefore, a negative relation between GDP and the size of the shadow economy is expected, as stated before. And this causes to expect a positive relation between the unemployment rate and the size of the shadow economy.

To conclude, for this thesis a positive relation is expected as unemployed are often working in the shadow economy.

PARTICIPATION RATE (IN %)

The participation rate is a percentage of the total labour force divided by the working age population. This is the population with the age between 15 and 64 years ("Labour force participation rate", 2021). As more individuals work in the shadow economy, the participation rate falls for the official economy (Kelmanson *et al, 2019*; Schneider & Enste, 2002). Therefore, to conclude, there is a negative relation between the participation rate and the size of the shadow economy. This is also based on the relation stated between the size of the shadow economy and the unemployment rate.

OVERVIEW OF THE DETERMINANTS

To conclude the section of the determinants influencing the size of the shadow economy, an overview is provided of the relations between the size of the shadow economy of the United States and the determinants. This is shown in figure 4.
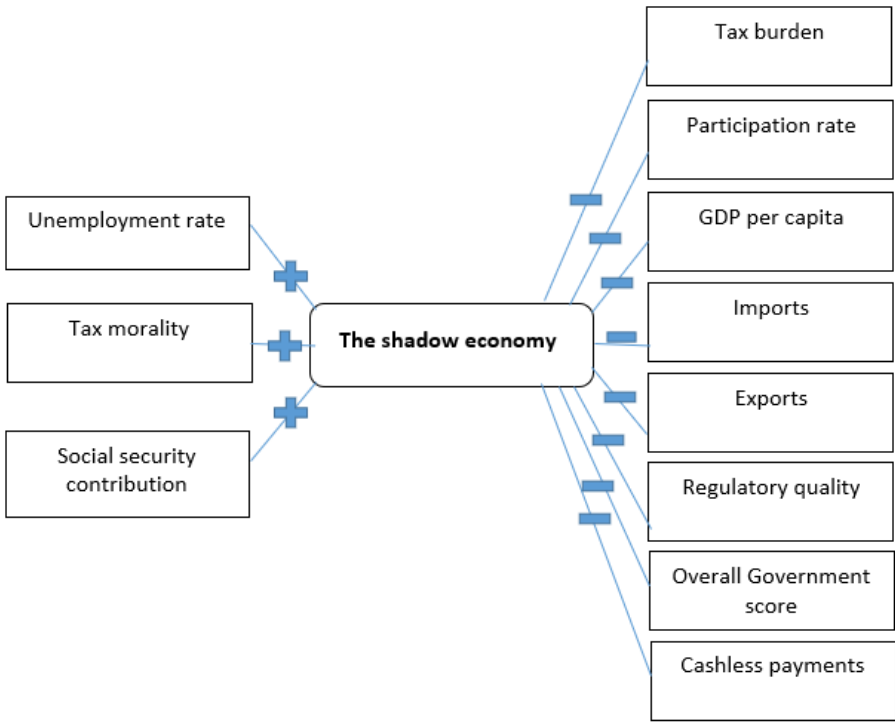
Figure 4: Overview of the positive and negative linear relations with the size of the shadow economy of the United States.

Approaches for measuring the size of the shadow economy and the determinants

The shadow economy is a very broad and difficult concept to perfectly define (Ion & Alexandru, 2009), let alone research the causes or measure it. As Kanniainen *et al* (2004) stresses, it is difficult to acquire accurate data on the shadow economy as every individual in the shadow economy does not want to be identified. This was stated by many authors, as they presented their methodology. They highlighted the fact that it is not easy to measure and there is not one perfect way to do it. The choice for a particular model or measure depends on the data availability and what the research objectives are (Kelmanson *et al,* 2019). Therefore, many different measures are used when measuring the shadow economy. According to Fleming *et al* (2000), the most common used measures of the shadow economy are the discrepancy between demand of services and goods, for

example currency, tax revenue and electricity, and the actual usage of these goods. Another very famous measure is surveys.

Appendix A provides an overview of the methods used to measure the size of the shadow economy. However, as the research objective for this thesis is not the measure the size of the shadow economy for the United States, the approaches will not be explained in detail.

In this thesis the data used for the dependent variable is provided by the research conducted by Medina & Schneider (2018). They used the MIMIC model for estimating the size of the shadow economy for multiple countries. For determining the reliability of the research and thus the data used, it is important to know how the data is retrieved. Therefore, the MIMIC model will be further explained.

Aigner *et al* (1986) reviewed multiple methods, however, they concluded that the best and most trustworthy approach is the MIMIC model (Multiple-Indicator Multiple-Causes model). Schneider & Enste (2002) share this view as they point out that this approach considers multiple effects and causes simultaneously. Therefore, this model is found many times in literature. Examples of authors which used the MIMIC model are Medina & Schneider (2018), Kelmanson *et al* (2019) and Ion & Alexandru (2009)**.** This model is a theory-based model which makes use of a latent variable, in this case is that the shadow economy. This model is used to confirm the impact of the exogenous variables have on the latent variable. To use this model, it is important to first define clearly what the relation between the exogenous variable is with the latent variable, to establish a structured base for the research. This model thus confirms more than it explains. The model can be seen to view it as two equations. The first equation shows the latent variable as the dependent variable and the causes of the shadow economy as the independent variables. The second situation shows the latent variable as the explanatory variable for the other indicators of informality. With this approach the fitted values of the latent variable are used to calculate the size of the shadow economy (Medina & Schneider, 2018; Kelmanson *et al,* 2019). This method is very useful as the dependent variable is unknown (Balcioglu, 2008), which is often the situation with measuring the shadow economy. However, as it is stated by Aigner *et al* (1986) as the perfect method, according to Kelmanson *et al* (2019) and Ion & Alexandru (2009) this approach has some short comings. This method is very sensitive to small changes in the sample, starting values and data. Furthermore, the combination of time series data and the MIMIC model can be difficult, calculation of the confidence interval is difficult, and it is hard to test the hypothesis of independence between the measurement errors and structural errors.

A variation on the MIMIC model is the DYMIMIC model. The expansion of the MIMIC model to the DYMIMIC model was done by Aigner *et al*. (1988). The DYMIMIC model is actually the MIMIC model in first differences (Balcioglu, 2008). Both models are popular for measuring and are used often in many studies (Balcioglu, 2008; Buehn & Schneider, 2008). From the literature analysed, the DYMIMIC model is also used by Enste & Schneider (2000) and Kanniainen *et al* (2004). Kanniainen *et al* (2004) first estimated the size of the shadow economy by using the currency demand method and then used this data in the DYMIMIC model.

Gasparênienè *et al,* 2016, however, used a totally different approach by applying first the Pearson correlation coefficient and then a multivariate regression analysis to measure the effect of the variables on the size of the shadow economy in the case of Ukraine. However, the research conducted by Gasparênienè *et al* (2016) does not state it clear whether the Pearson correlation coefficient is used to select all the variables initially used for the multivariate regression. For this thesis, it is assumed that this is the reason for applying the Pearson correlation coefficient.

The Pearson correlation, full name is Pearson product-moment correlation coefficient, measures the correlation, thus the strength and direction, between two variables which are continuous. The Pearson correlation tries to find a relation between the data of the two variables and tries to draw a line between of best fit. The Pearson coefficient, denoted as *r,* states how far these points are from the line of best fit. It can range from -1 till 1. When the value is 0, there is no relationship between the two variables (Laerd statistics, n.d.).

The second step applied by Gasparênienè *et al* (2016) is to take the independent variables where the Pearson correlation coefficient showed is statistically significant. The independent variables used in the multivariate regression analysis are: Total labour force, participation rate, GDP per capita, total taxes, import. The dependent variable is the level of shadow economy. A multivariate regression analysis is used to measure the impact of the independent variables on the dependent variable (the total variance) and the overall fit of the model on the data used (Laerd statistics, n.d.).

The reason for applying the multivariate regression analysis and Pearson correlation coefficient is because it is an approach which is not often used (Gasparênienè *et al*, 2016). While reading many literature work, there was indeed never one other paper which also applied this approach. Another benefit for applying this approach is that the multivariate regression analysis can identify anomalies or outliers and it can measure/give a in detail understanding of the relation between each independent variable with the dependent variable. However, not only the individual relation but also the relation between all the independent variables taken together on the dependent variable is measured (Weedmark, 2018; Keith & Marill, 2004). A disadvantage stated by Weedmark (2018) is the usage of the data. When a too small sample is used, the outcomes might not be correct or unreliable.

To conclude, to correct methodology for the research conducted depends highly on the data availability and the research objective that an academic wants to achieve. Therefore, when researching the shadow economy, every approach can be useful; however, then it is of importance to state why this approach is most beneficial for your research. The Pearson correlation coefficient and the multivariate regression analysis are beneficial as they measure the effect and strength of the effect of multiple independent variables on the dependent variable, which is the research objectives for this thesis. The research objective is to measure which variables have a relation with the dependent variable and how strong this relation is. The research objective is not to estimate the size of the shadow economy of the United States. It is furthermore interesting to know how the data for the dependent variable is acquired as this is done with the MIMIC model by Medina & Schneider (2018).

Hypothesis
Based on the literature and research objectives, the following two hypothesis are stated. These hypotheses provide the answer to the research question: *Which determinants influence the size of the shadow economy and what is the impact of these determinants for the United States in period of 1991 till 2015?*

Hypothesis 1: When unemployment rate, tax morality and social security contribution increase in size, the shadow economy will increase in size. There is a **positive** linear relationship.

Hypothesis 2: When tax burden, participation rate, GDP per capita, imports, exports, regulatory quality, overall government score and cashless payments increase, the shadow economy will decrease in size. There is a **negative** linear relationship.

These relations can be found in figure 4. This figure can be found at the conclusion of the section where the determinants were discussed.

Empirical strategy

As mentioned before, the methodology applied is inspired by the research conducted by Gasparênienè *et al* (2016). The method applied, will be further extended by adding new determinants. Furthermore, the research will be more detailed in explaining how to measure the effect of the determinants on the dependent variable, the size of the shadow economy of the United States in percentage of the official GDP. The determinants earlier presented in the theoretical framework will be the independent variables. The dependent variable and the independent variables used in the empirical research are shown in table 6.

Variable	Definition
SE	Size of shadow economy (in % of GDP)
Tax	Tax burden (in % of GDP)
Unemployment	Unemployment rate (in % of total labour force)
PR	PR (in %)
GDPcurrent	GDP (current $)
Import	Imports (in billion dollars)
Export	Exports (in billion dollars)
OGS	Overall government score index
RQ	Regulatory quality index world bank
CP	Cashless payments (in billions)
SSC	Social security contribution (in % of GDP)
TM	Tax morality (% too high)

Table 6: Definitions of the dependent variable and independent variables (imported from STATA 13.0).

For the empirical research the statistical software STATA version 13 is used. Furthermore, a significance level of 10% is applied. The data gathered for conducting this research will be explained in the next section. The research is conducted on the period of 1991 till 2015 for the United States.

The empirical methodology can be divided in two steps.

STEP 1: THE PEARSON CORRELATION COEFFICIENT

The first step is identifying the significant independent variables which have a linear associating with the dependent variable. This is done by measuring the Pearson correlation coefficient of each independent variable on the dependent variable (Laerd statistics, n.d.). The Pearson correlation coefficient has been explained earlier. Formula 1 (Jung *et al*, 2020) defines how the Pearson correlation coefficient is measured.

$$Rxy = \frac{\sum_{i}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i}^{n}(Y_i - \bar{Y})^2}} \tag{1}$$

The Pearson correlation coefficient can range from -1 to 1. When $r < 0$, there is a negative correlation between the two variables; when $r = 0$, there is no correlation; and when $r > 0$, there is a positive correlation. This is also shown in figure 5.
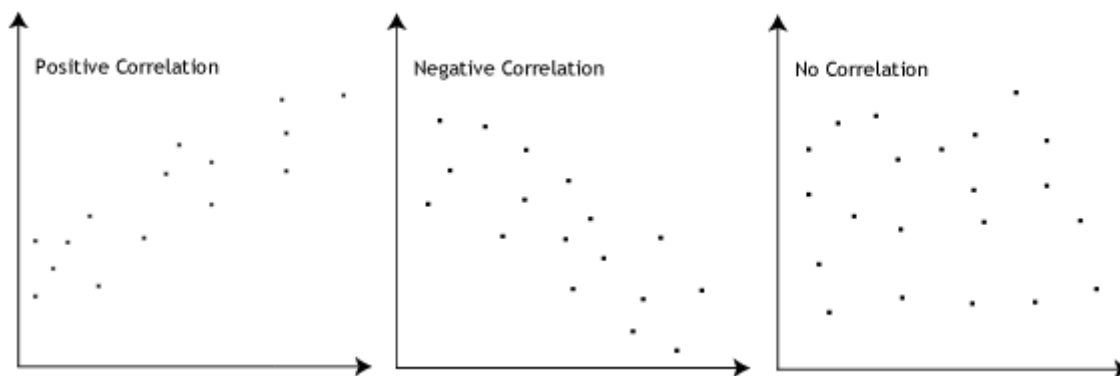
Figure 5: diagrams which show a positive, negative and no correlation between two variables (Source: Laerd statistics, n.d.).

The strength of the correlation depends on the steepness of line of fit which goes through/between the points presented in the diagrams. The steeper the line, the stronger the correlation (Laerd statistics, n.d.). The following table (7) is used to examine the strength of the correlation. The range for the strength of the correlation is equal for both the positive and negative coefficient.

0.1 < | r | < 0.3	-0.1 < | r | < -0.3	Weak correlation
0.3 < | r | < 0.5	-0.3 < | r | < -0.5	Medium correlation
| r | > 0.5	| r | > -0.5	Strong correlation

Table 7: Ranges of the type of correlation measured by the Pearson correlation coefficient (Source: Laerd statistics, n.d.).

After determining the strength of the correlation and the significance, the multivariate regression analysis can be executed with the significant variables.

STEP 2: THE MULTIVARIATE REGRESSION ANALYSIS

The multivariate regression analysis will be used to see the effect of each independent variables on the dependent variable, while holding the other independent variables constant (Studenmund, 2017). The structure of a formula with a multivariate regression is formulate as formula 2:

$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \cdots \beta_N * X_N + \in \qquad (2)$$

The variables which came out significant from the Pearson correlation analysis are placed in the formula. After regressing, there will be checked for multicollinearity by using the VIF. As multicollinearity influences the standard errors, variances, t-scores and even influence the expected signs for a coefficient (Studenmund, 2017). The Variance Inflation Factors (VIF) measures how much the variance has increased due to multicollinearity. The formula for the VIF is:

$$VIF\ (\hat{\beta}) = \frac{1}{(1-R_2)} \qquad (3)$$

To acquire decent results, it is important to determine if there is severe multicollinearity. According to Studenmund (2017) and Woolridge (2015), there is no formal table produced which provides the range for when there is no or severe multicollinearity. Therefore, there are different opinions on when there is severe multicollinearity. For example, STAT462 (2018) have as rule of thumb that there is severe multicollinearity as the VIF > 10. Wooldridge (2015) states carefully that the minimum for severe multicollinearity is above 10. For the thesis the following ranges (table 8) will be used based on what is stated by Studenmund (2017), woolridge (2015) and Stat462 (2018):

VIF value	Amount of multicollinearity
VIF = 1	No multicollinearity
1 < VIF ≤ 10	Moderate multicollinearity
VIF > 10	Severe multicollinearity

Table 8: Ranges for multicollinearity.

When the regression is tested for multicollinearity, the results can be defined on which and how much the determinants have an influence on the size of the shadow economy in the United States. Furthermore, the R-squared will be analysed to view how much the total sample variation of the dependent variable can be explained by the independent variables (Wooldridge, 2015).

Data

For the empirical method a numerous of data sources are used to determine the size of the shadow economy. The data ranges from indexes, percentages, currency ($) and number of payments in annual data. The data chosen for a variable is based on the reliability and availability of the data for the period 1991 till present of the United States. As different variables were chosen, based on literature, no dataset could be used. All the data had to be retrieved individually from databases. The data used can be found in Appendix B.

Dependent variable: the size of the shadow economy in percentage of the official GDP.

The data was found in a research done by Leandro Medina and Friedrich Schneider. This research was conducted for the International Monetary Fund working papers in the year 2018. In this paper, the size of the shadow economy was measured by using the MIMIC model for the period of 1991 till 2015 for 158 countries. Both authors are economist. Leandro Medina worked as economist for the IMF at the African Department, with a focus on Mozambique and Ghana, where he analysed the characteristics and size of the shadow economy. Furthermore, Medina also has a background in macro fiscal issues in South-east Asia ("Leandro Medina", n.d.). Prof. Dr. Friedrich Schneider is active at the Johannes Kepler University of Linz and, since 2006, active as research professor at the DIW in Berlin. Furthermore, he held multiple honorary positions at different universities and did many visits. He is well known for his publications on the shadow economy ("Friedrich Schneider", n.d.). He published articles for the The Quarterly Journal of Economics, Public Choice, The Economic Journal, The American Economic Review, Journal of Economic Literature and Kyklos. Furthermore, he also published multiple books and chapters ("Friedrich Schneider", n.d.; Medina & Schneider, 2018).

Independent variable: Tax burden in percentage of GDP.

The data for the United States for the variable Tax was found at the data bank of The World Bank. The world bank is a partnership of five institution which have as mission to promote shared prosperity and reduce poverty by providing sustainable solutions. The World Bank can be found in 130 countries and has 189 member countries. The five institutions are The International Development Association, The International Bank for Reconstruction and Development, The International Finance Corporation, The Multilateral Investment Guarantee Agency, and The International Centre for Settlement of Investment Disputes ("WHO ARE WE", n.d.).

The variable Tax represents the tax revenue in percentage of GDP, defined as the transfers to the government for public purposes. Social security contribution, fines and penalties are not included. Erroneously collected taxes and refunds are seen as negative tax revenue ("Tax revenue (% of GDP)", 2021). The data could be retrieved from 1991 till 2020.

Independent variable: Unemployment rate in percentage of the total labour force.

The unemployment rate of the United States is retrieved from the Federal Reserve Economic Data (FRED), which is an online database for international, national, private, and public data ("What is FRED?", n.d.). The data could be found for the period of 1991 till 2020. This variable represents the part of the total labour force which is unemployed between the age of 16 years and older (U.S. Bureau of Labor Statistics, 2021).

Independent variable: Participation rate in percentage.

The data for the variable participation rate of the United States is found at Statista. This is also an online database which is a provider of consumer and market data. With 2 million registered users, 23 million visits per month, 22,500 sources, over 1 million statistics, 8,000 topics and 170 different industries, they are the market leader. Statista is located in many countries, for example Germany, United Kingdom, United States, France, Singapore, Netherland, Japan and others ("About – us", n.d.)

This variable shows the American population which participated in the labour market, who are eligible to work. The data could be found for the period of 1991 till 2020 (Statista Research Department, 2021).

Independent variable: GDP per capita in current $

The data for the variable GDP per capita of the United States could also be retrieved from The World Bank for the period of 1991 till 2020. The GDP is the sum of gross value which is added by producer resident in the United States. The taxes on products are included and subsidies are not included ("GDP per capita (current US$)", 2021).

Independent variable: Export and import in billion $.

For both the variables, Export and Import, the data of the United States has also been recovered from FRED. It could be retrieved for the period of 1991 till 2020. Both exports and imports exist of all the goods and services send and retrieved by the United States (U.S. Bureau of Labor Statistics, 2021).

Independent variable: Overall government score as index.

The overall government score for the United States was retrieved from the 2021 index of economic freedom. This index is published by The Heritage Foundation. This index covers different types of freedom, ranging from financial freedom to property rights for 184 countries. The overall government score considers all the different types of freedom. The higher the index, the more freedom there is in that specific country ("Overall score", 2021). This data was also used by Kelmanson *et al* (2019). This index was found for the years 1995 till 2020.

Independent variable: Regulatory quality as index.

This variable is more of a spin-off of the overall government score; however, it has a more specific focus on the regulatory quality of the United States. The data for this variable is retrieved from The World Bank and provides information on how capable the government is with implementing decent policies and regulations. The data could be retrieved for the years 1996, 1998, 2000, 2002, 2003 till 2019 (Worldwide Governance Indicators, n.d.). The index can range from 2.5 to -2.5, which a lower/negative represents a low regulatory quality (Worldwide Governance Indicators, 2020).

Independent variable: Cashless payments.

The data for the variable cashless payments has also been retrieved from Statista. The data available ranged from 2000 to 2018. This variable presents all the non-cash payments done in the United States. The different non-cash payments are checks, credit cards, ACH debit transfers, non-prepaid debit cards, ACH credit transfers and prepaid debit cards (de Best, 2021).

Independent variable: Social security contribution as percentage of GDP.

The data for the variable social security contribution is retrieved from the database from the OECD. The OECD, Organization for Economic Co-operation, and Development, existing for 60 years has as main priority to shape policies which foster well-being, equality, prosperity, and opportunity. They possess a large database with indicators and publications. The data of the social security contribution of the United States provided by the OECD are the compulsory payments to the government, paid by employers and employees, which are given to institution that provide for example unemployment benefits, pensions, accident/injury/sickness benefits, reimbursements for hospital and medical expenses and family allowances. The data was retrieved for the period 1991 to 2019 ("Social security contribution", 2020)

Independent variable: Tax morality in percentage too high.

Also, the data for Tax morality of the United States was found in the database of Statista. The data provides an overview of how many Americans find their taxes on personal income too high. This data was gathered by surveys conducted through telephone interviews. Age range of the individuals called were 18 years and older. The data was provided for the years 2000 till 2021, except for the year 2002 (Duffin, 2021).

Results

The first step was measuring the Pearson correlation coefficient of the size of the shadow economy of the United States with each independent variable. The tables can be found in Appendix C.

The results show the strength of the relationship between the dependent and independent variable. A summary is given in table 9.

Independent variables	Strength of the correlation with dependent variable
Tax	Medium correlation
Unemployment	Medium correlation
PR	Medium correlation
GDPcurrent	Strong correlation
Import	Strong correlation
Export	Strong correlation
OGS	Weak correlation
RQ	Weak correlation
CP	Medium correlation
SSC	Medium correlation
TM	Weak correlation

Table 9: Strength of the correlation between the size of the shadow economy of the United States with the independent variables.

Furthermore, if there is a significant relationship between the two variables, this is recognizable due to the stars (* or ** or ***) denoted after the Pearson correlation coefficient (see Appendix C). The significance level, as stated before is at 10%. The significant variables are Tax, Unemployment rate, Participation rate, GDP per capita, Import, Export and Social security contribution. The variables which are not significant at 10% are Overall Government Score, Regulatory quality, Cashless payments, and Tax morality.

The multivariate regression analysis will therefore be done with the variables Tax, Unemployment, PR, GDPcurrent, Import, Export and SSC. This is also unfortunate, as no social-economic variables are in the regression.

MULTIVARIATE REGRESSION ANALYSIS

The next step is applying the multivariate regression analysis to the data acquired. This is done by regressing formula 4:

$$Y = \beta_0 + \beta_1 * Tax + \beta_2 * Unemployment + \beta_3 * PR + \beta_4 * GDPcurrent + \beta_5 * Import + \beta_6 * Export + \beta_7 * SSC + \in \quad (4)$$

The formula is regressed in STATA (version 13.0) and gave the following results (table 10):

Linear regression

SE	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tax	-.218	.138	-1.58	.131	-.509	.072	
Unemployment	.213	.114	1.86	.08	-.029	.454	*
PR	-.097	.099	-0.97	.344	-.306	.113	
GDPcurrent	-9.72e-07	3.95e-07	-2.46	.025	-1.81e-06	-1.39e-07	**
Import	7.60e-06	6.54e-06	1.16	.262	-6.21e-06	.0000214	
Export	-6.40e-06	8.70e-06	-0.74	.472	-.0000248	.000012	
SSC	.388	.34	1.14	.269	-.329	1.105	
Constant	.167	.071	2.34	.031	.017	.317	**

Mean dependent var		0.083	SD dependent var		0.008
R-squared		0.945	Number of obs		25
F-test		41.736	Prob > F		0.000
Akaike crit. (AIC)		-225.945	Bayesian crit. (BIC)		-216.194

*** $p<.01$, ** $p<.05$, * $p<.1$

Table 10: Regression results (imported from STATA 13.0).

The results present a few significant variables. The significant variables are Unemployment, GDPcurrent and the Constant. The significance level of 10% will be used. As the unemployment rate increases with 1%, the size of the shadow economy increases with 0.213%. For GDP, there is a negative relation with the size of the shadow economy. So, when GDP per capita increases with 1$, the size of the shadow economy decreases with -9.72e-07%. From the results it is seen that also tax, PR, and Export have a negative relation. So, for example, when taxes increase with 1%, the size of the shadow economy decreases with 0.218%. As mentioned before, Tax presents the tax revenue in percentage of GDP. The interpretation of this variable is in percentage; however, the variable is not logarithmic. This is due to the value of the used data is in percentage. So, the interpretation uses the official values of the determinants. The variables Tax, Unemployment, PR, and SSC are variables which are in percentage, therefore will be interpreted as percentages.

To check the obtained results for multicollinearity, the VIF will be used, which gave the following results (table 11):

Variance inflation factor

	VIF	1/VIF
Import	113.193	.009
Export	112.362	.009
GDPcurrent	70.367	.014
Unemployment	14.725	.068
Tax	11.739	.085
PR	8.493	.118
SSC	6.499	.154
Mean VIF	**48.197**	.

Table 11: VIF results (imported from STATA 13.0).

As seen in table 11, there are high levels of multicollinearity detected in every variable, except for PR and SSC. The next step is to remove or reduce the amounts of multicollinearity. The first step that can be taken is to add two variables together which are highly correlated.

Matrix of correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Tax	1.000						
(2) Unemployment	-0.723	1.000					
(3) PR	0.269	-0.114	1.000				
(4) GDPcurrent	-0.206	-0.096	-0.867	1.000			
(5) Import	-0.207	-0.023	-0.856	0.984	1.000		
(6) Export	-0.193	0.011	-0.918	0.973	0.980	1.000	
(7) SSC	0.372	-0.355	0.748	-0.758	-0.820	-0.826	1.000

Table 12: Correlation table (imported from STATA 13.0).

As seen in table 12, the correlation between Import and Export is very high. These two variables are added together which creates the new variable *tradebalance.* This variable is generated with the following command: gen tradebalance = Export – Import.

Table 13 represents the new correlation between the variables, now that tradebalance is added. There can be concluded that there is still high correlation between GDPcurrent and PR; tradebalance and GDPcurrent.

Matrix of correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)
(1) Tax	1.000					
(2) Unemployment	-0.723	1.000				
(3) PR	0.269	-0.114	1.000			
(4) GDPcurrent	-0.206	-0.096	-0.867	1.000		
(5) SSC	0.372	-0.355	0.748	-0.758	1.000	
(6) tradebalance	0.202	0.112	0.497	-0.803	0.627	1.000

Table 13: Correlation matrix with new variable tradebalance (imported from STATA 13.0).

Now let's try the regression again, however, this time with the new variable. The following results are found:

Linear regression

SE	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tax	-.205	.107	-1.91	.072	-.429	.02	*

	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Unemployment	.222	.098	2.26	.037	.015	.428	**
PR	-.104	.085	-1.22	.238	-.284	.075	
GDPcurrent	-9.15e-07	1.88e-07	-4.86	0	-1.31e-06	-5.20e-07	***
SSC	.367	.307	1.20	.247	-.277	1.012	
tradebalance	-7.737e-06	6.23e-06	-1.18	.252	-.0000205	5.71e-06	
Constant	.171	.066	2.60	.018	.033	.308	**

Mean dependent var	0.083	SD dependent var	0.008
R-squared	0.945	Number of obs	25
F-test	51.468	Prob > F	0.000
Akaike crit. (AIC)	-227.905	Bayesian crit. (BIC)	-219.373

*** $p<.01$, ** $p<.05$, * $p<.1$

Table 14: Regression with new variable tradebalance (imported from STATA 13.0).

As seen in the regression table 14, the t-values are higher compared to the last regression. Also, the variable Tax is now significant. The variable tradebalance shows a negative relation with the size of the shadow economy, as the tradebalance increases with 1$, the size of the shadow economy decreases with 7.737e-06%. This is as expected as trade is transparent and therefore easier to tax. Avoiding taxes by individuals is more difficult. The coefficient is very low, so the impact of this determinant, just as GDPcurrent, is not much. Tradebalance is however not significant. Furthermore, the R-squared from both regression is 94.5%. So, the model explains 94.5% of the variance in the size of the shadow economy. To check how high the levels of multicollinearity are with this regression, the new VIF values are estimated:

Variance inflation factor

	VIF	1/VIF
GDPcurrent	16.901	.059
Unemployment	11.47	.087
tradebalance	10.997	.091
Tax	7.473	.134
PR	6.653	.15
SSC	5.595	.179
Mean VIF	9.848	.

Table 15: VIF values of regression with new variable tradebalance (imported from STATA 13.0).

Despite the new variable, there are still too high levels of multicollinearity, especially with GDPcurrent, Unemployment and tradebalance (shown in table 15). The next possible step is to remove highly correlated variables in a way of trial and error. The three variables which multiple regression will be done are GDPcurrent, Unemployment and tradebalance.

After regressing multiple options, the following regression (table 16) with both GDPcurrent and Unemployment came out as the best one with the highest R-squard and lowest VIF value, while maintaining the highest number of variables. The "trial-and-error regressions" of the three variables can be found in appendix D with explanation.

Linear regression

SE	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tax	-.28	.087	-3.21	.005	-.462	-.097	***
Unemployment	.156	.082	1.91	.072	-.015	.327	*
PR	-.045	.07	-0.64	.527	-.192	.102	
GDPcurrent	-7.37e-07	1.14e-07	-6.45	0	-9.76e-07	-4.98e-07	***
SSC	.217	.282	0.77	.451	-.374	.808	
Constant	.149	.064	2.34	.03	.016	.282	**

Mean dependent var	0.083	SD dependent var	0.008
R-squared	0.941	Number of obs	25
F-test	60.209	Prob > F	0.000
Akaike crit. (AIC)	-228.031	Bayesian crit. (BIC)	-220.717

*** p<.01, ** p<.05, * p<.1

Table 16: Regression with lowest VIF value (imported from STATA 13.0).

This regression shows that Tax, Unemployment, GDPcurrent and the constant are significant. From the regression the following estimated regression equation can be constructed (formula 5):

$$\hat{Y} = 0.149 + -0.28 * Tax + 0.156 * Unemployment + -0.045 * PR + -7.37e - 07 * GDPcurrent + 0.217 * SSC + e \tag{5}$$

The model explains 94.1% of the variance of the dependent variable. This is 0.4% lower than the regressions done before. According to formula 5, there is always a constant base of 0.149% in size of the shadow economy. When tax revenue increases with 1% (Tax), the size of the shadow economy decreases with 0.28% as there are more revenues paid. When more revenues are paid, this means that less individuals are active in the shadow economy (**significant**). When the unemployment rate increases with 1%, the size of the shadow economy also increases with 0.156%. This can be explained due to more unemployment individuals are active in the shadow economy, as they can easily hide their activities more than employed (**significant**). The participation rate provides evidence for this because when the participation rate increases with 1%, the size of the shadow economy decreases with 0.045%. This represents the active population of working age which are employed. Therefore, when the number of employed individuals increases, the size of the shadow economy decreases. The change in GDP per capita is very small compared to the other variables. This can be explained as many factors influences the GDP (Kira, 2013). When GDP increases with 1$ the size of the shadow economy decreases with 7.37e-07%. This is due the fact that when more individuals work in the official economy, the size of the shadow economy decreases, and therefore, more value will be added to the GDP (**significant**). Lastly, there is a positive relation between the size of the shadow economy and the social security contribution. As individuals need to pay more social security contribution, the size of the shadow economy increases, because they try to avoid these payments. So, when the social security contribution payments increase with 1%, the size of the shadow economy increases with 0.217%.

To conclude, according to the empirical model used, the significant determinants which have influence on the size of the shadow economy of the United States are taxes, the amount of unemployment, and GDP per capita. Taxes and GDP per capita both have a negative influence on the size of the shadow economy of the United States and the amount individuals which are unemployment have positive effect on the size of the shadow economy. An overview is given in figure 6 (see next page). This resulted in the following conclusion on the hypothesis:

Variable/determinant	Expected linear relation	Conclusion on hypothesis
Unemployment rate	Positive	Proofed correct
Tax morality	Positive	Further research needed as no conclusion could be stated
Social security contribution	Positive	Further research needed as variable was insignificant
Tax burden	Negative	Proofed correct
Participation rate	Negative	Further research needed as variable was insignificant

GDP per capita	Negative	Proofed correct
Imports	Negative	Further research needed as no conclusion could be stated
Exports	Negative	Further research needed as no conclusion could be stated
Regulatory quality	Negative	Further research needed as no conclusion could be stated
Overall government score	Negative	Further research needed as no conclusion could be stated
Cashless payments	Negative	Further research needed as no conclusion could be stated

Table 17: Conclusion on hypothesis stated.
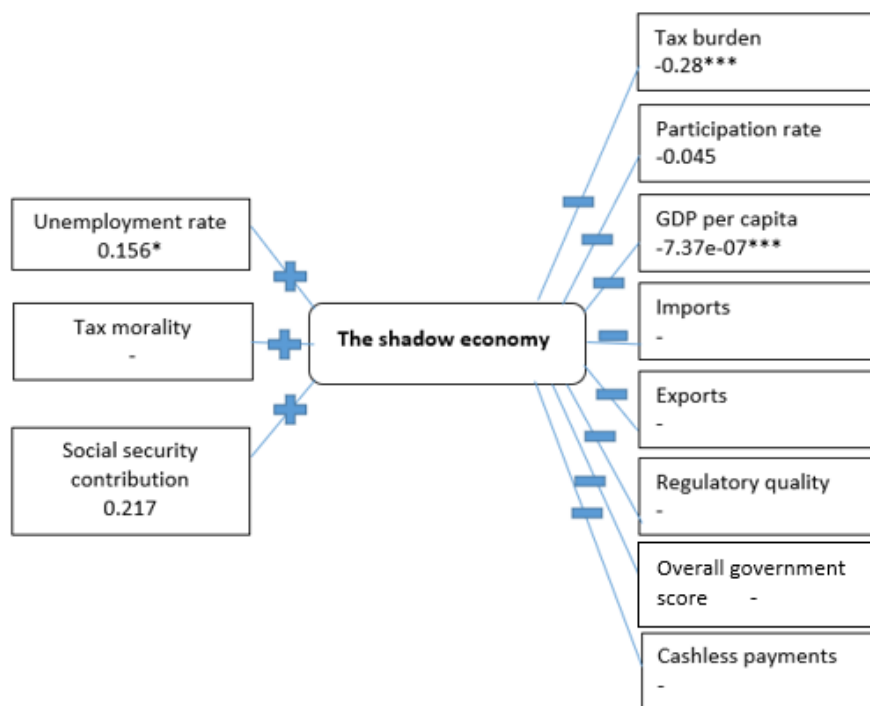
Figure 6: Outcome regression.

Discussion & Conclusion

The research conducted in the thesis contributes to the knowledge on which variables influence the size of the shadow economy in the United States and how strong this influence is. As the shadow economy is seen as a difficult topic due to the broadness of the concept, it is hard to determine the correct approach to measure the influence of the determinants on the shadow economy. It was therefore necessary to state the definition of the concept which provided a basis for the research conducted in this thesis. The following definition is constructed based on previous literature work: *The shadow economy covers all economic activity, legal and illegal, which individuals try to conceal from the government. This behaviour is caused due economic and social-economic factors like a high tax burden, high social security contribution, overburdening regulations, and weak institutional quality.*

35

The growth of the shadow economy can be seen as cycles based on increasing taxes and other payments, which decrease the government revenues and causes lower quality over public goods. Individuals are less willing to pay taxes due to low quality of public goods and enter the shadow economy. The shadow economy is seen as a very negative concept with negative side effects for firms and individuals. However, it also brings positivity as it creates more work opportunities and money is directly spend in the economy.

Countries can be specified in the categories developing, transition or developed. For every category there are different reasons why individuals enter the shadow economy. Furthermore, every category of country has different types of shadow activity which are more present.

The literature review and theoretical framework presented a basis to determine a decent methodology for the empirical research conducted. For this thesis, first, the Pearson correlation was applied to determine which variables have the strongest and significant relation with the dependent variable, the size of the shadow economy of the United States. The significant variables are Tax, Unemployment, PR, GDPcurrent, Import, Export, and SSC. The variables Tax, Unemployment, PR, and SSC have a medium correlation; while GDPcurrent, Import, and Export have a strong correlation with the size of the shadow economy according to the Pearson correlation coefficient.

Comparing this result to the research conducted by Gasparênienè *et al* (2016), some differences can be found. They applied the research on the variables total labour force, imports, participation rate, GDP per capita, and tax rate. It is interesting to compare this as the research in this thesis is based on their method. However, their method is extended with more determinants which are found to be of importance on the size of the shadow economy according to previous literature work. A comparison is shown in table 17.

Own research	Strength correlation	Gasparênienè *et al* (2016)	Strength correlation
Tax	Medium	Tax rate	Medium
Unemployment	Medium	Total labour force	Strong
PR	Medium	Participation rate	Strong
GDPcurrent	Strong	GDP per capita	Medium
Import	Strong	Import	Strong
Export	Strong		
SSC	Medium		

Table 17: Comparison table of strength of the correlation with dependent variable.

The similar variables used in both researches are Tax, Participation rate, GDP per capita and Import. Only the strength of the correlation between Tax and Import is in both researches the same. This can be explained by the usage of different data for a different country, as they applied their research on Ukraine. Another factor influencing this difference is the range of the research period. In this thesis, the research period is 1991 till 2015. In the research conducted by Gasparênienè *et al* (2016), a research period of 2003 till 2015 is applied.

To answer the research question of this thesis a multivariate regression analysis is applied with the variables which are significant according to the Pearson correlation. This approach achieves the research objectives stated in the research question of this thesis:

Which determinants influence the size of the shadow economy and what is the impact of these determinants for the United States in period of 1991 till 2015?

To conclude, the multivariate regression resulted in three variables having a significant influence on the size of the shadow economy of the United States. The determinants influencing the size of the shadow economy are GDP per capita, number of unemployed individuals, and taxes. Taxes influences the size of the shadow economy of the United States the most. This was also seen as the most important driver for the growth of the shadow economy by Enste (2018), Aigner *et al* (1986) and Enste & Schneider (2000). Following up taxes is the unemployment rate, and lastly the GDP per capita. The variables participation rate and social security contribution are not significant.

The relations between the independent variables and the dependent variable were as expected (see figures 4 and 6 for comparison). For the variable taxes was a negative relation expected as the data represents tax burden (tax revenue as % of GDP). This is because as taxes increase, more individuals enter the shadow economy, as the willingness to pay taxes decreases, thus lower tax revenues are available for the government (Laffer, 2004; Kelmanson *et al,* 2019 & Schneider & Enste, 2002). This can explain the negative relation between taxes and the size of the shadow economy of the United States, as this is also supported by the Laffer curve. It might have been better to take the tax rate on personal income, which would have provided a positive relation with the size of the shadow economy. This would have been interesting to provide proof for this relationship and thus the effect of an increase in taxes on the size of the shadow economy.

From research conducted by Gasparênienè *et al* (2016), Kanniainen *et al* (2004) and Medina & Schneider (2018), a positive relation between taxes and the size of the shadow economy is found. However, Kelmanson *et al* (2019) and Ion & Alexandru (2009) found a negative relation in their research conducted. As both used the variable tax burden, comparable with the research done in this thesis. Kelmanson *et al* (2019) used the fiscal freedom index from Heritage Foundation. Ion & Alexandru (2009) also used the tax-to-GDP ratio. It therefore highly depends on what type of data is used. When using tax revenue, a negative relation is found and when using the tax rate, a positive relation is found. Referencing back to figure 2, this shows the positive and negative relation.

For the variable unemployment rate, both in the research conducted by Kanniainen *et al* (2004) and Medina & Schneider (2018), a positive relation is found, comparable with the result in this thesis. This is due to lack of opportunity (mostly in developing countries) and individuals who are unemployed can easier hide their illegal income. Therefore, as more individuals are unemployed, there is a higher share of shadow economy.

The variable GDP per capita has a negative relation with the size of the shadow economy according to Gasparênienè *et al* (2016), Kanniainen *et al* (2004) and Kelmanson *et al* (2019). This can be explained as more economic activity and productivity is in the official economy, less is going to the shadow economy. It could also be the other way around. When more individuals enter the shadow economy, less value is added to the official GDP.

The variables which were not significant according to the Pearson correlation coefficient were Overall Government Score, Regulatory quality, Cashless payments, and Tax morality. This is unfortunate as the goal of this thesis is an interdisciplinary approach, which is reached by these variables. When viewing the data used (see appendix B), these are the only variables with some years of data missing. This is most likely the reason for not having a significant correlation with the size of the shadow economy of the United States. This is therefore a limitation in the thesis. If more data would have been available, these variables possibly would have been significant and could have been used in the multivariate regression. This was also mentioned by Weedmark (2018) as a disadvantage of using the multivariate regression. When a too small sample is used, the outcomes might not be correct and unreliable.

When reviewing my literature and theoretical part, an improvement could be to find more recent data on the size of the shadow economy for developing, transition and developed countries. There is the possibility that the situation in size of the shadow economy has changed in the past two/three decades.

Furthermore, due to the high levels of multicollinearity, doing the multivariate regression was difficult. This problem could be solved if a larger research period was used, thus more data. Furthermore, while dealing with the multicollinearity, another possible solution is using the Principal Component Analysis (PCA). This method is useful to remove multicollinearity and to use all the variables significant according to the Pearson correlation coefficient. For the current research conducted, the used solution was dropping variables which caused high levels of multicollinearity. A way to keep these variables and produce decent results is using the PCA. The steps for applying the PCA approach are presented in appendix E on the data for the research in this thesis.

It would also have been interesting if I estimated the size of the shadow economy of the United States myself by using the MIMIC model, with my own variables and data. Now some bias could exist as other variables and data is used to determine the size of the shadow economy of the United States. As mentioned before, the data was retrieved from the research conducted by Medina & Schneider (2018). However, as seen before, the results from the empirical research in this thesis are comparable with the literature presented and therefore, the research conducted in this thesis is decent enough to state a few accurate conclusions.

Furthermore, with an eye on the future, possible future research can be on the relation between the digital world and the shadow economy. As the digital world grows rapidly and more illegal techniques to hide shadow activities are constructed it is interesting and important to keep track of these developments. Therefore, it was also interesting to add the variable cashless payments to the model to see the effect of this variable on the size of the shadow economy of the United States.

Another interesting possible future research is conducting a large-scale research to determine which determinants influence the size of the shadow economy of developing, transitioning and developed countries. From literature work became clear that different shadow activities are presented in these countries. A more detailed research would provide an informative overview for economists, politicians, and organisation. Reducing the size of the shadow economy might be easier if there is an understanding on which determinants and activities are more present.

Other future research should be a more specific focus on the relation between overburdening regulations and the size of the shadow economy, as this determinant is multiple times stressed in the literature as one of the important reasons for individuals to enter the shadow economy. This implies further interdisciplinary research and shows the importance of this research for also the world of politics. It is important to determine for policymakers and decisionmakers to know when the regulations are overburdening. This is linked with the complexity of a tax system as this has according to Schneider & Neck (1993) an impact on individuals entering the shadow economy to avoid taxes. Yet not much attention has been paid to this matter, nor I found many literature on this aspect. Therefore, further research on this topic might also be beneficial for politicians and policymakers.

Appendix

A. APPROACHES USED TO MEASURE SIZE OF THE SHADOW ECONOMY

Type of approach	Name of the approach	Definition	Strength and weaknesses
Direct approach	Tax auditing	A measure which estimates the size of the shadow economy by using auditing measurements of undeclared taxable income (Schneider & Enste, 2002).	See Survey.
Direct approach	Survey	This measure is on micro level which provides knowledge on the actual perception and participation of individuals in the shadow economy activities (Medina & Schneider, 2018; Schneider & Enste, 2002).	The advantage of these methods is actual already stated. These measures give detailed information on the structure of the shadow economy and direct perceptions of individuals (Kelmanson *et al,* 2019; Aigner *et al,* 1986). However, these measures are on micro level and will not cover all the shadow economy. These measures also provide only point estimates, which makes it unable to provide estimates over a time period. This data is thus only on micro level and might be not representative for other countries or cross-country analysis (Kelmanson *et al,* 2019; Aigner *et al,* 1986). Furthermore, the surveys can have issues with self-reporting bias and a sampling error (Fleming *et al,* 2000).
Indirect approach	Discrepancy between national account statistics	The difference between the income and expenditure statistics	The assumptions made with this measure are sensitive to elasticity, the base year, and the velocity of money and the expenditure side of the national statistics are determined with no errors. However, these errors can have flaws which make them not reliable measures

			(Aigner *et al,* 1986; Medina & Schneider, 2018; Schneider & Enste, 2002; Kelmanson *et al,* 2019).
Indirect approach	Labour force statistics difference	Difference between the officially measured labour force and the actual labour force participation (Schneider & Enste, 2002).	The assumption stated with this measure is the total labour force participation is assumed to be constant, so when there is a decrease in the official labour participation, this decrease is due to more individuals entering the shadow economy (Aigner *et al,* 1986; Medina & Schneider, 2018; Schneider & Enste, 2002). This view seems a little but too vague. This decrease in labour participation can have multiple reasons. For example, individuals who retire, education or individuals finding a new job (Medina & Schneider, 2018). Aigner *et al* (1986) also addresses the fact that individuals can have a job in the official and unofficial economy at the same time. In this situation, the participation rate does not change, however the shadow economy still increases in size.
Indirect approach	Transaction approach	The amount of monetary transaction done in the economy (Schneider & Enste, 2002).	It is stated that there is a constant relation over time between the official GNP and the amount of transaction done within the economy. This is called the quantity theory of money. To determine what amount of the GNP belongs to the shadow economy, the total nominal GNP will be subtracted by the officially measured GNP. For this approach, a base year needs to be determined when there was no shadow

			economy and thus constant overtime and the relation between nominal transactions and total nominal GDP is perceived as normal. This also one of the main weaknesses of this approach as it is not realistic that there is a base year without shadow economy (Aigner *et al*, 1986).
Indirect approach	Currency demand approach	This approach takes the assumption that the transactions within the shadow economy are done in cash to stay untraceable by the government. This increases the demand for currency. To measure this difference in currency use, the situation with a low tax burden will be compared to the actual tax burden (Aigner *et al,* 1986; Medina & Schneider, 2018; Schneider & Enste, 2002; Fleming *et al,* 2000).	However, Aginer *et al* (1986) and Fleming *et al* (2000) stress the fact that a large flaw in this approach is not considering that many payments in the shadow economy can also be done using cheques, credit cards or other payments.
Indirect approach	Electricity approach	This approach thus compares the growth in the official GDP with the growth in the consumption of electricity. The difference between the indicators shows the size of shadow economy (Medina & Schneider, 2018; Schneider & Enste, 2002; Fleming *et al,* 2000).	However, electricity does not only change due to change in economic activity. This can also be caused by seasonal changes and the structure of the economy evolves (Arora & Lieskovsky, 2014). Some have concluded, according to Arora & Lieskovsky (2014), that it is no longer an efficient approach for measuring, as the economic activity and energy consumption are not in the same line. Furthermore, not all shadow activities make

			use of electricity (Medina & Schneider, 2018).

B. DATA USED

Year	SE	Taxes	Unemployment	PR	GDP	Imports	Exports	OGS	RQ	CP	SSC	TM
1991	10.12%	10.32%	6.9%	66.2%	$ 24,342.26	$ 623.54	$ 594.931				6.74%	
1992	10.00%	9.98%	7.5%	66.4%	$ 25,418.99	$ 667.79	$ 633.053				6.71%	
1993	9.69%	10.31%	6.9%	66.3%	$ 26,387.29	$ 719.97	$ 654.799				6.68%	
1994	9.23%	10.62%	6.1%	66.6%	$ 27,694.85	$ 813.42	$ 720.937				6.69%	
1995	8.91%	10.98%	5.6%	66.6%	$ 28,690.88	$ 902.57	$ 812.810	76.7			6.68%	
1996	8.90%	11.37%	5.4%	66.8%	$ 29,967.71	$ 963.97	$ 867.589	76.7	1.59		6.61%	
1997	8.23%	11.83%	4.9%	67.1%	$ 31,459.13	$ 1,055.77	$ 953.803	75.6			6.59%	
1998	8.00%	12.33%	4.5%	67.1%	$ 32,853.68	$ 1,115.69	$ 952.979	75.4	1.67		6.65%	
1999	7.82%	12.26%	4.2%	67.1%	$ 34,513.56	$ 1,248.61	$ 992.778	75.5			6.63%	
2000	7.60%	12.97%	4.0%	67.1%	$ 36,334.91	$ 1,471.31	$ 1,096.255	76.4	1.76	72.5	6.66%	63.0%
2001	8.01%	11.86%	4.7%	66.8%	$ 37,133.24	$ 1,392.57	$ 1,024.636	79.1		75.6	6.69%	65.0%
2002	8.54%	9.86%	5.8%	66.6%	$ 38,023.16	$ 1,424.14	$ 998.741	78.4	1.6	78.8	6.62%	
2003	8.40%	9.39%	6.0%	66.2%	$ 39,496.49	$ 1,539.30	$ 1,036.177	78.2	1.6	82	6.53%	50.0%
2004	8.43%	9.54%	5.5%	66.0%	$ 41,712.80	$ 1,796.71	$ 1,177.631	78.7	1.57	85.5	6.50%	50.0%
2005	7.86%	10.68%	5.1%	66.0%	$ 44,114.75	$ 2,026.42	$ 1,305.225	79.9	1.61	89.7	6.38%	51.0%
2006	7.47%	11.31%	4.6%	66.2%	$ 46,298.73	$ 2,243.54	$ 1,472.613	81.2	1.64	93.6	6.32%	48.0%
2007	8.00%	11.31%	4.6%	66.0%	$ 47,975.97	$ 2,379.28	$ 1,660.854	81.2	1.49	97.2	6.28%	53.0%
2008	7.76%	10.32%	5.8%	66.0%	$ 48,382.56	$ 2,560.14	$ 1,837.055	81	1.53	100.7	6.32%	52.0%
2009	9.18%	7.92%	9.3%	65.4%	$ 47,099.98	$ 1,978.45	$ 1,581.996	80.7	1.4	110.4	6.25%	46.0%
2010	8.71%	8.60%	9.6%	64.7%	$ 48,466.66	$ 2,360.18	$ 1,846.281	78	1.45	114.9	6.12%	48.0%
2011	8.23%	9.57%	8.9%	64.1%	$ 49,882.56	$ 2,682.46	$ 2,102.995	77.8	1.46	119.4	5.45%	50.0%
2012	7.83%	9.80%	8.1%	63.7%	$ 51,602.93	$ 2,759.85	$ 2,191.280	76.3	1.3	123.8	5.44%	46.0%
2013	7.66%	10.52%	7.4%	63.2%	$ 53,106.54	$ 2,764.21	$ 2,273.428	76	1.27	130.4	6.14%	50.0%
2014	7.04%	10.95%	6.2%	62.7%	$ 55,049.99	$ 2,879.36	$ 2,371.704	75.5	1.28	137	6.14%	52.0%
2015	7.00%	11.21%	5.3%	62.7%	$ 56,863.37	$ 2,792.43	$ 2,265.862	76.2	1.26	143.5	6.17%	51.0%
2016		10.89%	4.9%	62.8%	$ 58,021.40	$ 2,739.69	$ 2,227.174	75.4	1.5	139.9	6.16%	57.0%
2017		11.71%	4.4%	62.9%	$ 60,109.66	$ 2,930.09	$ 2,374.560	75.1	1.63	163.3	6.17%	51.0%
2018		9.90%	3.9%	62.9%	$ 63,064.42	$ 3,138.16	$ 2,528.704	75.7	1.62	174.2	6.07%	45.0%
2019		10.02%	3.7%	63.1%	$ 65,279.53	$ 3,125.22	$ 2,514.751	76.8	1.35		6.10%	45.0%
2020		9.97%	8.1%	61.7%	$ 63,543.58	$ 2,771.96	$ 2,127.128	76.6				46.0%

C. PEARSON CORRELATION COEFFICIENTS

Pearson correlations

Variables	(SE)	(Tax)
SE	1.000	
Tax	-0.412**	1.000

*** p<0.01, ** p<0.05, * p<0.1

Pearson correlations

Variables	(SE)	(Unemployment)
SE	1.000	
Unemployment	0.404**	1.000

*** p<0.01, ** p<0.05, * p<0.1

Pearson correlations

Variables	(SE)	(PR)
SE	1.000	
PR	0.412**	1.000

*** p<0.01, ** p<0.05, * p<0.1

Pearson correlations

Variables	(SE)	(GDPcurrent)
SE	1.000	
GDPcurrent	-0.731***	1.000

*** p<0.01, ** p<0.05, * p<0.1

Pearson correlations

Variables	(SE)	(Import)
SE	1.000	
Import	-0.720***	1.000

*** p<0.01, ** p<0.05, * p<0.1

Pearson correlations

Variables	(SE)	(Export)
SE	1.000	
Export	-0.675***	1.000

*** p<0.01, ** p<0.05, * p<0.1

Pearson correlations

Variables	(SE)	(OGS)
SE	1.000	
OGS	0.196	1.000

*** p<0.01, ** p<0.05, * p<0.1

Pearson correlations

Variables	(SE)	(RQ)
SE	1.000	
RQ	0.267	1.000

*** p<0.01, ** p<0.05, * p<0.1

Pearson correlations

Variables	(SE)	(CP)
SE	1.000	
CP	-0.367	1.000

*** p<0.01, ** p<0.05, * p<0.1

Pearson correlations

Variables	(SE)	(SSC)
SE	1.000	
SSC	0.412**	1.000

*** p<0.01, ** p<0.05, * p<0.1*

Pearson correlations

Variables	(SE)	(TM)
SE	1.000	
TM	-0.263	1.000

*** p<0.01, ** p<0.05, * p<0.1*

D. TRAIL-AND-ERROR REGRESSIONS

Trail-and-error regression 1:

This is the regression presented as the most optimal one in the text. In this regression the variable tradebalance is left out. With a R-squared of 94.1%, lowest VIF values and for almost every variable higher t-values.

Linear regression

SE	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tax	-.28	.087	-3.21	.005	-.462	-.097	***
Unemployment	.156	.082	1.91	.072	-.015	.327	*
PR	-.045	.07	-0.64	.527	-.192	.102	
GDPcurrent	0	0	-6.45	0	0	0	***
SSC	.217	.282	0.77	.451	-.374	.808	
Constant	.149	.064	2.34	.03	.016	.282	**

Mean dependent var	0.083	SD dependent var		0.008
R-squared	0.941	Number of obs		25
F-test	60.209	Prob > F		0.000
Akaike crit. (AIC)	-228.031	Bayesian crit. (BIC)		-220.717

*** p<.01, ** p<.05, * p<.1*

Variance inflation factor

	VIF	1/VIF
Unemployment	7.8	.128
GDPcurrent	6.103	.164
Tax	4.857	.206
SSC	4.642	.215
PR	4.387	.228
Mean VIF	5.558	.

Trail-and-error regression 2:

In this regression, GDPcurrent is left out. The relation between the independent variables and dependent variable is the same negative or positive relation, except for PR. This is interesting to see, as concluded from the literature research conducted, this relation is negative. In this regression the relation is positive, which seems odd. This makes the regressed model doubtful. Furthermore, the R-squared for this model is 87.3%, which is lower than the first regression. This model thus explains less

of the variance of the dependent variable than the other regression. Also, the VIF is higher than the first regression, which implies more multicollinearity.

Linear regression

SE	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tax	-.322	.154	-2.09	.051	-.644	.001	*
Unemployment	.162	.144	1.12	.276	-.14	.464	
PR	.225	.077	2.92	.009	.064	.386	***
SSC	.416	.454	0.92	.37	-.533	1.366	
tradebalance	0	0	3.04	.007	0	0	***
Constant	-.06	.067	-0.90	.379	-.2	.08	
Mean dependent var		0.083	SD dependent var			0.008	
R-squared		0.873	Number of obs			25	
F-test		26.037	Prob > F			0.000	
Akaike crit. (AIC)		-208.949	Bayesian crit. (BIC)			-201.636	

*** p<.01, ** p<.05, * p<.1*

Variance inflation factor

	VIF	1/VIF
Unemployment	11.289	.089
Tax	7.094	.141
SSC	5.589	.179
tradebalance	3.971	.252
PR	2.471	.405
Mean VIF	6.083	.

Trail-and-error regression 3:

This is the regression with leaving out the variable Unemployment. This model explains 92.9% of the variance of the dependent variable. This is less than the first regression and larger than the second regression. However, this time the SSC has a negative relation, although a positive is expect according to the literature research. As the payment for the social security contribution increase, more individuals try to avoid these payments and enter the shadow economy. This relation is therefore odd. Furthermore, the VIF value is also higher than the two regressions presented earlier.

Linear regression

SE	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Tax	-.424	.049	-8.68	0	-.527	-.322	***
PR	-.108	.094	-1.14	.268	-.305	.09	
GDPcurrent	0	0	-4.18	.001	0	0	***
SSC	-.091	.254	-0.36	.723	-.622	.439	
tradebalance	0	0	0.10	.919	0	0	
Constant	.24	.064	3.76	.001	.106	.374	***
Mean dependent var		0.083	SD dependent var			0.008	
R-squared		0.929	Number of obs			25	
F-test		49.964	Prob > F			0.000	
Akaike crit. (AIC)		-223.670	Bayesian crit. (BIC)			-216.357	

*** p<.01, ** p<.05, * p<.1*

Variance inflation factor

	VIF	1/VIF
GDPcurrent	16.635	.06
tradebalance	7.478	.134

PR	6.651	.15
SSC	3.144	.318
Tax	1.284	.779
Mean VIF	7.038	.

E. PRINCIPAL COMPONENT ANALYSIS (PCA)

Here are the steps shown to perform a Principal Component Analysis. It will not be discussed in detail as it is not the focus of the thesis. However, as this is a solution against multicollinearity, without dropping the variables, this is approach would be optional for removing the issues found in the results.

Component loadings
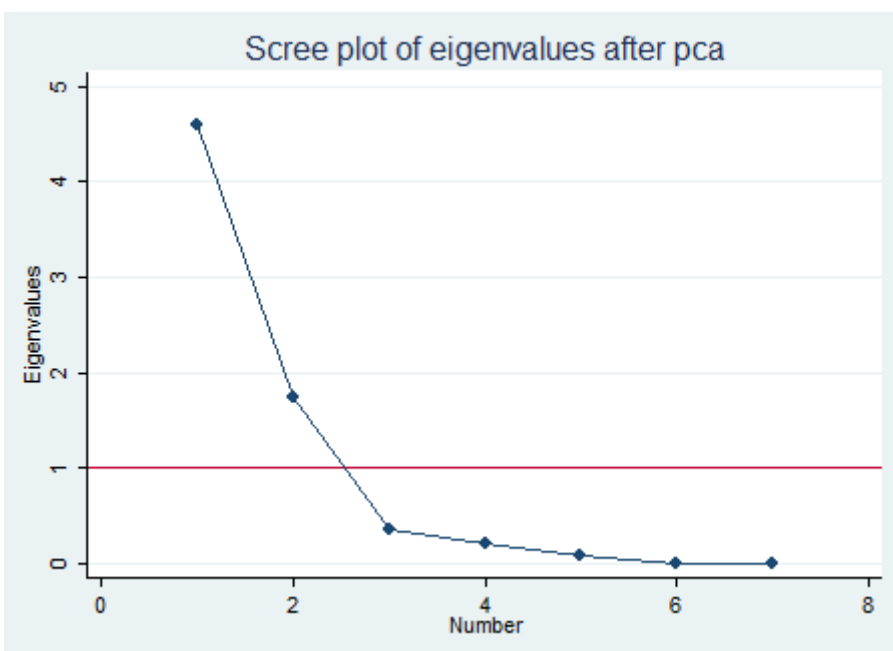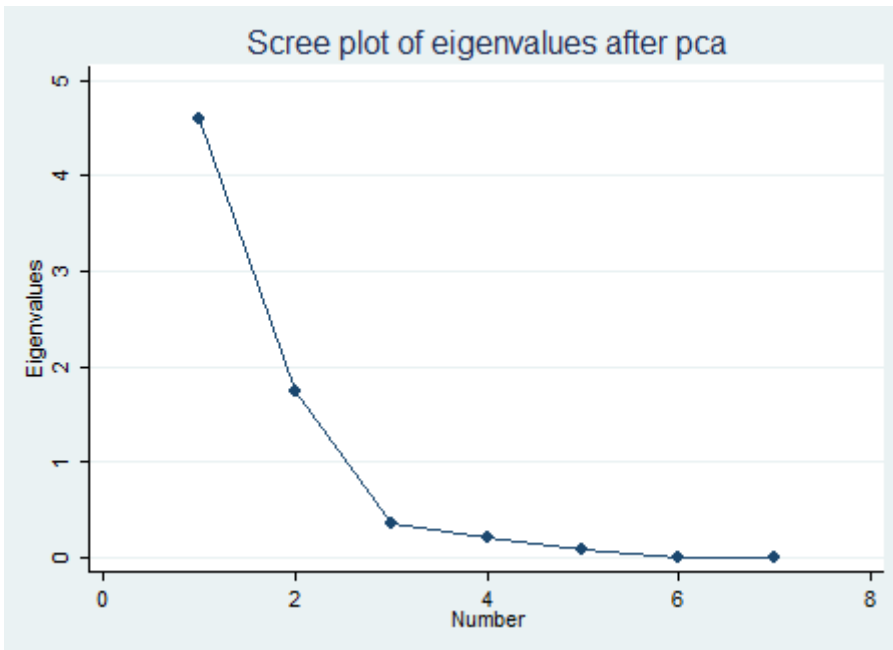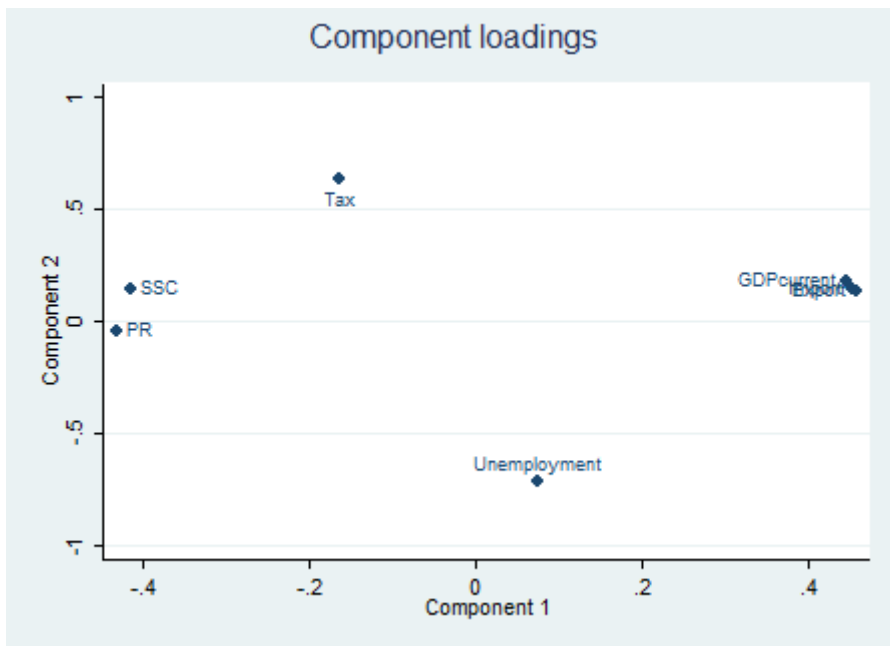
. pca Tax Unemployment PR GDPcurrent Import Export SSC

Principal components/correlation

Number of obs	=	29
Number of comp.	=	7
Trace	=	7

Rotation: (unrotated = principal)   Rho   =   1.0000

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	4.60542	2.86962	0.6579	0.6579
Comp2	1.7358	1.38308	0.2480	0.9059
Comp3	.352722	.145195	0.0504	0.9563
Comp4	.207527	.123651	0.0296	0.9859
Comp5	.0838758	.0761152	0.0120	0.9979
Comp6	.00776069	.000867358	0.0011	0.9990
Comp7	.00689334	.	0.0010	1.0000

Principal components (eigenvectors)

Variable	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6	Comp7	Unexplained
Tax	-0.1657	0.6353	0.6744	-0.1881	0.2457	-0.1090	0.0807	0
Unemployment	0.0743	-0.7124	0.4196	-0.2348	0.4874	-0.0941	0.0967	0
PR	-0.4309	-0.0414	0.0394	0.7893	0.3727	0.2024	0.0906	0
GDPcurrent	0.4445	0.1747	-0.2523	0.1417	0.2677	-0.3328	0.7112	0
Import	0.4511	0.1389	-0.0607	0.2427	0.3748	-0.3307	-0.6818	0
Export	0.4565	0.1299	0.0219	-0.0714	0.2275	0.8469	0.0154	0
SSC	-0.4146	0.1429	-0.5476	-0.4499	0.5471	0.0321	-0.0714	0

```
. pca Tax Unemployment PR GDPcurrent Import Export SSC, mineigen(1)

Principal components/correlation                  Number of obs     =          29
                                                  Number of comp.   =           2
                                                  Trace             =           7
      Rotation: (unrotated = principal)           Rho               =      0.9059
```

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	4.60542	2.86962	0.6579	0.6579
Comp2	1.7358	1.38308	0.2480	0.9059
Comp3	.352722	.145195	0.0504	0.9563
Comp4	.207527	.123651	0.0296	0.9859
Comp5	.0838758	.0761152	0.0120	0.9979
Comp6	.00776069	.000867358	0.0011	0.9990
Comp7	.00689334	.	0.0010	1.0000

Principal components (eigenvectors)

Variable	Comp1	Comp2	Unexplained
Tax	-0.1657	0.6353	.1729
Unemployment	0.0743	-0.7124	.0936
PR	-0.4309	-0.0414	.1419
GDPcurrent	0.4445	0.1747	.03697
Import	0.4511	0.1389	.02936
Export	0.4565	0.1299	.01114
SSC	-0.4146	0.1429	.1729

```
. rotate, varimax
```

Principal components/correlation

Number of obs   =   29
Number of comp.   =   2
Trace   =   7

Rotation: orthogonal varimax (Kaiser off)   Rho   =   0.9059

Component	Variance	Difference	Proportion	Cumulative
Comp1	4.50896	2.67671	0.6441	0.6441
Comp2	1.83226	.	0.2618	0.9059

Rotated components

Variable	Comp1	Comp2	Unexplained
Tax	-0.0464	0.6549	.1729
Unemployment	-0.0576	-0.7140	.0936
PR	-0.4312	0.0383	.1419
GDPcurrent	0.4690	0.0902	.03697
Import	0.4689	0.0538	.02936
Export	0.4725	0.0440	.01114
SSC	-0.3814	0.2165	.1729

Component rotation matrix

	Comp1	Comp2
Comp1	0.9831	-0.1833
Comp2	0.1833	0.9831

```
. estat loadings
```

Principal component loadings
    component normalization: sum of squares(column) = 1

	Comp1	Comp2
Tax	-.1657	.6353
Unemployment	.0743	-.7124
PR	-.4309	-.04138
GDPcurrent	.4445	.1747
Import	.4511	.1389
Export	.4565	.1299
SSC	-.4146	.1429

```
. predict pc1 pc2, score

Scoring coefficients
    sum of squares(column-loading) = 1
```

Variable	Comp1	Comp2
Tax	-0.1657	0.6353
Unemployment	0.0743	-0.7124
PR	-0.4309	-0.0414
GDPcurrent	0.4445	0.1747
Import	0.4511	0.1389
Export	0.4565	0.1299
SSC	-0.4146	0.1429

```
. estat kmo

Kaiser-Meyer-Olkin measure of sampling adequacy
```

Variable	kmo
Tax	0.3796
Unemployment	0.3407
PR	0.7893
GDPcurrent	0.6964
Import	0.7765
Export	0.7355
SSC	0.8996
Overall	0.7000

The Kaiser-Meyer-Olkin measure shows that this approach could be applied on the data gathered and used in this thesis. However, as the values of Tax and Unemployment are low, for these two variables it might not be perfect (Katchova, 2014).

References

"About us – the company behind the successful product". (n.d.). *Statista.* Retrieved from: https://www.statista.com/aboutus/

Aigner, D. J., Schneider, F., Ghosh, D. (1986). Me and my shadow: Estimiating the size of the U.S. underground economy from time series data. MRG working paper #M8615. *Department of Economics, University of Southern California.* Retrieved from: https://ageconsearch.umn.edu/record/296050/

Arora, V. & Lieskovsky, J. (2014). Electricity Use as an Indicator of U.S. Economic Activity. *U.S. Energy Information Administration.* Retrieved from: https://www.eia.gov/workingpapers/pdf/electricity_indicator.pdf

Awasthi, R. & Engelschalk, M. (2018). Taxation and the Shadow Economy: How the Tax System Can Stimulate and Enforce the Formalization of Business Activities. *World Bank Group*. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3153229

Balcioglu, H. B. (2018). Estimation of underground economy in North Cyprus: new results with MIMIC model approach and brief model comparison. *ECOFORUM, 7(1), n.d.*

de Best, R. (2021). Number of cashless payments in the U.S. 2000-2018, by type. *Statista.* Retrieved from: https://www.statista.com/statistics/580750/number-cashless-payments-usa-by-type/

Buehn, A. & Schneider, F. (2008). MIMIC Models, Cointegration and Error Correlation: An Application to the French Shadow Economy. *IZA.* Discussion paper No. 3306. Retrieved from: http://ftp.iza.org/dp3306.pdf

Coendet T. (2021). The Concept of Regulatory Arbitrage. *Economic Analysis of Law in European Legal Scholarship*, 11, 159-181.

Duffin, E. (2021). Do you consider the amount of federal income tax you have to pay as too high, about right, or too low? *Statista.* Retrieved from: https://www.statista.com/statistics/223274/americans-perceptions-on-their-income-taxes/

Enste, D. H. (2018). The shadow economy in industrial countries. *Iza World of Labour.* Retrieved from: https://wol.iza.org/articles/shadow-economy-in-industrial-countries/long

Enste, D. H. & Schneider, F. (2000). Shadow Economies: Size, Causes and Consequences. *Journal of Economic Literature,* 38(1), 77-114. https://www.researchgate.net/publication/4724886_Shadow_Economies_Size_Causes_and_Consequences

Fernando, S. (2009). Issues of E-Learning in Third World Countries. *Encyclopedia of Information Science and Technology*, Second Edition, pp. 2273-2277. IGI Global. http://doi:10.4018/978-1-60566-026-4.ch360

Fleming, M. H., Roman, J. & Farrell, G. (2000). The Shadow Economy. *Journal of International Affairs,* 53(2), 387-409.

"Friedrich Schneider". (n.d.). *World Economic Journal.* Retrieved from: https://www.world-economics-journal.com/Authors/Friedrich%20Schneider.author?AID=Friedrich%20Schneider

"Friedrich Schneider". (n.d.). *IZA.* Retrieved from: https://www.iza.org/person/206/friedrich-schneider

Gasparênienè, L., Remeikienè, R. & Heikkila, M. (2016). Evaluation of the impact of shadow economy determinants: Ukrainian case. *Intellectual Economics*, 10(2), 108-113.

"GDP per capita (current US$) – United States". (2021). *The World Bank.* Retrieved from: https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?end=2020&locations=US&start=1991&view=chart

Global Standards Proportionality (GSP) Working Group. (2018). GSP WG 2018 Member Survey, Global standards and financial inclusion: shadow banking. *Afi.* Retrieved from: https://www.afi-global.org/publications/gsp-wg-2018-member-survey-global-standards-and-financial-inclusion-shadow-banking/

Golic, Z. (2019). Problem of Financing Women Entrepreneurs: Experience of Women Entrepreneurs in Post-Conflict Bosnia and Herzegovina. *Women Entrepreneurs and Strategic Decision Making in the Global Economy*, pp. 278-304. IGI Global. http://doi:10.4018/978-1-5225-7479-8.ch014

"Gross domestic product (GDP)". (2021). *OECD Data*. Retrieved from: https://data.oecd.org/gdp/gross-domestic-product-gdp.htm

Ion, D. & Alexandru, A. (2009). Estimating the size of the shadow economy in Japan: A structural model with latent variables. *The Bucharest Academy of Economic Studies.* Retrieved from: https://www.researchgate.net/publication/24075681_Estimating_the_size_of_the_shadow_economy_in_Spain_A_structural_model_with_latent_variables

Johnson, S., D. Kaufmann, & P. Zoido-Lobatón. (1998). Regulatory discretion and the unofficial economy. *The American Economic Review*, 88(2),387-392.

Jung, S., Park, S., Jung, S. & Hwang, E. (2020). Monthly Electric Load Forecasting Using Transfer Learning for Smart Cities. *Sustainability,* 12(16), 1-18.

Kanniainen, V., Pääkkönen, J. & Schneider, F. (2004). Determinants of shadow economy: theory and evidence. *ResearchGate.* Retrieved from: https://www.researchgate.net/publication/228592221_Determinants_of_shadow_economy_theory_and_evidence

Katchova, A. (2014). Principal Component Analysis. *Econometrics Academy.* Retrieved from: https://sites.google.com/site/econometricsacademy/econometrics-models/principal-component-analysis

Keith, A. & Marill, M. D. (2004). Advanced Statistics: Linear Regression, Part II: Multiple Linear Regression. *Academic Emerging Medicine,* 11(1), 94-102.

Kelmanson, B., Kirabaeva, L., Medina, L., Mircheva, B. & Weiss, J. (2019). Explaining the shadow economy in Europe: size, causes and policy options. *International Monetary Fund.* Retrieved from: https://www.imf.org/en/Publications/WP/Issues/2019/12/13/Explaining-the-Shadow-Economy-in-Europe-Size-Causes-and-Policy-Options-48821

Kira, A. R. (2013). The Factors Affecting Gross Domestic Product (GDP) in Developing Countries: The Case of Tanzania. *European Journal of Business and Management,* 5(4), 148-158.

Kodres, L. E. (2013). What is Shadow Banking? *Finance & Development,* 50(2), 42-42.

Konno, S., Teramoto, A. & Mera, Y. (2013). Compiling statistics of shadow banking. *Bank for International Settlements,* 36, 76-88.

"Labour force participation rate". (2021). *OECD Data.* Retrieved from: https://data.oecd.org/emp/labour-force-participation-rate.htm

Laerd statistics. (n.d.). Multiple Regression Analysis using Stata. *Lund Research Ltd.* Retrieved from: https://statistics.laerd.com/stata-tutorials/multiple-regression-using-stata.php

Laerd statistics. (n.d.). Pearson's Correlation using Stata. *Lund Research Ltd.* Retrieved from: https://statistics.laerd.com/stata-tutorials/pearsons-correlation-using-stata.php

Laerd statistics. (n.d.). Pearson Product-Moment Correlation. *Lund Research Ltd.* Retrieved from: https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php

Laffer, A. B. (2004). The Laffer Curve: Past, Present and Future. *The Heritage Foundation,* No. 1765. Retrieved from: https://iife.edu.vn/wp-content/uploads/2020/04/Laffer-Couver-Last-Present-and-Future-bg1765.pdf

"Leandro Medina". (n.d.). *World Economic Journal.* Retrieved from: https://www.world-economics-journal.com/Authors/Leandro%20Medina.author?AID=1185

Medina, L. & Schneider, F. (2018). Shadow Economies Around the World: What Did We Learn Over the Last 20 Years? *IMF Working Papers.* Retrieved from: https://www.imf.org/en/Publications/WP/Issues/2018/01/25/Shadow-Economies-Around-the-World-What-Did-We-Learn-Over-the-Last-20-Years-45583

Moon, S., RØttingen, J. & Frenk, J. (2017). Global public goods for health: weaknesses and opportunities in the global health system. *Health Economics, Policy and Law,* 12(2), 195-205.

Nesvetailova, A. (2018). *Shadow Banking*. Oxon: Routledge.

"Overall score". (2021). *2021 INDEX OF ECONOMIC FREEDOM*. Retrieved from: https://www.heritage.org/index/visualize

Ray, K., Das, R. C., & Das, U. (2016). Convergence Aspect of Capital Formation: A Study on Major Countries. *Handbook of Research on Global Indicators of Economic and Political Convergence*, pp. 51-66. IGI Global. http://doi:10.4018/978-1-5225-0215-9.ch003

"Regulatory quality". (2019). *The World Bank.* Retrieved from: https://govdata360.worldbank.org/indicators/h5083f593?country=USA&indicator=394&countries=BRA&viz=line_chart&years=1996,2019

Schneider, F. (2013). Size and development of the shadow economy of 31 European and 5 other OECD countries from 2003 to 2013: a further decline. *Researchgate.* Retrieved from: https://www.researchgate.net/publication/268185661_Size_and_Development_of_the_Shadow_Economy_of_31_European_and_5_other_OECD_Countries_from_2003_to_2012_Some_New_Facts

Schneider, F. & Buehn, A. (2016). Estimating the Size of the Shadow Economy: Methods, Probblems and Open Questions. *IZA.* Discussion paper No. 9820. Retrieved from: http://ftp.iza.org/dp9820.pdf

Schneider, F. & Enste, D. (2002). Hiding in the Shadows, The Growth of the Underground Economy. *International Monetary Fund.* Retrieved from: https://www.imf.org/external/pubs/ft/issues/issues30/

Schneider, F. & Neck, R. (1993). The Development of the Shadow Economy under Changing Tax Systems and Structures. *Journal of Economic Literature*, 50(3), 344-369.

"Social security contribution". (2020). *OECD.* Retrieved from: https://data.oecd.org/tax/social-security-contributions.htm

Statista Research Department. (2021). Civilian labor force participation rate in the United States from 1990 to 2020. *Statista.* Retrieved from: https://www.statista.com/statistics/191734/us-civilian-labor-force-participation-rate-since-1990/

STAT462. (2018). 10.7 – Detecting Multicollinearity Using Variance Inflation Factors. *PennState Eberly College of Science.* Retrieved from: https://online.stat.psu.edu/stat462/node/180/

"Stopping Crime and Violence in Latin America: A Look at Prevention from Cradle to Adulthood". (2019). *The World Bank.* Retrieved from: https://www.worldbank.org/en/results/2018/05/17/fin-a-la-violencia-en-america-latina-una-mirada-a-la-prevencion-desde-la-infancia-hasta-la-edad-adulta

Studenmund, A.H. (2017). *A Practical Guide to Using Econometrics*. Harlow: Pearson Education Limited.

"Tax Morale: What Drives People and Businesses to Pay Taks?". (2019). *OECD Publishing*. Retrieved from: https://www-oecd-org.proxy.library.uu.nl/tax/tax-morale-f3d8ea10-en.htm

"Tax revenue". (2020). *OECD data.* Retrieved from: https://data.oecd.org/tax/tax-revenue.htm

"Tax revenue (% of GDP) – United States". (2021). *The World Bank.* Retrieved from: https://data.worldbank.org/indicator/GC.TAX.TOTL.GD.ZS?end=2020&locations=US&start=1991

United Nations. (2014). Country classification. *World Economic Situation and Prospects.* Retrieved from: https://www.un.org/en/development/desa/policy/wesp/wesp_current/2014wesp_country_classification.pdf

U.S. Bureau of Economic Analysis. (2021). Exports of Goods and Services. *FRED.* Retrieved from: https://fred.stlouisfed.org/series/EXPGS

U.S. Bureau of Economic Analysis. (2021). Imports of Goods and Services. *FRED.* Retrieved from: https://fred.stlouisfed.org/series/IMPGS#0

U.S. Bureau of Labor Statistics. (2021). Unemployment rate. *FRED.* Retrieved from: https://fred.stlouisfed.org/series/UNRATE#

Weedmark, D. (2018). The Advantages & Disadvantages of a Multiple Regression Model. *Sciencing.* Retrieved from: https://sciencing.com/advantages-disadvantages-multiple-regression-model-12070171.html

"What is FRED?". (n.d.). *FRED.* Retrieved from: https://fredhelp.stlouisfed.org/fred/about/about-fred/what-is-fred/

"WHO ARE WE". (n.d.). *The World Bank.* Retrieved from: https://www.worldbank.org/en/who-we-are

Williams, C. C. (2015). Out of the shadows: Classifying economies by the extent and nature of employment in the informal economy. *International Labour Review*, 154(3), 331-351.

Wiseman, T. (2013). US shadow economies: a state-level study. *Constitutional Political Economy* 24, 310-335. Retrieved from: https://link-springer-com.proxy.library.uu.nl/article/10.1007/s10602-013-9146-7#citeas

Wooldridge, J. M. (2015). *Introduction to Econometrics.* Hampshire, UK: Cengage Learning.

Worldwide Governance Indicators. (2020). Regulatory Quality: Estimate. *The World Bank.* Retrieved from: https://datacatalog.worldbank.org/regulatory-quality-estimate

Worldwide Governance Indicators. (n.d.). Regulatory Quality. *The World Bank.* Retrieved from: https://govdata360.worldbank.org/indicators/h5083f593?country=USA&indicator=394&viz=line_chart&years=1996,2019