VILNIUS UNIVERSITY
FACULTY OF MATHEMATICS AND INFORMATICS

MASTER THESIS

Modeling Corruption and Shadow Economy via SEM

Korupcijos ir Šešėlinės Ekonomikos Modeliavimas taikant Struktūrinių Lygčių Modelius

Giedrė Maksimavičiūtė

VILNIUS 2020

MATEMATIKOS IR INFORMATIKOS FAKULTETAS
TAIKOMOSIOS MATEMATIKOS INSTITUTAS

Darbo vadovas	Prof., Habil. dr. Vydas Čekanavičius
Darbo recenzentas	dr. Dmitrij Celov

Darbas apgintas 2020 sausio 13d.

Registravimo NR.

Modeling Corruption and Shadow Economy via SEM

Abstract

The main new idea of the present work is that the Corruption Index is not equivalent to corruption itself. Moreover, the Corruption Index has a slight biasedness due to certain subjectivity of its construction. Since corruption cannot be measured directly in this work it is modeled as some latent variable. The Corruption Index is used just like one of the indicators of corruption. The classical SEM, when the witness manifestation of corruption through various indicators proved insufficient for our purposes, therefore, SEM MIMIC containing various exogenous variables (possible sources for corruption) was constructed. Corruption is not equivalent to the Shadow Economy, but two are closely related. Therefore, a separate SEM model was constructed for the Shadow Economy. Finally, one joint MIMIC model for corruption and the Shadow Economy was obtained. One of the main problems related to data was variables with different frequency, but it was solved applying Denton - Chollete disaggregation.

Key words : (Corruption, Shadow Economy, SEM, MIMIC)

Korupcijos ir Šešėlinės Ekonomikos Modeliavimas taikant Struktūrinių Lygčių Modelius

Santrauka

Pagrindinė ir nauja darbo mintis yra ta, kad korupcijos indeksas nėra lygus pačiai korupcijai. Be to, korupcijos indeksas gali būti truputį šališkas dėl tam tikro subjektyvumo. Kadangi korupcijos negalima įvertinti tiesiogiai, šiame darbe ji bus modeliuojama kaip latentinis kintamasis. Korupcijos indeksas bus naudojamas kaip vienas iš rodiklių. Klasikinis SEM modelis, kai korupcija yra išreiškiama įvairiais indikatoriais, pasirodė nepakankamas mūsų tikslams, dėl to buvo sukonstruotas SEM MIMIC modelis, kuriame naudojami ir įvairūs egzogeniniai kintamieji (galimi korupcijos šaltiniai). Korupcija neprilygsta šešėlinei ekonomikai, tačios jos dvi yra glaudžiai susijusios. Todėl šešėlinei ekonomikai buvo sukurtas atskiras SEM modelis. Galiausiai buvo gautas vienas bendras korupcijos ir šešėlinės ekonomikos MIMIC modelis. Viena iš pagrindinių problemų, susijusių su duomenimis, buvo skirtingas jų dažnis, dėl to buvo taikytas duomenų deagregavimas naudojant Denton - Cholette metodą.

Raktiniai žodžiai : (Korupcija, Šešėlinė Ekonomika, SEM, MIMIC)

Contents

1 Introduction

It is well known that corruption and the Shadow Economy exist. The investigation of the Shadow Economy goes back for more than a few decades. A term corruption is much older, it is even hard to find when it appeared for the first time. You can read about both of them in newspapers, see announcements in the news or even hear some comments on a bus on Your way to work. Corruption and the Shadow Economy are widely considered as significant problems in our society and yields serious challenges for all countries, especially the ones with developing economies. Lithuania is not an exception. The problem is that they cannot be measured directly.

There are different thoughts about the development of corruption and the Shadow Economy in Lithuania. If we will be able to establish main causes and indicators of the Shadow Economy and corruption, this will help to understand better the choices of further actions and decisions for government and provide a better insight into the problem for economists.

The main aim of the research is to find those causes and indicators, which could affect one of the previously mentioned problems: corruption and the Shadow Economy in Lithuania. We begin with an overview of possible causes of corruption and known models for other countries. Then we choose an appropriate model for analysis. Finally, a relationship between the corruption model and the Shadow Economy model is investigated.

2 Theoretical framework

2.1 The Shadow Economy

The Shadow Economy, as can be expected from its name, is not directly observable and has a few different definitions. Some of them are related to taxes ("revenue not reported to, and not discovered by, the tax authorities produced in underground activities"); some are related to the economy ("national production or income that is missed by the statistical offices when they calculate the value of national product", see Tanzi (1999)). One of the most common definitions is that of Smith's: "market - based production of goods and services whether legal or illegal, that escapes detection in official estimates of GDP" or "market-based production of goods and services, whether legal or illegal, that escapes detection by the tax authorities", see Smith (1997). Schneider and Enste (2002) define the Shadow Economy slightly differently: "...includes not only illegal activities but also unreported income from the production of legal goods and services, either from monetary or barter transactions".

Though definitions are slightly different, at the core of all of them is the same idea - in general, the Shadow Economy consists of all economic activities, which are not registered in the official GDP.

Note that, in our models, we do not include any direct violent criminal activities, such as burglary, drug dealing, human trafficking, etc.

2.1.1 Causes of the Shadow Economy

Tax Burden. When the Shadow Economy is discussed, tax related concepts are most frequently considered (e.g., tax burden, tax rates, tax evasion, etc.). That is because taxes, or to be more precise, tax burden, is one of the main drivers of the unobserved economy, see Schneider and Enste (1999).

One of the first to research the Shadow Economy are Allingham and Sandmo (1972). The main idea of their work is that with the increase of the tax burden and tax evasion, the unobserved economy also increases. The tax burden is described as one of the three major causes of the Shadow Economy. For example, Scandinavia has high marginal tax rates, but a lower tax burden and this is a supposed reason for it to have a lower share of the unobserved economy. In Russia, the situation is different: low marginal tax rates and high tax burden causes a high share of the unofficial economy in GDP, see Johnson et al. (1998).

Unemployment Rate. It is natural to expect that if the unemployment rate in the country is high, the Shadow Economy is larger too. For this reason, most researchers take into account the percentage of the unemployed workforce.

Tanzi writes: "relation between the underground economy and the unemployment rate is ambiguous'", see Tanzi (1999). He observes that a group of people who are counted as 'hidden' workers can be retired people, illegal immigrants, those who have official and unofficial works, housewives who are not part of the official workforce, etc. That is, very different people. The same ambiguity is noted in "Unemployment and the Shadow Economy in the OECD", see Bajada and Schneider (2009); meaning that, you can have a 'normal' official work, at the same time being a part of the Shadow Economy.

Regulation. One of the most important causes of the Shadow Economy is regulation and its intensity. Regulations include such things as minimum working age, overtime payments, minimum number of holidays, etc. It is shown that strong evidence exists on the relationship between lower regulation and lower share of the unofficial economy, see Johnson et al. (1998).

Political Stability. Political stability is not one of the most important causes of the Shadow Economy but it still can be related to it. It is shown by Elbahnasawy and Ellis, that: "Political instability, political polarization, and authority pattern seem to play a role in determining the incentives of government to develop tax capacity of state and the development of the informal economy. The empirical analysis yields insights into the intensity of political instability and polarization required for the informal economy to expand", see Elbahnasawy et al. (2016).

2.1.2 Indicators of the Shadow Economy

Gross Domestic Product. Gross Domestic Product (GDP) is one of the best - known Shadow Economy indicators (indicator is a variable, with which one latent variable can be reflected). According to economic logic, GDP defines an economic development level of the country and usually is assumed to be an indicator of wealth. It is quite obvious that poor countries have a much bigger part of the Shadow in their economies. It is noted, that the Informal Economy increases, when GDP decreases, due to that the Informal Economy represents a "life jacket" for firms and individuals in financial troubles, see Dell'Anno et al. (2004). However, a positive relationship also exists between GDP and Shadow Economy, see Giles (1999). The hidden economy increase efficiency, expand entrepreneurship and affect growth in the official economy.

Money Supply. Mostly, transactions in the Shadow Economy are carried out using cash instead of credit cards, checks, etc. to avoid any evidence of illegal activities. "As a rule, countries, where the use of electronic money is more widespread see substantially lower volumes of the Shadow Economy", see Krstić and Schneider (2015). So the monetary approach to estimate the size of an illegal economy is based on this assumption.

Labor force. The total labor force or the labor force participation rate can also be an immensely important indicator of the Shadow Economy. Changes in them can reflect a flow of resources between the official and unofficial economies. Although the labor force participation rate is one of the main indicators of the illegal economy, its impact can be various: sometimes positive, sometimes negative. In the last years, the structural composition of the labor force has changed. For example, female participation in a workforce grows, etc., see Dell'Anno (2007).

2.2 Corruption

Corruption, just like the Shadow Economy, does not have a unique definition. For example, Huberts writes: "It is crucial to be clear about the actual interpretation of the corruption concept. Is it bribing and being bribed, is it private profit from (public) power; is it un-ethical (public) behavior?", see Huberts (2010). The scale of corruption can depend on differences in culture, history, legal systems, traditions, the mentality of the people, etc.

One of the most popular descriptions is by Transparency International: "Corruption is the abuse of entrusted power for private gain. It can be classified as grand, petty and political, depending on the amounts of money lost and the sector where it occurs", see Transparency (2010).

2.2.1 Causes of Corruption

Rule of law. The rule of law is crucial for a stable democracy and social justice. That is why the role of the legal system is featured prominently in many studies and can be marked as a cause of corruption, see North and North (1992). In countries where a higher level of corruption is expected, the lower respect for the rule of law exists. As remarked by Buehn and Schneider (2009): "While strong and efficient legal systems protect property rights and provide a stable framework for economic activity, weak legal systems fail to provide such an environment". Less opportunity for corruption is left in a country where more laws are proclaimed to protect private and public interests.

Government. A government always plays a very important role in the development of the economy of a country. It also can be viewed as one of the possible causes of corruption: a more democratic government with more transparent decision rules usually means lower corruption in a country. Štefan Šumah wrote: "...there are many opportunities to manipulate public spending and it is carried out by high-level officials to get bribes (meaning more government spending or a large budget gives more opportunities for corruption).", see Sumah et al. (2018).

2.2.2 Indicators of Corruption

Corruption Perceptions Index. Every year since 1995 Transparency International presents the Corruption Perceptions Index. It ranks 180 countries and territories by their perceived levels of public sector corruption according to experts and businesspeople by using different surveys and analysis. The assumption was held that the Corruption Perceptions Index is an indicator of the real corruption and when the latter increases, the former acts the same. Due to its construction, the Corruption index is affected by some subjectivity of the experts.

Political Stability. Corruption affects political stability. This indicator was presented as a cause earlier because it can vary as the Shadow Economy changes. However it was also found that political stability is related to corruption: "...while times of political instability are bad for economic growth, they might be taken as an opportunity to improve institutional quality and combat the spread of corruption", see Abdel-Latif et al. (2018).

2.3 The Shadow Economy versus Corruption

It might seem that corruption and the Shadow Economy go hand in hand and essentially represent the same thing. However, it is interesting to ask if a relationship between these two different concepts exists. There is no universally accepted answer. For example, Schneider (2007) wrote: "...corruption and the Shadow Economy can be either complements or substitutes". Both corruption and the Shadow Economy can be mixed up because of common circumvention of regulations, payment of taxes, revenues, increase in public expenditures and hamper on productivity and growth, see Borlea et al. (2017). On the other hand, corruption and the Shadow Economy possess many differences.

3 Methology

3.1 Structural Equation Model

This chapter is prepared by using V. Čekanavičius and G. Murauskas book "Statistika III", see Čekanavičius and Murauskas (2009) and A .Buehn and F. Schneider paper "MIMIC models, Cointegration and Error Correction: An Application to the French Shadow Economy", see Buehn and Schneider (2008).

Structural equation modeling (SEM) is a multivariate statistical analysis technique that is used to analyze structural relationships among variables. This technique is a combination of factor analysis and multiple regression analysis.

The creation of the model consists of 4 steps:

- Description of the model.
- Identification of the model.
- Estimation of parameters.
- Interpretation of results and improvement of the model.

Latent variables are not monitored directly, like imagination, tolerance, etc. By creating a SEM model it is assumed that latent variables impact can be defined by observed values. Variables that are related to (explained by) factors are called those factors' indicators. The rest are directly observed exogenous variables.

Types of variables in SEM:

- Latent factors.
- Indicators.
- Exogenous variables.

Assumptions of SEM Before creating a structural equation model, six assumptions should be fulfilled:

- Missing values - it is expected no missing values in the data.
- Outliers - data should be free of outliers.
- Sample size - most researchers prefer a 200 to 400 sample size with 10 to 15 indicators. As a rule of thumb, 10 to 20 times as many cases as variables are expected.
- Multicollinearity of observed variables - correlations between variables should not be very close to zero.
- Linearity - the dependencies of all variables are linear.
- Normality - observed variables must have a normal distribution. In practice, this requirement is frequently replaced by a requirement for variables to be measured on an interval scale.

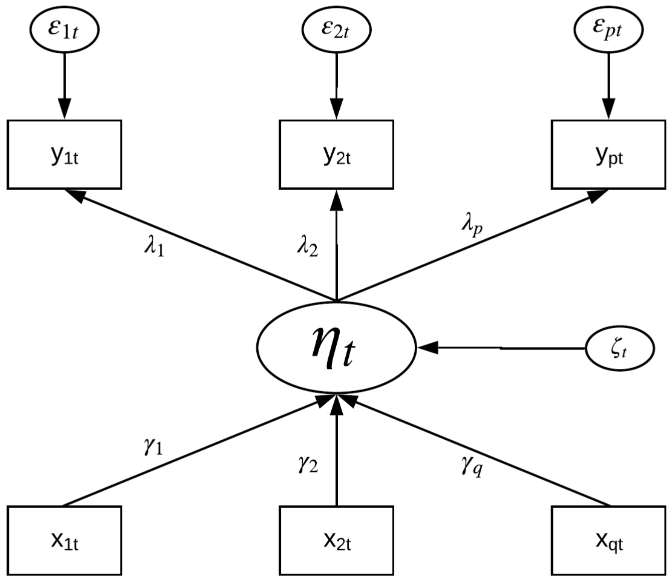

Figure 1: General Structure of a MIMIC model

3.2 Multiple Indicators and Multiple Causes Model

This research is focused on Multiple Indicators and Multiple Causes (MIMIC) models, which are a part of SEM models, where a latent variable is caused by some variables and reflected by others. For the first time these models were presented almost 50 years ago, see Zellner (1970). They became quite popular and are widely used nowadays because several causes and indicators can be applied at the same time.

The main idea of the MIMIC model is to explain the relationship between observable variables and an unobservable variable by minimizing the distance between the sample covariance matrix and the covariance matrix predicted by the model. This model consists of two types of equations: structural and measurement. The structural equation model defines the relationships between the indicators and the latent variable. It is given by

$$\eta_t = \gamma^{'}\mathbf{x_t} + \zeta_t, \tag{1}$$

where $\mathbf{x_t}^{'} = (x_{1t}, x_{2t}, ..., x_{qt})$ is a $(1 \times q)$ vector of time series as indicated by the subscript t. Each time series $x_{it}, i = 1, ..., q$ is a potential cause of latent variable η_t and a 'causal' relationships between the latent and its causes is described by a vector of coefficients in the structural model $\gamma^{'} = (\gamma_1, \gamma_2, ..., \gamma_q)$. Since the structural equation model only partially explains the latent varibale η_t, the error term ζ_t represents the unexplained component. The MIMIC model assumes that the variables are measured as deviations from their means and that the error term does not correlate to the causes, that is $\mathbf{E}(\eta_\mathbf{t}) = \mathbf{E}(\mathbf{x_t}) = \mathbf{E}(\zeta_\mathbf{t}) = 0$ and $\mathbf{E}(\mathbf{x_t}\zeta_\mathbf{t}^{'}) = \mathbf{E}(\zeta_\mathbf{t}\mathbf{x_t}^{'}) = \mathbf{0}$. The variance of ζ_t is the $(q \times q)$ covariance matrix of the causes x_t.

The measurement model represents the link between the latent variable and its indicators, which means, that the latent unobservable variable is expressed in terms of observable endogenous variables. It is specified by a linear relationship

$$\mathbf{y_t} = \lambda\eta_t + \varepsilon_t, \tag{2}$$

where $\mathbf{y_t'} = (y_{1t}, y_{2t}, ..., y_{pt})$ is a $(1 \times p)$ vector of individual time series variables $y_{jt}, j = 1, ..., p$. $\varepsilon_\mathbf{t} = (\varepsilon_{1t}, \varepsilon_{2t}, ..., \varepsilon_{pt}$ is a $(p \times 1)$ vector of disturbancies, where every $\varepsilon_{jt}, j = 1, ..., p$ is a white noise error term. Their $(p \times p)$ covariance matrix is given by $\mathbf{\Theta}_\varepsilon$. The single $\lambda_j, j = 1, ..., p$ in the $(p \times 1)$ vector of regression coefficients λ, represents the magnitude of the expected change of the respective indicator for a unit change in the latent variable. Like the MIMIC model's causes, the indicators are directly measurable and expressed as deviations from their means, that is, $\mathbf{E}(\mathbf{Y_t}) = \mathbf{E}(\varepsilon_\mathbf{t}) = 0$. Moreover, it is assumed that the error terms in the measurement model do not correlate either to the causes $\mathbf{x_t}$ or to the latent variable η_t, hence, $\mathbf{E}(\mathbf{x_t}\varepsilon_\mathbf{t}') = \mathbf{E}(\varepsilon_\mathbf{t}\mathbf{x_t}') = 0$ and $\mathbf{E}(\eta_\mathbf{t}\varepsilon_\mathbf{t}') = \mathbf{E}(\varepsilon_\mathbf{t}\eta_\mathbf{t}') = 0$. And the final assumption is that the ε_t does not correlate to ζ_t, that is $\mathbf{E}(\varepsilon_\mathbf{t}\zeta_\mathbf{t}') = \mathbf{E}(\zeta_\mathbf{t}\varepsilon_\mathbf{t}') = 0$. Figure 1 shows the general structure of the MIMIC model.

MIMIC model's covariance matrix describes the relationship between the observed variable in terms of their covariances. From equations (1) and (2) model's covariance matrix can be derived as

$$\Sigma = \begin{pmatrix} \lambda(\gamma'\Phi\gamma + \psi) + \Theta_\varepsilon & \lambda\gamma'\Phi \\ \Phi\gamma\lambda' & \Phi \end{pmatrix}, \tag{3}$$

where Σ is a function of the parameters λ, γ and the covariances contained in Φ, Θ_ε, and ψ.

Since the latent variable is not observable, its size is unknown, and the parameters of the model must be estimated using the links between the observed variables' variances and covariances. Thus, the goal of the estimation procedure is to find values for the parameters and covariances that procedure an estimate for Σ that is as close as possible to the sample covariance matrix for the observed causes and indicators, thas is the $x_t s$ and $y_t s$.

Although the estimates obtained by the MIMIC model are the most reliable compared to other methods, drawbacks still exist. The main difficulties in applying this model can be identified:

- Application with small samples.
- Complex calculation of latent variable's confidence intervals.
- Hypothesis testing about structural and measurement errors independence.
- Complex to convert index from SEM to real values that are analyzed.

Furthermore, it is noticed that most of the macroeconomic variables do not satisfy stationarity condition and by using their differences long term information is lost. Therefore, stationarity should be given special attention as it is a crucial condition for good regression.

4 Empirical Application

4.1 Data Description

The data used in this research comes from various sources and covers only one country - Lithuania. The time period from 1999 to 2018 is taken. The particular time period is chosen because of the importance of the Corruption Perceptions Index variable and its data availability. A list of variables used in this analysis is provided in Appendix A.

As mentioned before, the greatest attention was paid to the Corruption Perceptions Index. It is an index, which is created by Transparency International, using a survey method and should reveal corruption. It is assumed that corruption can be related to this index, but cannot be perceived as a 'real corruption' due to its subjectivity. Normally, the scale of an index is from 0 (highly corrupt) to 100 (very clean), but for an easier interpretation this index was rescaled: 0 means no corruption and 100 - highly corrupt.

In the analysis, 3 variables from Worldwide Governance Indicators about the government were also used: government effectiveness, political stability and the rule of law. They also were rescaled from 0 (weak) to 5 (strong) governance performance (earlier it was from -2.5 (weak) to 2.5 (strong)).

The tax burden was calculated as the tax revenue percentage of GDP. An intensity of economic regulations was calculated as a percentage of GDP related to government employment.

Furthermore, variables, such as GDP, labor force and its participation rate, openness to trade, M1 and unemployment rate were taken from different open sources: Eurostat, Official Statistics Portal of Lithuania, etc.

4.2 Data Disaggregation

In practice, differences in data frequency are experienced when data is collected from various sources. Therefore, before modeling, it is necessary to make sure that the data included in the model equation have the same frequency. According to their frequency, the data can be divided into two categories:

- Low frequency. This is rarely observed data - annual or less frequent. Observing such sizes there are difficulties due to data collection, aggregation, and interpreting and this also takes a long time.

- High frequency. These are the more commonly observed sizes - quarterly, monthly, weekly, daily. In most cases, monitoring of these sizes is automated and does not require additional resources for processing or interpreting them.

To equalize the frequencies of the available data, it is necessary to decide whether the change from the high - frequency data to the low (aggregate) or vice versa (disaggregate) is needed.

One of the SEM models' assumption is a sufficient amount of data. Variables such as rule of law, political stability, openness to trade, government effectiveness and Corruption Perceptions Index had only annual data and 20 observations might be not enough for a plausible model, therefore, disaggregation was used. There are many different methods of disaggregation, see Sax and Steiner (2013). Denton - Chollete disaggregation was applied since it gave better results in comparison to the ones obtained by Chow - Lin method. Moreover, this type of disaggregation is usually suggested.

Labor force and labor force participation rate had quarterly data only from 2001, so disaggregation was also used for the first 2 years (1999 and 2000).

4.3 Correlations, Normality and Stationarity Testing

The empirical analysis started by pre-testing the data. In the first step, each series' stationarity was checked. For the annual data, all variables were I(1) except M1, tax burden and labor force participation rate (see table 6 in Appendix B). For the quarterly data situation was almost the same, only M1 was not I(1) (see table 7 in Appendix B). Stationarity is not a must assumption for the SEM model, that is why, for neatness, first differences were used for all the variables, even though, M1, tax burden and labor force participation rate were not stationary after differencing.

Secondly, correlations between variables must be checked, due to the assumption that multicollinearity is not allowed, meaning that correlations between variables should not be very strong. Correlations were calculated for the first differences of our variables; see Appendix C. All variables seemed to fit for the MIMIC model. This assumption was fulfilled.

Finally, one of the main assumptions of the SEM models is data normality. As can be seen in the 1st table, Shapiro - Wilk test was used, also kurtosis and skewness were analyzed. First differences of the data were used because such transformed variables were analyzed in the model. For quarterly data, none of the observed variables were normally distributed (though political stability and government effectiveness are very close to the acceptancy of the normality hypothesis). Situation for annual data is better, more variables can be viewed as normally distributed, however, there are only 20 observations in the sample and special

	Annual Data			Quarterly Data		
Variable	Shapiro - Wilk Test	Kurtosis	Skewness	Shapiro - Wilk Test	Kurtosis	Skewness
Corruption Perceptions Index	0.0299	0.1312	0.9478	0.0002	1.2315	−0.9436
Economic Regulations	0.1191	−0.7909	0.1335	0.0005	1.3842	0.7109
GDP	6.65e-07	8.8604	2.7763	2.2e-16	65.7315	−8.0075
Governement Effectiveness	0.0458	−1.5784	0.1789	0.0485	−1.0430	0.1298
Labor Force	0.1742	−1.2596	−0.3884	2.792e-05	5.1130	0.8276
Labor Force Participation Rate	0.0487	−1.2759	−0.3792	7.282e-05	5.3116	1.1097
M1	0.7142	−0.3742	−0.4246	0.0001	2.1719	0.0962
Openness to Trade	0.8695	−0.2748	−0.1995	0.0042	2.3501	−0.2742
Political Stability	0.7302	−0.5600	−0.3104	0.0495	0.3994	−0.4654
Rule of Law	0.0284	0.9894	−1.0623	0.0031	1.6111	−0.8769
Tax Burden	0.0033	2.1720	−1.5546	0.0020	0.0517	−0.7712
Unemployment Rate	0.0004	0.7459	1.4073	6.195e-05	1.6527	0.8771

Table 1: Normality Testing for the First Differences

caution was taken evaluating those results' trustworthiness. For kurtosis and skewness, the results are similar to ones obtained from the Shapiro - Wilk test.

Transformations, such as logarithms, square roots, Box - Cox, Tukkey's Ladder were tested, however, even all these transformations did not succeed in making all variables normally distributed. One more way to avoid the normality assumption is to use bootstrap. Regrettably, the data sample used in this research is too small for bootstrap. Other methods, which can be used to avoid violation of normality assumption are Satorra - Bentler and Yuan - Bentler transformations. These methods were used in further research.

4.4 Expected Results

The analysis is based on similar researches by other authors. According to theoretical considerations, unemployment rate, openness to trade, rule of law, the intensity of economic regulation, M1, labor force and its participation rate, government effectiveness, GDP, Corruption Perceptions Index, political stability and the tax burden were used. It was already mentioned that the tax burden is one of the major causes of the Shadow Economy. It affects a choice between labor and leisure, increases a labor force supply in the unobserved economy, therefore, a positive sign is expected. An increase in the intensity of economic regulations reduces the free choice for individuals, that is why the Shadow Economy should be affected positively. The unemployment rate's influence can be either positive or negative, depending on the income and possibility of unemployed workers to turn to the unofficial economy. In the model a negative sign is expected for the rule of law (both for the Shadow Economy and corruption), the bigger respect for the rule of law and better institutional quality should reduce corruption. Openness to trade is an important part of economic development, a negative relationship is anticipated. If individuals feel that their preferences and interests are properly presented in the political institutions, public voice and accountability are high, this will result in a lower Shadow Economy and negative sign between them in the model. The same expectation is from government effectiveness in the corruption model.

Usually, cash is used in the Shadow Economy. Therefore, theoretically, a positive relationship with the M1 indicator is expected. The faster the official economy grows, the faster the Shadow Economy does the same, because better conditions appear for both economies, for a short term negative and for a long term positive signs are expected for GDP. Talking about the labor force, the situation is similar to the unemployment rate, because depending on the situation it can increase or decrease both the Shadow Economy and corruption. The relationship between political stability and corruption should be negative since bigger political stability means fewer wars, more clarity and less corruption. And the last indicator of corruption is the Corruption Perceptions Index which is published annually and includes surveys' results. The expected dependency is, of course, positive.

All the expectations for the specific models can be found in Appendix D.

4.5 Final Models and Results

Logarithms of M1, GDP, Corruption Perceptions Index and openness to trade were used. Yuan - Bentler transformation suggested better results, so it was decided to use it instead of Satorra - Bentler.

4.5.1 The Shadow Economy Models

Two models for the Shadow Economy were made (see figures 2 and 3): one with annual data, another with quarterly. The tax burden, as mentioned, is one of the main causes of the Shadow Economy. In Lithuania's model it is a very significant variable for quarterly data, but not for the annual one. The same situation is with the intensity of economic regulations. Most countries show a great impact of this variable to the unofficial economy but in Lithuania, only the model with annual data contains this variable. As it is seen from figures 2 and 3, the unemployment rate is involved in both models and has negative signs, which means that an increase in the unemployment rate has a negative influence on the Shadow Economy, which can happen since lower income means lower supply. Openness to

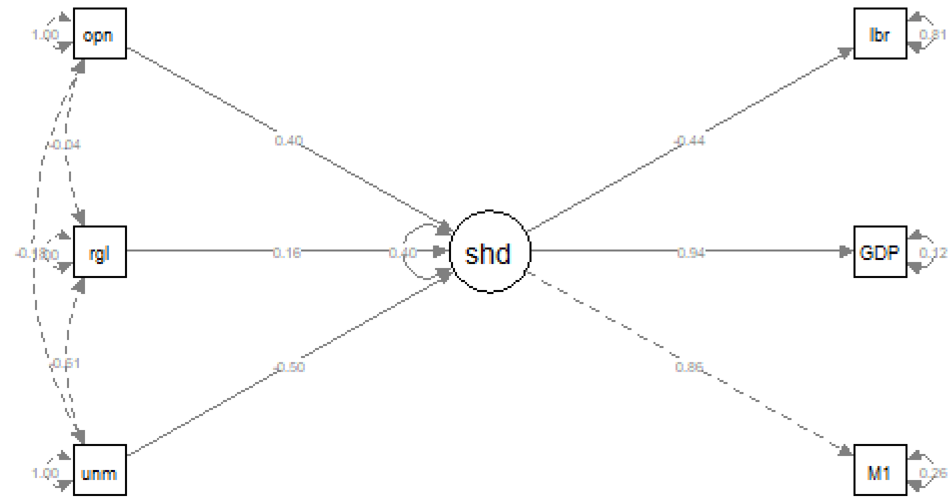

Figure 2: Model for the Shadow Economy, Annual Data

	RMSE	SRMR	χ^2	CFI
Annual model	0.00	0.075	0.582	1.00
Quarterly model	0.00	0.010	0.611	1.00

Table 2: The Shadow Economy Models' Significance Testing

trade is also included as a cause of the informal economy, although in other countries the sum of export and import does not seem to be related to the Shadow Economy. Two variables were chosen as indicators: M1 and GDP. Considering GDP, for a short period a negative impact was expected and for a long - positive. However, in both cases, a relationship is that if the Shadow Economy grows, the same does GDP. In both models, a positive relationship between M1 and the informal economy was established: when the Shadow Economy grows, cash transactions also increases. The labor force participation rate was also used as an indicator and had a negative sign for the Shadow Economy.

Considering which model is better, it is not easy to choose. Of course, quarterly one has a bigger sample size which is a merit for a MIMIC model, but looking at the statistics (see table 2) it is seen that both models fit the data. Root mean square errors (RMSE) are equal and show perfect results, but this significance test is sensitive to sample size and special attention should be paid to others. Confirmatory fit index (CFI) is perfect for both models, because $CFI \geq 0.90$, means that the model fits data. Standardized root means square residual (SRMR) is better for a model with quarterly data because lower it explains data more precisely. And the last criteria is χ^2, in the 2nd table p-value of it is demonstrated, which means that both models fit the data, but one with quarterly data - better.

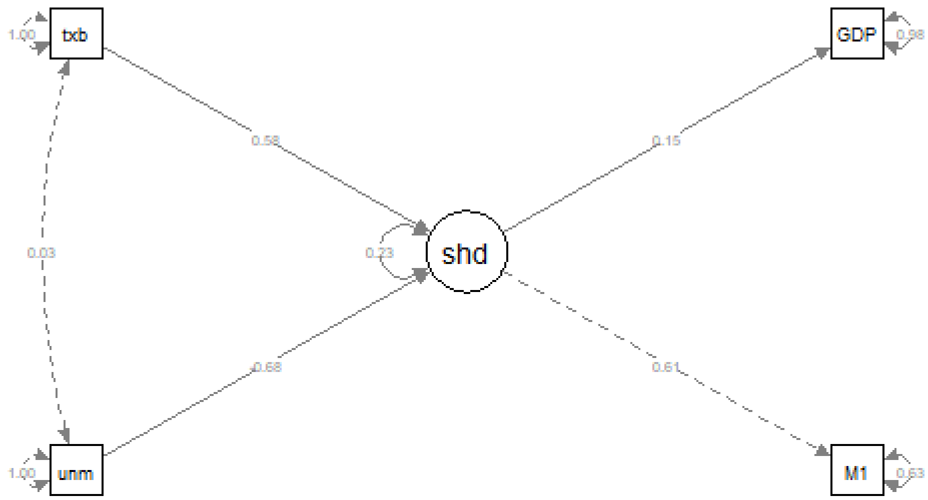

Figure 3: Model for the Shadow Economy, Quarterly Data

4.5.2 Corruption Models

Similarly to the Shadow Economy ones two models for corruption were built (see them in Appendix E). Both annual and quarterly models have the same causes and indicators and their coefficients are almost the same. As corruption causes, the rule of law and government effectiveness were analyzed. As expected, in both cases, the rule of law had a negative impact, which means that the bigger respect for the rule of law reduces corruption. It was expected that the increase in government effectiveness would decrease corruption. However, as it follows from the results in Lithuania the government effectiveness also increases corruption. This is one of the most surprising findings in this research. Corruption Perceptions Index was chosen as an indicator of corruption. Usually, people understand this index as real corruption, but it is not purely correct as was mentioned above. Therefore, this observed variable was included as an indicator. As was expected, a positive increase in the 'real corruption' means a positive growth in Transparency International's counted index. Considering political stability, it can be seen that when the political system is more stable in Lithuania, corruption gets bigger, contrarily to the expected behavior. It seems that in Lithuania, unlike other countries, corruption decrease is not related to political stability growth.

Testing which model is more significant, the same situation as with the Shadow Economy was observed. Results from significance testing can be found in the 3rd table. RMSE and CFI both show a very good fit for the data. As before, the quarterly model gives better results for SRMR and χ^2. Thus, the model containing quarterly data seems to be more accurate and fitting the data better, but the one with annual data is also acceptable.

	RMSE	SRMR	χ^2	CFI
Annual model	0.00	0.044	0.504	1.00
Quarterly model	0.00	0.004	0.650	1.00

Table 3: Corruption Models' Significance Testing

	RMSE	SRMR	χ^2	CFI
Annual model	0.145	0.113	0.167	0.917
Quarterly model	0.123	0.110	0.008	0.712

Table 4: Joint Models' Significance Testing

4.5.3 Joint Models

Two joint models with two latent variables were created (see Appendix E). Causes and indicators are almost the same as for the models with separate latent variables. Indicators for corruption remain the same in both models and retain signs of the paths. Considering causes, it is observed that both joint models contain the rule of law and government expenditure. Openness to trade was also added to the joint model with annual data. It shows the following relationship: when import and export in Lithuania grows, corruption is growing as well. For the Shadow Economy part of the model, M1 is chosen as an indicator in both models, but the labor force is included only in the model with annual data. The same is true for the model with quarterly one with an exception that GDP is used instead of the labor force. As a cause of the informal economy, openness to trade shows a positive impact on both models. The tax burden is a cause in a joint model with quarterly data as well as in the one with only one latent variable and is replaced by the intensity of economic regulation in the annual model.

A consideration which model is better according to significance tests gives no final answer (see table 4). None of the models fit data as good as expected. Evaluation of χ^2 statistic implies that the joint model with quarterly data does not fit the data. For the one with annual data, the situation is hardly better because neither RMSE nor SRMR is acceptable. There are two explanations for that: sample sizes are too small and correct results cannot be obtained (however, models for corruption and the Shadow Economy separately did not meet this problem). Another explanation might be that corruption and the Shadow Economy in Lithuania are not closely related to each other. This point of view is supported by the fact that a correlation between these two latent factors is not significant.

5 Conclusions

The main new idea of the present work is that the Corruption Perceptions Index is not equivalent to corruption itself. It is modeled as a latent variable of the SEM MIMIC model. By using available data we also explore possible causes and sources of corruption in Lithuania. Model weights allow us to estimate the impact of each of them. Since corruption is closely related to the Shadow Economy, two SEM MIMIC models - one for the Shadow Economy itself and one joint model for corruption and Shadow Economy - are also built. Preparing data for models we encountered the problem of different data frequency. This problem was solved by applying Denton - Chollete disaggregation. Note that SEM models for the Shadow Economy are well - known for many countries. On the other hand, it seems that the SEM model for corruption is rarely considered.

Comparing our model for the Shadow Economy with known models for other countries, we found that surprisingly in Lithuania the intensity of economic regulation and tax burden are not among its most significant causes. In Lithuania the most significant source of the Shadow Economy is unemployment. Though openness to trade is not usually included in the models for the Shadow Economy, it proved to be a significant reason for Lithuania Model.

Unlike our expectations, it turned out that models for the Shadow Economy and corruption are not closely related, the fact still requires economic interpretation.

References

Abdel-Latif, H., Elgohari, H., and Mohamed, A. (2018). Corruption, political instability and growth: Evidence from the arab spring. *Available at SSRN 3240211.*

Allingham, M. G. and Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of public economics*, 1(3-4):323–338.

Bajada, C. and Schneider, F. (2009). Unemployment and the shadow economy in the oecd. *Revue économique*, 60(5):1033–1067.

Borlea, S. N., Achim, M. V., and Miron, M. G. A. (2017). Corruption, shadow economy and economic growth: An empirical survey across the european union countries. *Studia Universitatis „Vasile Goldis" Arad–Economics Series*, 27(2):19–32.

Buehn, A. and Schneider, F. (2008). Mimic models, cointegration and error correction: An application to the french shadow economy.

Buehn, A. and Schneider, F. (2009). Corruption and the shadow economy: a structural equation model approach.

Čekanavičius, V. and Murauskas, G. (2009). *Statistika ir jos taikymai III dalis.*

Dell'Anno, R. (2007). The shadow economy in portugal: An analysis with the mimic approach. *Journal of Applied Economics*, 10(2):253–277.

Dell'Anno, R. et al. (2004). Estimating the shadow economy in italy: A structural equation approach. Technical report.

Elbahnasawy, N. G., Ellis, M. A., and Adom, A. D. (2016). Political instability and the informal economy. *World Development*, 85:31–42.

Giles, D. E. (1999). Measuring the hidden economy: Implications for econometric modelling. *The Economic Journal*, 109(456):370–380.

Huberts, L. W. (2010). A multi approach in corruption research: towards a more comprehensive multi-level framework to study corruption and its causes.

Johnson, S., Kaufmann, D., and Zoido-Lobaton, P. (1998). Regulatory discretion and the unofficial economy. *The American economic review*, 88(2):387–392.

Krstić, G. and Schneider, F. (2015). Formalizing the shadow economy in serbia. *Belgrade: Springer Open.*

Kwiatkowski, D., Phillips, P. C., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, 54(1-3):159–178.

MacKinnon, J. G. (1996). Numerical distribution functions for unit root and cointegration tests. *Journal of applied econometrics*, 11(6):601–618.

North, D. C. and North, D. C. (1992). *Transaction costs, institutions, and economic performance.* ICS Press San Francisco, CA.

Sax, C. and Steiner, P. (2013). Temporal disaggregation of time series. *The R Journal*, 5(2):80–87.

Schneider, F. and Enste, D. (1999). Shadow economies around the world: size, causes and consequences.

Schneider, F. and Enste, D. (2002). *Hiding in the shadows: the growth of the underground economy*, volume 30. International Monetary Fund.

Schneider, F. G. (2007). Shadow economies and corruption all over theworld: New estimates for 145 countries. *Access mode: http://ideas. repec. org/a/zbw/ifweej/5744. html*.

Smith, P. M. (1997). Assessing the size of the underground economy: The statistics canada perspective. *The underground economy: Global evidence of its size and impact*, pages 11–37.

Sumah, S. et al. (2018). Corruption, causes and consequences. *Chapters*.

Tanzi, V. (1999). Uses and abuses of estimates of the underground economy. *The Economic Journal*, 109(456):338–347.

Transparency, I. (2010). Index, corruption perceptions. *URL: https://www.transparency.org/what-is-corruption*.

Zellner, A. (1970). Estimation of regression relationships containing unobservable independent variables. *International Economic Review*, pages 441–454.

A List of Variables

Variable	Definition	Measurement	Source
Corruption Perceptions Index	Index which aggregates data from several different sources that provide perceptions by business people and country experts of the level of corruption in the public sector.	Scale from 0 to 100, where 0 means very clean and 100 - highly corrupt	Transparency International
Employment in General Government	All employment of the general government.	Thousands of people	ILOSTAT database
Gross Domestic Product	Monetary measure of the market value of all the final goods and services produced in a specific period time	EUR million	Eurostat Database
Government Effectiveness	Captures perceptions of the quality of public and civil services and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.	Scale from 0 to 5, where higher values correspond to better outcome	Worldwide Governance Indicators
Labor Force	Number employed people plus the unemployed who are looking for work	Thousands of people	Official Statistics Portal of Lithuania
M1	The money supply that is composed of physical currency and coin, demand deposits, travelers' checks, other checkable deposits, and negotiable order of withdrawal accounts.	EUR Million	Trading Economics Database
Openness to trade	Sum of imports and exports.	Percentage of GDP	Worldwide Governance Indicators
Political Stability and Absence of Violence/ Terrorism	Captures perceptions of the likelihood of political instability and/or politically motivated violence, including terrorism.	Scale from 0 to 5, where higher values correspond to better outcome	Worldwide Governance Indicators

Rule of Law	Captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence.	Scale from 0 to 5, where higher values correspond to better outcome	Worldwide Governance Indicators
Tax Revenue	Tax revenue is defined as the revenues collected from taxes on income and profits, social security contributions, taxes levied on goods and services, payroll taxes, taxes on the ownership and transfer of property, and other taxes.	EUR million	CEIC Database
Unemployment Rate	Unemployed individuals share in the total labor force.	Percentage of total labor force	Eurostat Database

B Stationarity Testing

Variable	Test Equation	Level			Test Equation	First Differences		
Causes		ADF	PP	KPSS		ADF	PP	KPSS
Corruption Perceptions Index	C & T	0.4295	0.4427	0.094	C	0.0053	0.0290	0.0943
Economic Regulation	C & T	0.4979	0.3692	0.0914	C	0.0009	0.2191	0.0741
Governement Effectiveness	C & T	0.1998	0.6309	0.0811	C	0.0218	0.0405	0.105
GDP	C & T	0.5126	0.1254	0.1573	C	0.0826	0.0540	0.361
Labor Force	C & T	0.2781	0.6884	0.1719	C	0.0372	0.0947	0.273
Labor Force Participation Rate	C & T	0.7512	0.9672	0.2471	C & T	0.4006	0.01685	0.1006
M1	C & T	0.99	0.99	0.17	C & T	0.3061	0.4742	0.1068
Openness to Trade	C & T	0.4478	0.413	0.0752	C	0.0053	0.0290	0.0943
Political Stability	C	0.9952	0.583	0.1645	C	0.0018	0.0559	0.1791
Rule of Law	C & T	0.3443	0.129	0.0981	C	0.0700	0.0371	0.0946
Tax Burden	C & T	0.2271	0.5692	0.0804	C	0.4444	0.1551	0.1587
Unemployment Rate	C	0.3367	0.6057	0.0916	C	0.0292	0.2614	0.1156

Table 6: Stationarity Testing for Annual Data [1]

Variable	Test Equation	Level			Test Equation	First Differences		
Causes		ADF	PP	KPSS		ADF	PP	KPSS
Corruption Perceptions Index	C & T	0.3675	0.5419	0.167	C	0.00	0.01	0.0963
Economic Regulation	C & T	0.5581	0.7101	0.3663	C	0.0331	0.01	0.2622

[1] For the Augmented Dickey-Fuller (ADF) test and Phillips-Perron (PP) test, the MacKinnon one-sided p values are given, see MacKinnon (1996), whereas test statistics are reported for the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. Its critical values are taken from Kwiatkowski et al. work, see Kwiatkowski et al. (1992). For a test equation with constant (C) the critical values are: 0.347 (10% level), 0.463 (5% level) and 0.739 (1% level) whereas for a test equation with constant and trend (C & T) the critical values are: 0.119 (10% level), 0.146 (5% level) and 0.216 (1% level)

Government Effectiveness	C & T	0.0997	0.6479	0.1541	C	0.00	0.6479	0.0752
GDP	C & T	0.4404	0.4849	0.1278	C	0.03443	0.01	0.0609
Labor Force	C & T	0.5689	0.3446	0.3594	C	0.0138	0.0278	0.2156
Labor Force Participation Rate	C & T	0.5738	0.9091	0.6243	C	0.01	0.01	0.0622
M1	C & T	0.9801	0.99	0.4055	C & T	0.1192	0.01	1.011
Openness to Trade	C & T	0.5815	0.5068	0.1293	C	0.00	0.0111	0.0782
Political Stability	C	0.03407	0.5137	0.2689	C	0.00	0.0244	0.1216
Rule of Law	C & T	0.7835	0.3972	0.1527	C	0.00	0.0586	0.1005
Tax Burden	C & T	0.7204	0.01	0.1672	C	0.01697	0.01	0.1626
Unemployment Rate	C & T	0.3724	0.8181	0.1807	C	0.00	0.01	0.11

Table 7: Stationarity Testing for Quarterly Data [1]

C Correlations Matrix

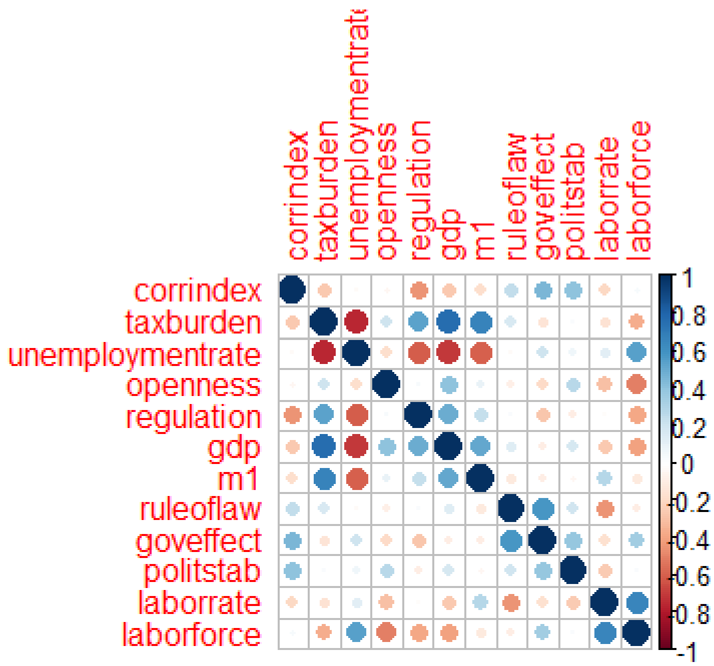

Figure 4: Correlations Matrix for Annual Data

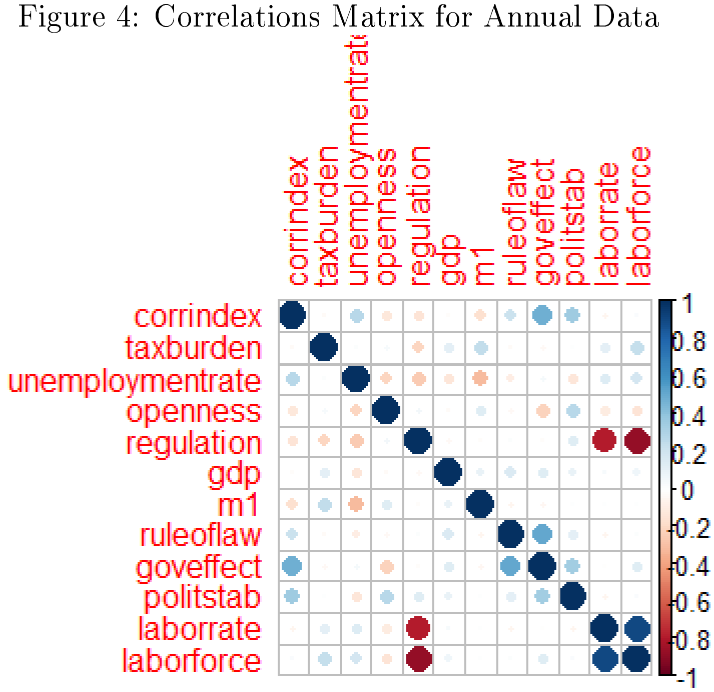

Figure 5: Correlations Matrix for Quarterly Data

D Expected Signs

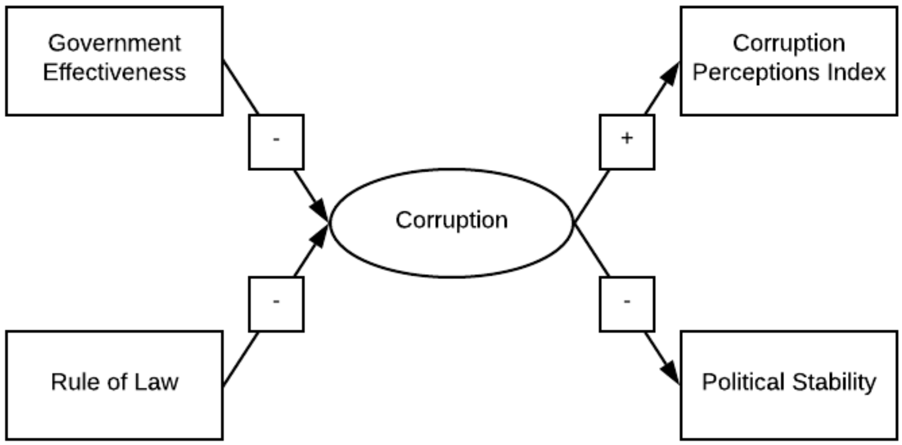

Figure 6: Expectations for Corruption

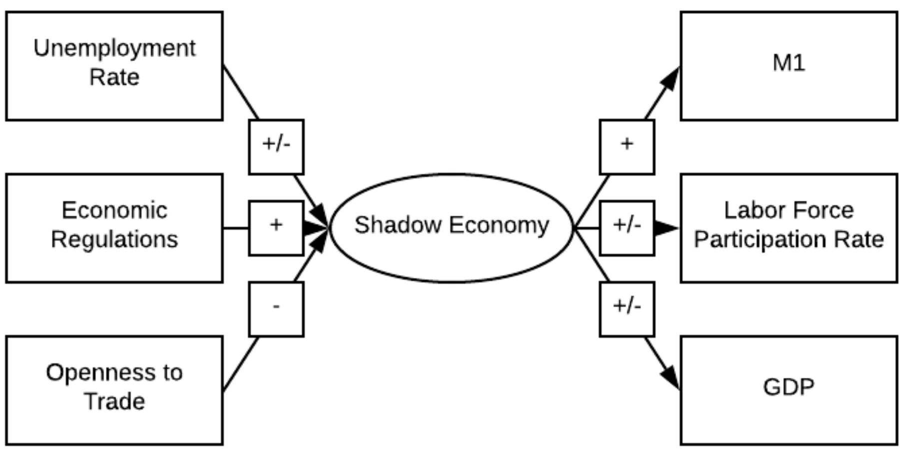

Figure 7: Expectations for the Shadow Economy. Annual Data

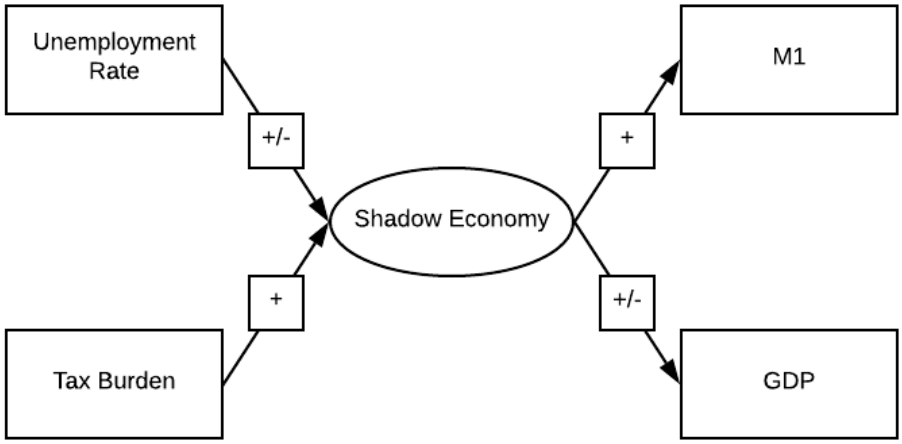

Figure 8: Expectations for the Shadow Economy. Quarterly Data

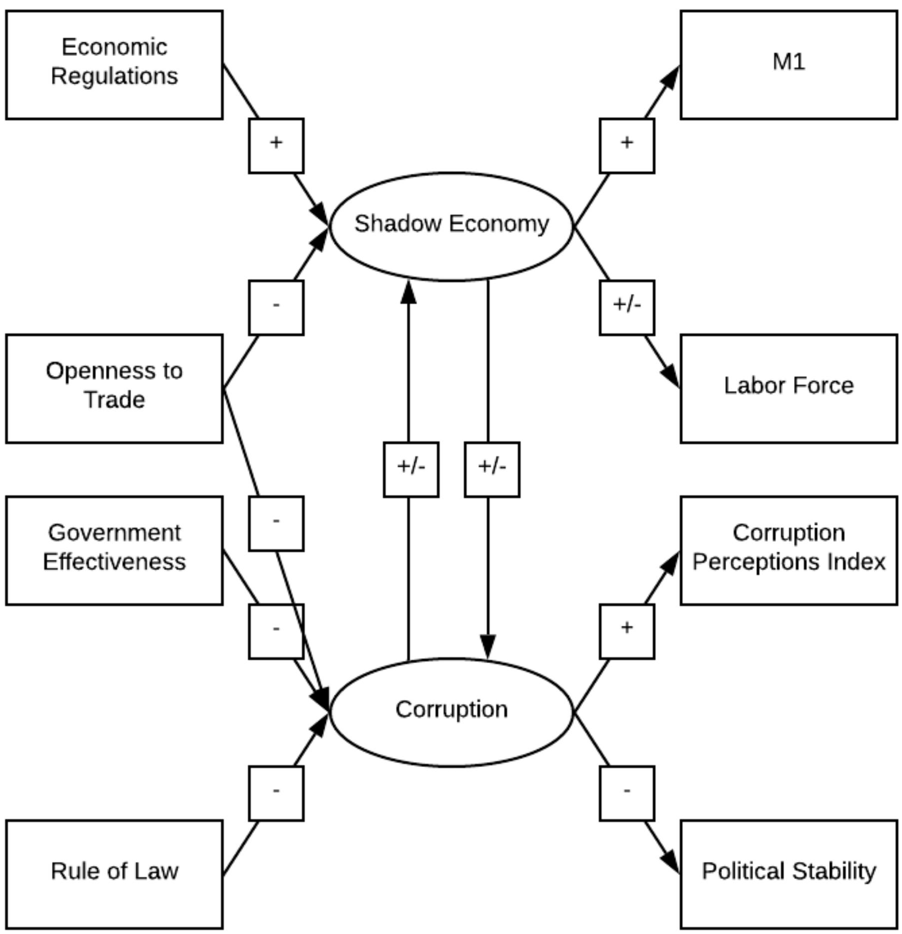

Figure 9: Expectations for Joint Model. Annual Data

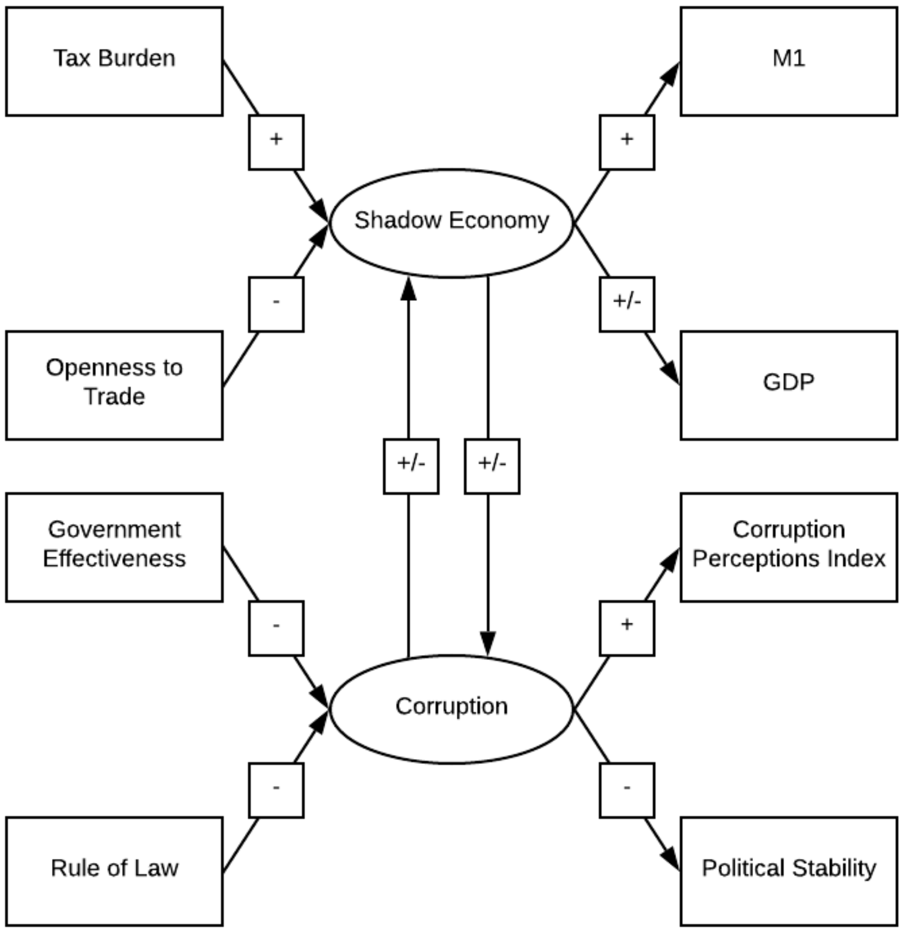

Figure 10: Expectations for Joint Model. Quarterly Data

E Plots of the Models

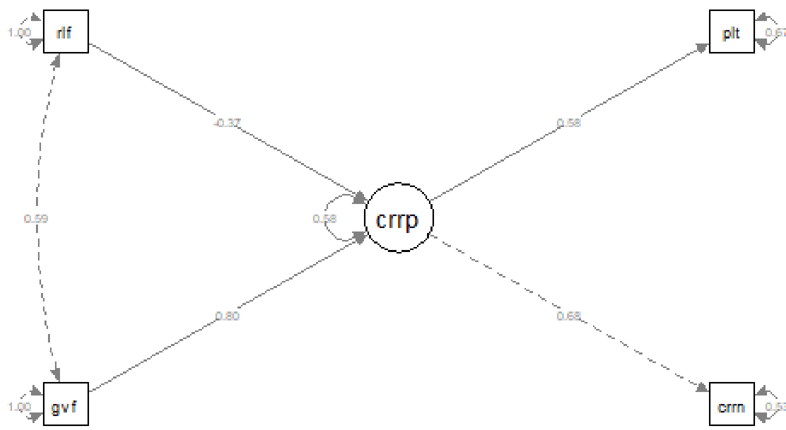

Figure 11: Model for Corruption, Annual Data

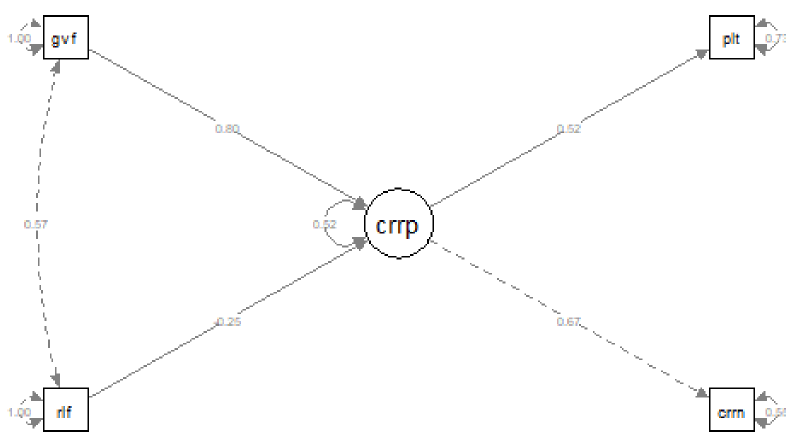

Figure 12: Model for Corruption, Quarterly Data

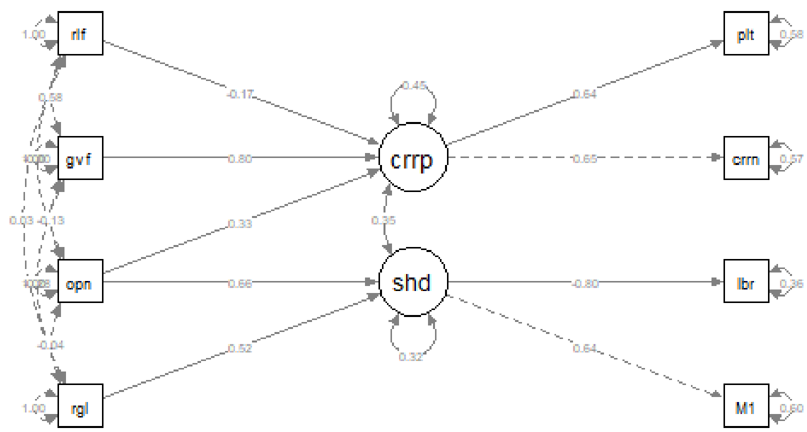

Figure 13: Joint Model, Annual data

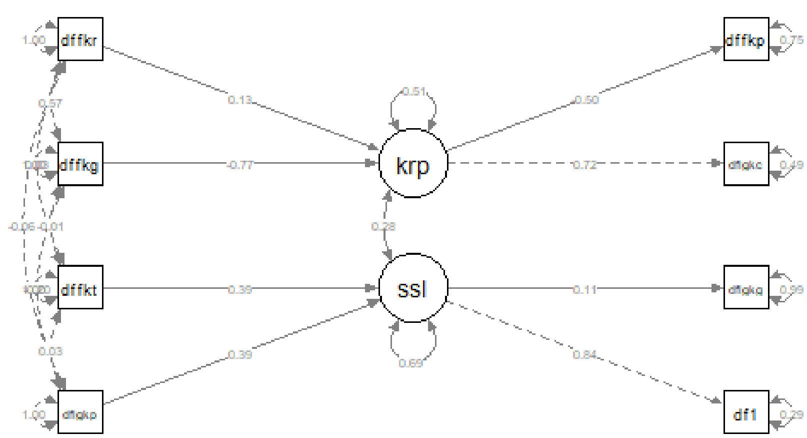

Figure 14: Joint Model, Quarterly data

F Code in R

```
# Data dissagregation
library(tempdisagg)
library(stats)
# variables such as corruptionindex, goveffect, politstab, ruleoflaw, openness were uploaded
corruptiontime<- ts(corruptionindex, start = 1999, end = 2018)
disagrcorrindex<- td(corruptiontime 1, to = "quarterly", method = "denton-cholette",conversion
= "average")
dcorrindex<- predict(disagrcorrindex)
goveffecttime<- ts(goveffect, start = 1999, end = 2018)
disagrgoveffect<- td(goveffecttime 1, to= "quarterly", method= "denton-cholette", conver-
sion = "average")
dgoveffect<- predict(disagrgoveffecttime)
politstabtime<- ts(politstab, start = 1999, end = 2018)
disagrpolitstab<- td(politstabtime 1, to= "quarterly", method= "denton-cholette", conver-
sion = "average")
dpolitstab<- predict(disagrpolitstab)
ruleoflawtime<- ts(ruleoflaw, start = 1999, end = 2018)
disagrruleoflaw<- td(ruleoflawtime 1, to= "quarterly", method= "denton-cholette", conver-
sion = "average")
druleoflaw<- predict(disagrruleoflaw)
opennesstime<- ts(openness, start = 1999, end = 2018)
disagropenness<- td(opennesstime 1, to= "quarterly", method= "denton-cholette", conver-
sion = "average")
dopenness<- predict(disagropenness)
laborforce<- c(1457.1, 1381.7, 1343.1, 1395.4, 1442.7)
laborforcetime<- ts(laborforce, start = 1999)
disagrlaborforce<- td(laborforcetime 1, to = "quarterly", method = "denton-cholette", con-
version = "average")
dlaborforce<- predict(disagrlaborforce)
laborrate<- c(61.114, 60.123, 58.732, 58.074, 60.149)
laborratetime<- ts(laborrate, start = 1999)
disagrlaborrate<- td(laborratetime 1, to = "quarterly", method = "denton-cholette", con-
version = "average")
dlaborrate<- predict(disagrlaborrate)
# all annual variables were uploaded: mtaxburden, munemploymentrate, mopenness, mreg-
ulation, mgdp, mm1, mlaborforce, mruleoflaw, mgoveffect, mpolitstab, mcorrindex, mlabor-
rate
# all quarterly variables were uploaded: ktaxburden, kunemploymentrate, kopenness, kreg-
ulation, kgdp, km1, klaborforce, kruleoflaw, kgoveffect, kpolitstab, kcorrindex, klaborrate
# for stationarity testing, the same functions were applied for all variables:
library(forecast)
library(urca)
```

auto.arima(variable)
adf.test(variable)
pp.test(variable)
kpsstau<-ur.kpss(variablewithtrend, type = 'tau')
summary(kpsstau)
kpssmu<-ur.kpss(variablewithconstant, type = 'mu')
summary(kpssmu)
# correlations matrix were get from (firstly, names of variables were changed):
library(corrplot)
annualdata<- cbind(corrindex, taxburden, unemploymentrate, openness, regulation, gdp, m1, ruleoflaw, goveffect, politstab, laborrate, laborforce)
annualdata<- cor(annualdata)
corrplot(annualdata)
quarterlydata<- cbind(corrindex, taxburden, unemploymentrate, openness, regulation, gdp, m1, ruleoflaw, goveffect, politstab, laborrate, laborforce)
quarterlydata<- cor(quarterlydata)
corrplot(quarterlydata)
# for normality testing the same functions were applied for all variables:
library(rockchalk)
shapiro.test(variable)
kurtosis(variable)
skewness(variable)
library(ggplot2)
$ggplot(data.frame, aes(x = variable)) + geom_{h}istogram(binwidth = 5)$
hist(variable)
qqnorm(variable)
# models:
library(lavaan)
library(semPlot)
# first differences of variables were taken using diff function: difmtaxburden, difmunemploymentrate, difmopenness, difmregulation, difmgdp, difmm1, difmlaborforce, difmruleoflaw, difmgoveffect, difmpolitstab, difmcorrindex, difmlaborrate, difktaxburden, difkunemploymentrate, difkopenness, difkregulation, difkgdp,difkm1, difklaborforce, difkruleoflaw, difkgoveffect, difkpolitstab, difkcorrindex, difklaborrate diflogmgdp<- diff(log(mgdp))
diflogmopenness<- diff(log(mopenness))
diflogmm1<- diff(log(mm1)) shadowannual<- cbind(diflogmgdp, difmregulation,diflogmopenness, difmlaborrate, diflogmm1, difmunemploymentrate)
$matiniaip < -'seselis =\sim diflogmm1 + diflogmgdp + difmlaborrate$
$seselis \sim difmunemploymentrate + difmregulation + diflogmopenness'$
matiniaipcfa<- cfa(model= matiniaip, data=shadowannual, test = "Yuan-Bentler")
summary(matiniaipcfa, standardized=TRUE, fit.measures=TRUE)
semPaths(matiniaipcfa, whatLabels = "std", rotation = 2)
diflogmcorrindex<- diff(log(mcorrindex))
corrannual<- cbind(difmpolitstab,diflogmcorrindex, difmcorrindex, difmruleoflaw, difmgoveffect)

$korupcija < -'korupcija =\sim diflogmcorrindex + difmpolitstab$
$korupcija \sim difmgoveffect + difmruleoflaw'$
korupcijacfa<- cfa(model= korupcija, data=corrannual, test = "Yuan-Bentler")
summary(korupcijacfa, standardized=TRUE, fit.measures=TRUE)
semPaths(korupcijacfa, whatLabels = "std", rotation = 2)
diflogmlaborforce<- diff(log(mlaborforce))
jointannual<- cbind(diflogmm1,diflogmcorrindex, diflogmopenness, difmregulation, difmpolitstab, diflogmlaborforce, difmunemploymentrate, difmgoveffect, difmruleoflaw)
$bendrasis < -'seselis =\sim diflogmm1 + diflogmlaborforce$
$seselis \sim diflogmopenness + difmregulation$
$korupc =\sim diflogmcorrindex + difmpolitstab$
$korupc \sim difmgoveffect + difmruleoflaw + diflogmopenness$
$seselis \sim\sim korupc'$
bendrasiscfa<- cfa(model= bendrasis, data=jointannual, test = "Yuan-Bentler")
summary(bendrasiscfa, standardized=TRUE, fit.measures=TRUE)
semPaths(bendrasiscfa, whatLabels = "std", rotation = 2)
diflogkm1<- diff(log(km1))
diflogkgdp<- diff(log(kgdp))
shadowquarterly<- cbind(diflogkm1, diflogkgdp, difkunemploymentrate, difktaxburden)
$seselisk < -'seselis =\sim diflogkm1 + diflogkgdp$
$seselis \sim difkunemploymentrate + difktaxburden'$
seselisk<- cfa(model= seselisk, data=shadowquarterly, test = "Yuan-Bentler")
summary(seselisk, standardized=TRUE, fit.measures=TRUE)
semPaths(seselisk, whatLabels = "std", rotation = 2)
diflogkcorrinex<- diff(log(kcorrindex))
corrquarterly<- cbind(diflogkcorrindex, difkpolitstab, difkruleoflaw, difkgoveffect)
$korupcijak < -'korupcija =\sim diflogkcorrindex + difkpolitstab$
$korupcija \sim difkruleoflaw + difkgoveffect'$
korupcijak<- cfa(model= korupcijak, data=corrquarterly, test = "Yuan-Bentler")
summary(korupcijak, standardized=TRUE, fit.measures=TRUE)
semPaths(korupcijak, whatLabels = "std", rotation = 2)
diflogkopenness<- diff(log(kopenness))
jointquarterly<- cbind(diflogkm1,diflogkgdp,diflogkopenness, difktaxburden,diflogkcorrindex, difkpolitstab, difkgoveffect, difkruleoflaw)
$bendras < -'seselis =\sim diflogkm1 + diflogkgdp$
$seselis \sim diflogkopenness + difktaxburden$
$korupc =\sim diflogkcorrindex + difkpolitstab$
$korupc \sim difkgoveffect + difkruleoflaw$
$seselis \sim\sim korupc'$
bendrascfa<- cfa(model= bendras, data=jointquarterly, test = "Yuan-Bentler")
summary(bendrascfa, standardized=TRUE, fit.measures=TRUE)
semPaths(bendrascfa, whatLabels = "std", rotation = 2)