



## IDENTIFICATION OF THE SHADOW ECONOMY DETERMINANTS FOR THE EUROZONE MEMBER STATES: APPLICATION OF THE MIMIC MODEL

Rita REMEIKIENĖ<sup>1\*</sup>, Ligita GASPARĖNIENĖ<sup>2</sup>,  
Viktoras CHADYŠAS<sup>3</sup>, Martin CEPEL<sup>4</sup>

<sup>1</sup>, <sup>2</sup>*Lithuanian Institute of Agrarian Economics, V. Kudirkos 18-2, LT-03105, Vilnius, Lithuania*

<sup>3</sup>*Vilnius Gediminas Technical University, Saulėtekio al. 11, LT-10223 Vilnius, Lithuania*

<sup>4</sup>*LIGS University LLC Honolulu, Richards Street, Suite 836, Hawaii, USA*

Received 10 December 2017; accepted 19 September 2018

**Abstract.** This article is aimed at identification of the shadow economy's causal factors and indicators in 19 Eurozone member states over the period from 2005 to 2016. Application of the MIMIC model has allowed to identify the following causal factors of the shadow economy in the Eurozone: employment rate, gender wage gap and income inequalities (expressed as the GINI index). All of these causal factors of the shadow economy in the Eurozone are attributable to the group of labour market determinants, which proposes that a reasonably arranged labour market mechanism can substantially diminish the probability of the shadow economy emergence. On the other hand, it has been found that the level of the shadow economy determines a positive/negative degree of the public trust in the EU authorities. The novelty of the research lies in the disclosure of the main causal factors of the shadow economy in the geographical area that covers different countries with a single currency. The findings of this research may contribute to the development of the shadow economy reduction strategies in 19 Eurozone member states.

**Keywords:** shadow economy, the MIMIC model, Eurozone member states, causal factors, indicators, labour market determinants.

**JEL Classification:** E26, O17.

### Introduction

According to Schneider, Raczkowski, and Mróz (2015), the shadow economy is a natural element of our economic and social life. For a comprehensive understanding of the roots of the shadow economy, it is advisable not only to estimate the current size of this phenomenon, but also to identify its most influential causal factors. Identification of the shadow economy's causal factors and indicators is one of the key issues in the area of the shadow economy

---

\*Corresponding author. E-mail: [rita.remeikiene@laei.lt](mailto:rita.remeikiene@laei.lt)

research and a real help for public authorities that have been trying to combat this phenomenon. Perceiving the shadow economy's causal factors and indicators, public authorities would be capable of developing purposeful countermeasures for the reduction of the level of the shadow economy in particular countries or regions.

The relevance of this topic as well as its practical applicability are confirmed by the newest scientific studies in this area: Schneider et al. (2015) found that the main determinants of the shadow economy in 31 European country over the period 2003–2014 include unemployment, self-employment and tax morale. With reference to Medina and Schneider's (2017) estimations, the main factors that determined the fluctuations in the size of the shadow economy in 158 countries worldwide over the period 1990–2015 covered trade openness, unemployment rate, GDP per capita, the size of the governmental sector, fiscal freedom and corruption rate. The research in the average driving forces of the shadow economy in 38 OECD member countries with application of the MIMIC model revealed that personal income tax (13.8%), indirect taxes (14.1%), tax morale (14.5%), unemployment (14.7%), self-employment (14.5%), growth of GDP (14.3%) and business freedom index (14.2%) more or less evenly contribute to the level of the shadow economies in the OECD-38 (Elgin & Schneider, 2016). Achim, Borlea, Gaban, and Cuceu (2018) emphasised the importance of human happiness (according to the authors, happy taxpayers are characterized by a lower propensity towards the shadow economy in the EU countries). In general, the main drivers of the shadow economy include tax and social security burdens, tax morale, the quality of public institutions and the regulation of the labour market. A reduction in the tax burden is therefore likely to lead to a reduction in the size of the shadow economy (Schneider & Williams, 2016).

The Eurozone member states, selected for this research, share some common features: first, before accessing the Eurozone, they had to meet the convergence criteria (particular economic and legal requirements established in the Treaty of Maastricht in 1992); second they use the single currency. Hence, the results of this research could contribute to the development of the shadow economy prevention measures which, in their turn, could help to reduce the level of the shadow economy in all Eurozone member states. The scientific problem can be stated as follows: which causal factors affect the level of the shadow economy in the Eurozone?

The main purpose of this article is to identify the shadow economy's causal factors and indicators in 19 Eurozone member states over the period from 2005 to 2016. For the fulfilment of the defined purpose, the following objectives were raised: 1) to conduct the theoretical analysis of the methods to measure the size of the shadow economy; 2) to introduce the main principles of the MIMIC model; 3) to empirically verify which of the causal factors had the most significant impact on the level of the shadow economy in 19 Eurozone member states over the period from 2005 to 2016.

The novelty of the research: 1) the research, based on the MIMIC model, has revealed that the level of the shadow economy in the Eurozone member states is not affected by such causal factors as currency inflation, governmental spending, the size of government or government consumption, while in less developed countries without a single currency the above-mentioned factors are main determinants of the shadow economy; 2) the indicators

of the labour market are important causals of the shadow economy in both the euro and non-euro area countries; in this case, the shadow economy in the Eurozone is positively affected by low employment rates, whereas the shadow economy in non-euro area countries is positively affected by the level of unemployment. The Eurozone member states are facing the problems which manifest as income inequalities and the decline in the middle class. These problems, in their turn, promote wage gaps and cause the at-risk-of-poverty rate to rise. The causals of the shadow economy, disclosed by this research, determine such high levels of this phenomenon that employment of individual strategies cannot diminish their destructive impacts. Hence, consideration of the results of this research may contribute to the development of joint shadow economy prevention strategies.

The methods of the research include scientific literature analysis, statistical data analysis and the MIMIC model.

## **1. The methodologies to estimate the size of the shadow economy: literature review**

“The existence of shadow economies has potentially serious implications for economic performance” (Bose, Capasso, & Wurm, 2012, p. 620) since shadow economies refer to illegal, unreported or unrecorded activities “driven by profit, tax evasion, gain or circumvention of legal regulations” (Gasparėnienė, Remeikienė, & Schneider, 2017, p. 275). Scientific literature proposes a wide variety of the definitions of the shadow (underground) economy, but different interpretations make the accurate estimation of the size of this phenomenon a real challenge (Remeikienė, Gasparėnienė, & Schneider, 2018). While there remains little doubt that underground activities amount to comparatively large shares of many economies (whether developed or developing), scientists are still having discussions concerning the appropriate measure of the level of the shadow economy. The vast majority of the scientific studies focus on the role of public policy and public administration, and highlight the importance of such causal factors as tax burden, social security contribution burden, complexity and perceived unfairness of the local tax system, bureaucracy, complexity and instability of laws, and corruption (Schneider, 2007; Williams, 2009, 2010; Bose et al., 2012; Buehn & Schneider, 2012; Schneider & Buehn, 2013; Schneider et al., 2015 and others). Nevertheless, different methodologies lean on slightly different causal factors and indicators, which, according to Bose et al. (2012), is determined by the hidden nature of the shadow economy itself. With reference to Dey, Russell, and Thomson (2011) and Tregidga (2017), the shadow economy includes not only undeclared activities, but also the practice of presenting shadow reports “prepared according to different institutional and ideological rules” (Dey et al., 2011, p. 64), but actually mirroring the scopes of entities’ “social, environmental and/or economic performance and impacts” (Tregidga, 2017, p. 512). The hidden nature of the shadow economy makes it difficult to capture and measure.

The analysis of scientific literature has revealed that thus far the methods of the estimation of the size of the shadow economy have ranged from direct surveys to indirect methods and models based either on statistical data analysis or employment of representative indicators (Williams, 2010; Williams & Nadin, 2012; Buehn & Schneider, 2012; Juškieienė, 2015, etc.).

The detailed summary of the main approaches, methods and data types to help to estimate the size of the shadow economy is presented in Table 1.

Table 1. The main shadow economy estimation methodologies (source: compiled by the authors)

Approach	Data type	Methods		Scientific sources
Direct approach	Qualitative	Sample surveys, interviews, expert evaluations, micro-surveys of informal sector, compliance methods, in-depth audit		Williams (2009); Williams and Nadin (2012, 2013)
Indirect approach	Quantitative	Monetary	Denomination of bank notes, cash contribution ratio, currency ratio/demand method, velocity of money, transaction method	M. D. Fethi, S. Fethi, and Katircioglu (2006); Georgiou (2007); Williams (2009); Schneider and Buehn (2013); Williams and Nadin (2012); Putniņš and Sauka (2014)
		Income-expenditure	GDP discrepancies, income/expenditure discrepancies, consumer expenditure, national accounting aggregates	Williams et al. (2007); Georgiou (2007); Sookram et al. (2009); Hatipoglu and Ozbek (2011); AT Kearney (2013); Schneider et al. (2015); Juškienė (2015)
		Non-monetary	Ranking, electricity consumption method, detection-controlled estimation, the number of small and medium enterprises, employment discrepancies	Kaufmann and Kaliberda (1996); Lacko (1998); Sookram et al. (2009); Schneider and Buehn (2013, 2016); Schneider et al. (2015)
Multiple causes approach	Multiple	DYMIMIC, MIMIC, DGE, SEM		Elgin and Oztunali (2012); Elgin and Schneider (2013); Schneider and Buehn (2013, 2016); Trebicka (2014); Schneider et al. (2015)

The data in Table 1 show that the main methodologies of the estimation of the size of the shadow economy can be attributed to the categories of direct, indirect and multiple-causes approach. The methods of direct approach are referred to as microeconomic methods because they investigate agents' behaviour, the changes in business resources and market structures, and rely on respondents' voluntary replies or tax auditing (Schneider & Buehn, 2013). As the methods of direct approach (sample surveys, interviews, expert evaluations, etc.) allow to identify the shadow economy's determinants, the types of goods and services commonly traded in the "shadow", the channels of the shadow economy, the characteristics of the shadow economy participants (households and business entities) and the plausible changes in the general economics (predicted size of the shadow economy and the size of the shadow economy in particular sectors), they are considered to be capable of disclosing the detailed information about the structure of the shadow economy (Williams, Round, & Rodgers, 2007; Williams & Nadin, 2012; Schneider & Buehn, 2013; Schneider et al., 2015). What is more, application of a direct method may help to develop new shadow economy indices which, unlike the indices derived from macroeconomic indicators, require few assumptions

and can be used through time and across particular sectors or economies. For instance, by employing the method of the survey of business managers, Putninš and Sauka (2015) derived the index that combines the estimates of misreported business income, unregistered or hidden employees and unreported wages, and measures the size of the shadow economy as a percentage of GDP. This method provides the comprehensive information on the structure of the shadow economy, in particular, in the service and manufacturing sectors. The Labour Force Survey (LFS) in combination with enterprise surveys or structural business statistics (SBS) help to generate a harmonized database and develop the Labour Input Method that estimates the scale of undeclared work by measuring the discrepancy between the reported labour supply and the reported use of labour (Williams et al., 2017).

It should also be noted that the use of microeconomic methods is not restricted within representative surveys, interviews or expert evaluations. Lichard, Hanousek, and Filer (2014) estimated the size of the shadow economy in the Czech and Slovak Republic by using the household data based on the consumption-income gap. As it was indicated by Schneider (2017), the main advantage of this method lies in prevention of “making the unrealistic assumption which leads to underestimating the size of the shadow economy” (Schneider, 2017, p. 9). In addition, the authors (Lichard et al., 2014) avoid the problem of arbitrary *a priori* assignment of individuals to the groups of “evading” or “non-evading” declaration of their income.

Despite the advantages of the methods of direct approach, they are criticized for high costs (Pocius, 2015), data unreliability (respondents express their subjective opinions on the issues under research) (Schneider & Buehn, 2013), sensitivity of the results to wording of survey questions (Schneider et al., 2015), applicability only for small-scale studies (Herwartz, Tafenau, & Schneider, 2013) and failure to cover the groups of hidden population truly acting in the “shadow” (Pocius, 2015).

The methods of indirect approach are referred to as macroeconomic methods because they allow to compare the values of general economic indicators (GDP, national income and national expenditure balance, unemployment rate, currency rates, imports, exports, consumption rate, consumer income and expenditure, average tax rate, etc.) (Sookram, Watson, & Schneider, 2009; Schneider & Buehn, 2013, etc.) within a country or a sector under research. According to Schneider and Buehn (2013), the key advantage of the methods of indirect approach lies in their ability to provide an insight in the indicators that reflect the magnitude of the shadow economy in the course of time. On the other hand, the methods of indirect approach are criticized for data disparities and discrepancies (Williams, 2010; Williams & Nadin, 2012; Schneider et al., 2015), a comparatively narrow scale of applicability (Schneider et al., 2015), reliance on the base year’s statistical data (Pocius, 2015) and failure to consider the indicators of digital shadow economy (e.g. payments in virtual money (cryptocurrencies)) (Gasparėnienė & Remeikienė, 2016).

Finally, the methods of multiple-causes approach (DYMIMIC, MIMIC, DGE, SEM) are referred to as a mixture of micro and macroeconomic methods because they cover a variety of general economic (e.g. GDP per capita), tax environment (e.g. shares of direct and indirect taxation, social insurance contributions), legal environment (e.g. state regulation), societal (e.g. tax morale), labour market (e.g. employment and unemployment rates, average

working time per week) and monetary (e.g. the changes in local currency per capita) indicators of the shadow economy. According to Trebicka (2014), the methods of multiple-causes approach are based on a two-sector (household and business) dynamic general equilibrium. The strategic equations in these methods allow to model the causal relationship between the shadow economy's observable causes (e.g. a change in a tax rate) and unobservable latent variables, while the measurement equations link the unobservable latent variables with their indicators (e.g. the demand for a national currency) (Bose et al., 2012). Inclusion of a comparatively wide dataset is considered to be one of the key advantages of the methods of multiple-causes approach (Bose et al., 2012; Elgin & Schneider, 2013, etc.). Nevertheless, they earn some criticism for reliance on national income statistics (Elgin & Oztunali, 2012), high probability of double counting (Juškienė, 2015) and dependency of the results on the set of selected variables (Pocius, 2015).

Previous studies on the determinants of the shadow economy in non-euro area countries reveal the significance of the impact of these determinants: for instance, in Tanzania (Dell'Anno, Davidescu, & Balele, 2018), some African states (namely, Kenya, Namibia, Ghana and Nigeria) (Nchor & Adamec, 2015), Taiwan (Wang, Lin, & Hu, 2006) and Egypt (Hassan & Schneider, 2016b), the growth in the level of the shadow economy was found to be determined by currency and currency inflation, while the significance of currency and currency inflation to the level of the shadow economy in the Eurozone has not been confirmed thus far. As Tanzania, African states, Taiwan and Egypt are developing countries, the other causal factors, which promote the shadow economies in these countries, include government spending, the size of government and government consumption (all of the above-mentioned causals reflect the ability of a government to efficiently manage public finance). In Egypt, African states and Tanzania, a labour market indicator – unemployment rate – was found to significantly increase the level of the shadow economy.

Summarising, although scientific literature offers a variety of the methodologies for estimation of the size of the shadow economy, the impact of various determinants on the total size of this phenomenon is still hard to measure. Thus far, the methods of multiple-causes approach (in particular, the MIMIC model) have been considered the most reliable in this regard as they rely on a variety of measures rather than on a single measure (which lets reduce the probability of measurement error) and allow to identify “the slope coefficients between the size of the shadow economy and its cause variables without directly observing the latent variable” (Bose et al., 2012, p. 622). This way, the changes in causal variables help to predict the changes in the total size of the shadow economy. For the above-explicated reasons, we select the MIMIC model for our empirical research.

## **2. Theoretical principles and criticism of the MIMIC model**

The MIMIC multiple-causes model is considered the most comprehensive methodology developed for estimation of the size of the shadow economy. In this model, the shadow economy is treated as a latent variable which, on one side, is related to the set of observed indicators, and on the other side – to the set of causal variables that have a considerable impact on the multitude of the researched phenomenon. When a sufficient quantity of indicative

and causal data is available, the model is developed by employing the standard procedures of econometrics.

The shadow economy ( $\eta$ ) is a scalar variable which is linearly described by a set of directly observed variables  $X_1, X_2, \dots, X_q$  and scalar random noise ( $\zeta$ ) (Schneider & Buehn, 2016):

$$\eta' = Y_1 X_1 + Y_2 X_2 + \dots + Y_q X_q + \zeta. \tag{1}$$

The latent (hidden) variable ( $\eta$ ), in its turn, directly describes endogenous variables  $Y_1, \zeta Y_2, \dots, Y_p$ , which are dependent on the levels of scalar noise  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ :

$$\begin{aligned} Y_1 &= \lambda_1 \eta + \varepsilon_1; \\ Y_2 &= \lambda_2 \eta + \varepsilon_2; \\ &(\dots) \\ Y_p &= \lambda_p \eta + \varepsilon_p. \end{aligned} \tag{2}$$

Structural noise ( $\zeta$ ) and estimation errors  $\varepsilon$  have a normal distribution and are linearly independent. Then, the following marking is introduced:

- $X^T = (x_1, x_2, \dots, x_q)$  – observed exogenous variables (causes);
- $Y^T = (Y_1, Y_2, \dots, Y_q)$  – structural parameters (structural model);
- $y^T = (y_1, y_2, \dots, y_q)$  – observed endogenous variables (indicators);
- $\lambda^T = (\lambda_1, \lambda_2, \dots, \lambda_q)$  – structural parameters (estimation model);
- $\varepsilon^T = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_q)$  – estimation errors;
- $\nu^T = (\nu_1, \nu_2, \dots, \nu_q)$  – standard deviation of the estimation errors.

Formulas (1) and (2) can be rewritten as:

$$\eta'_t = Y^T x_t + \zeta_t; \tag{3}$$

$$Y_t = \lambda \eta'_t + \varepsilon_t. \tag{4}$$

It is presumed that  $E(\zeta \varepsilon^T) = 0$ ,  $E(\zeta^2) = \sigma^2$ , and  $E(\varepsilon^T) = \Theta^2$ .

$\Theta_{p \times p}$  refers to a diagonal matrix with  $\nu$  inherent to its diagonal.

The model can be converged into a reduced form, i.e. into a function of observed variables:

$$y = \lambda \left( Y^T x + \zeta \right) + \varepsilon = \Pi x + \nu, \tag{5}$$

here  $\Pi = \lambda Y^T$ , and  $\nu = \lambda \zeta + \varepsilon$ .

This way, the matrix of the covariation of the model is developed:

$$\Sigma = \left( \lambda \left( Y^T \Phi Y + \Psi \right) \right) / \Phi Y \lambda^T + \Theta_\varepsilon \lambda Y^T \Phi / \Phi. \tag{6}$$

The latent (hidden) variable ( $\eta$ ) is invisible, and its value remains unknown. The other parameters of the model have to be evaluated by analysing the links between the observed variables in the dispersion and covariation equations. The main aim is to find the values of parameters  $Y$  and  $\lambda$ , and estimate  $\Sigma$ .

The interaction between the causal variables  $X_q$ , multitude of the shadow economy (i.e. latent variable  $\eta$ ) and indicative variables  $Y_p$  over a particular period is depicted in Figure 1.



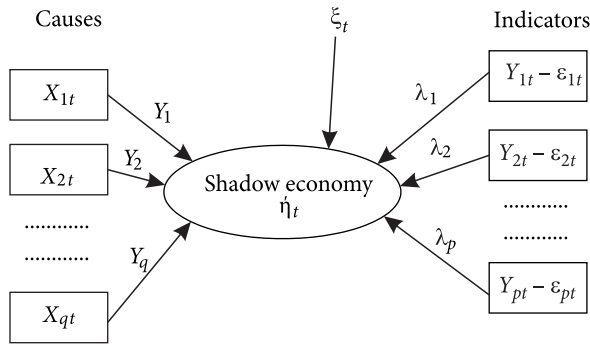


Figure 1. General structure of the MIMIC model

According to Schneider and Buehn (2016), the MIMIC model has its limitations. The weightiest criticism of the model lies in its sampling, debatable reliability and comparatively narrow applicability. The critics of the model mainly disagree on the choice of the latent variables (Helberger & Knepel, 1988; Dell'Anno, 2003, etc.). “The confirmatory rather than exploratory nature of this approach means that one is more likely to determine whether a certain model is valid than to “find” a suitable model” (Schneider & Buehn, 2016, p. 23). In addition, the MIMIC is suspected to plausibly cover the illegal economic activities (e.g. drug trafficking, cigarette smuggling, etc.) which are hard to dissociate from the analysis of the general shadow economy, i.e. the risk of data duplication may emerge. The criticism of the MIMIC model for a high probability of data duplication is common in scientific literature because blind observation of the theoretical presumptions makes it difficult to select appropriate variables and avoid the empirical limitations related to data unavailability.

Helberger and Knepel (1988) note that the MIMIC model generates unstable coefficients due to the changes in a sample size and alternative specifications. Dell' Anno (2003) proved that the stability of the coefficients is lost with an increasing sample size. Another problem with application of the MIMIC model is unavailability of time series data since only simple analytical tools, such as  $q$  and steamleaf plots, are available for residue feature analysis (Dell'Anno, 2003).

One more disadvantage of the MIMIC model lies in its procedure which requires obtainment of “the real world” numbers that reflect the magnitude of shadow activities (Breusch, 2005a, 2005b). When the latent variable and its measures fail to be accurately estimated, the MIMIC may present a set of the coefficients that could be employed for estimation of the index showing the dynamics of unobservable variables. In general, scientific literature proposes that application of the MIMIC for economic estimations is fairly complicated. Nevertheless, the MIMIC model is treated as a valuable tool for identification of the shadow economy's causes and indicators. Moreover, the above-explicated critical arguments should serve as an incentive for further research rather than a reason to reject the method.



### 3. The results of the empirical research

In the first stage of the empirical research, the data on 21 causal factor and 8 indicators in 19 Eurozone member states were collected. The data set included some missing values (see Table 2).

Table 2. Statistics of the shadow economy's causal factors and indicators

Causal factors		N	Mean	St. Deviation	Missing	
					Count	Percent
$X_1$	Employment	201	64.53	5.698	8	3.8
$X_2$	Unemployment	209	832.52	1262.016	0	0
$X_3^*$	Minimum wage	156	855.08	6.550	53	25.4
$X_4$	Gender pay gap	163	15.29	4228.637	46	22
$X_5$	Expenditure of social protection, percent of GDP	174	6654.65	254596.837	35	16.7
$X_6$	Exports	209	173513.25	221456.378	0	.0
$X_7$	Imports	209	167648.43	.189	0	.0
$X_8$	GDP per capita in PPS	209	.92	47850.579	0	.0
$X_9$	International balance in payments <sup>35</sup>	207	8304.35	4.755	2	1.0
$X_{10}$	Labour cost index	201	3.01	5319.179	8	3.8
$X_{11}$	People at risk of poverty	200	3849.11	4.658	9	4.3
$X_{12}$	GINI coefficient	200	47.97	1132.497	9	4.3
$X_{13}$	Expenditure social protection index	172	1822.08	15179.097	37	17.7
$X_{14}$	Expenditure disability index	174	9831.47	24562347.893	35	16.7
$X_{15}$	Population	209	17634233.86	34424.311	0	.0
$X_{16}^*$	Expenditure on education	104	25201.63	15.954	105	50.2
$X_{17}$	Household level of internet access	209	69.06	16.487	0	.0
$X_{18}$	Individuals using the internet	209	27.70	15.495	0	.0
$X_{19}$	Individuals using the internet for interaction with public authorities	171	46.09	127193.644	38	18.2
$X_{20}$	Emigration	188	87737.91	213231	21	10
$X_{21}$	Immigration	188	132723.14	372	21	10
$X_{22}$	Overcrowding rate	181	16.11	14.885	28	13.4
Indicators						
$Y_1$	Shares of environmental and labour taxes in total tax revenues from taxes and social contributions	190	2.51	.577	19	9.1

End of Table 2

Causal factors		N	Mean	St. Deviation	Missing	
					Count	Percent
Y <sub>2</sub>	Level of citizens' confidence in EU institutions	171	52.27	10.458	38	18.2
Y <sub>3</sub>	Young people in employment	209	13.783	4.9036	0	0
Y <sub>4</sub> *	Inflows of FDI	124	136301.07	201390.327	86	41.1
Y <sub>5</sub>	Self-employment level	209	1058.75	1445.235	0	0
Y <sub>6</sub>	Poverty rate after social transfers	200	2822.33	4006.848	9	4.3
Y <sub>7</sub>	Real labour productivity	209	101.31	5.356	0	.0
Y <sub>8</sub>	Real effective exchange rate	209	.140	2.7454	0	.0

Note: \* will be deleted in the further calculations.

In the second stage of the empirical research, the preconditions for the development of the MIMIC model were verified. The data in Table 3 show that the correlation between the causal variables is strong since the values of some correlation coefficients are close to unit (e.g. the correlation between variables X<sub>5</sub> and X<sub>13</sub>, or X<sub>5</sub> and X<sub>8</sub>, etc.), while the correlation between the indicative variables is less strong since the value of the largest coefficient is equal to only -0.439 (see Table 4).

Table 3. Correlation matrix of causal variables

	X <sub>1</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>8</sub>	X <sub>12</sub>	X <sub>13</sub>	X <sub>17</sub>	X <sub>18</sub>	X <sub>19</sub>
X <sub>1</sub> Pearson Correlation	1	.457**	.314**	.284**	-.170*	.404**	.335**	.464**	.518**
Sig. (2-tailed)		.000	.000	.000	.019	.000	.000	.000	.000
N	201	156	166	201	192	164	201	201	165
X <sub>4</sub> Pearson Correlation	.457**	1	-.138	-.098	.140	-.104	.016	.105	.271**
Sig. (2-tailed)	.000		.093	.215	.074	.209	.634	.164	.001
N	156	163	150	163	163	148	163	163	144
X <sub>5</sub> Pearson Correlation	.314**	-.138	1	<b>.905**</b>	.070	<b>.968**</b>	.592**	.749**	.572**
Sig. (2-tailed)	.000	.093		.000	.361	.000	.000	.000	.000
N	166	150	174	174	174	172	174	174	136
X <sub>8</sub> Pearson Correlation	.784**	-.098	<b>.905**</b>	1	.065	.900**	.447**	.616**	.436**
Sig. (2-tailed)	.000	.215	.000		.361	.000	.000	.000	.000
N	201	163	174	209	200	172	209	209	171
X <sub>12</sub> Pearson Correlation	-.170*	.140	.070	.065	1	.062	.074	.094	-.021
Sig. (2-tailed)	.019	.074	.361	.361		.417	.301	.184	.787
N	192	163	174	200	200	172	200	200	162

End of Table 3

	X <sub>1</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>8</sub>	X <sub>12</sub>	X <sub>13</sub>	X <sub>17</sub>	X <sub>18</sub>	X <sub>19</sub>
X <sub>13</sub> Pearson Correlation	.404**	-.104	<b>.968**</b>	<b>.900**</b>	.062	1	.592**	.774**	.598**
Sig. (2-tailed)	.000	.200	.000	.000	.417		.000	.000	.000
N	164	148	172	172	172	172	172	172	134
X <sub>17</sub> Pearson Correlation	.335**	.016	.592**	.447**	.074	.592**	1	.888**	.838**
Sig. (2-tailed)	.000	.934	.000	.000	.301	.000		.000	.000
N	201	163	174	209	200	172	209	209	171
X <sub>18</sub> Pearson Correlation	.464**	.105	.749**	.616**	.094	.774**	.888**	1	.808**
Sig. (2-tailed)	.000	.184	.000	.000	.184	.000	.000		.000
N	201	163	174	209	200	172	209	209	171
X <sub>19</sub> Pearson Correlation	.518**	.271**	.572**	-.021	-.021	.598**	.638**	.808**	1
Sig. (2-tailed)	.000	.001	.000	.787	.787	.000	.000	.000	
N	165	144	136	162	162	134	171	171	171

Table 4. Correlation matrix of indicative variables

	Y <sub>1</sub>	Y <sub>2</sub>	Y <sub>3</sub>
Y <sub>1</sub> Pearson Correlation	1	-0.06	-.077
Sig. (2-tailed)		.441	.294
N	190	171	190
Y <sub>2</sub> Pearson Correlation	-0.06	1	<b>-.439**</b>
Sig. (2-tailed)	.441		.000
N	171	171	171
Y <sub>3</sub> Pearson Correlation	-.077	-.439**	1
Sig. (2-tailed)	.294	.000	
N	190	171	209

Multicollinearity between the variables is also disclosed by VIF (Variance inflation index): if multicollinearity between the variables is not observed,  $VIF < 4$ . Otherwise, one of the variables is eliminated. The calculations showed that  $VIF > 4$  for variables X<sub>5</sub>, X<sub>8</sub>, X<sub>13</sub>, X<sub>17</sub>, X<sub>18</sub>, X<sub>19</sub>. Variable X<sub>13</sub> with the highest VIF value ( $VIF = 21.439$ ) is eliminated. After elimination of X<sub>13</sub>, the regression equation for the indicative variables is developed. Then, the variables with higher VIF values are eliminated (the correlation coefficients are also considered). After elimination of variables X<sub>5</sub>, X<sub>18</sub> and X<sub>19</sub>, a new regression equation is developed (see Table 5).

In the third stage of the research, the MIMIC model was developed by employing LISREL 9.30 for Windows. The initial MIMIC 5-1-3 covers all the causal and indicative variables (see Figure 2). The description of the calculation procedure is presented in the annex.

The selection of the indicative variables was based on the following reasons: *shares environmental and labour taxes* was selected as the main indicative variable because the level of the shadow economy significantly affects the amount of the taxes collected, especially in the labour market. The findings of some previous studies (Dell'Anno et al., 2018; Hassan & Schneider, 2016a) disclosed that the labour market, labour taxes and the shadow economy

directly affect each other; *the level of citizen’s confidence in EU institutions* was selected as the second indicative variable because the level of the shadow economy affects citizens’ confidence in public institutions (although the links between citizens’ confidence in public institutions and the shadow economy have hardly been researched in previous studies, the authors of this article are of the opinion that if a state’s government tolerates the shadow economy and the related phenomena, it inevitably leads to the reduction of the public morale and the decision to participate in shadow activities); *young people in employment* was selected as the third indicative variable minding the fact that the high youth unemployment rate was one of the most urgent problems which promoted the EU Youth Employment Initiative (2016).

Table 5. VIF values after elimination of  $X_5$ ,  $X_{18}$  and  $X_{19}$

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
Constant	2.639	.681		3.877	.000		
$X_1$	.033	.009	.318	3.569	.000	.617	1.621
$X_4$	-.033	.007	-.398	-4.785	.000	.707	1.415
$X_8$	-.142	.251	-.049	-.567	.572	.664	1.506
$X_{12}$	-.045	.009	-.347	-4.815	.000	.940	1.064
$X_{17}$	.008	.003	.193	2.324	.022	.709	1.410

Note: a – Dependent Variable:  $Y_1$ .

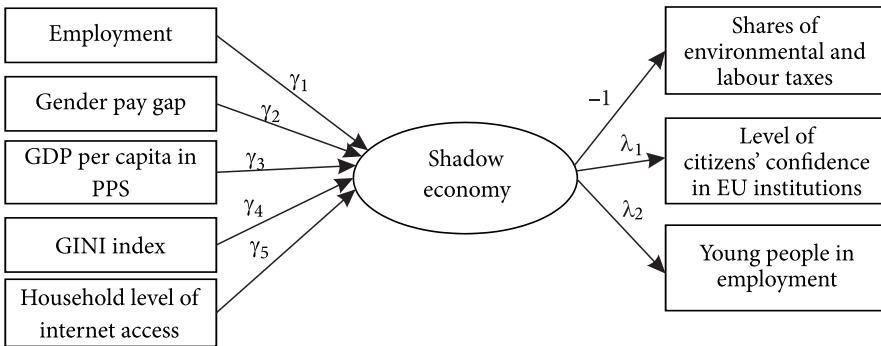


Figure 2. Development of the MIMIC 5-1-3 model for the Eurozone member states

The summary of the developed MIMIC models with their statistical values and different test results is presented in Table 6.

The significance of the variables in the models was verified by employing t-statistics:  $|t\text{-statistic}| > 1.96$  indicated significant variables. As variable  $X_8$  did not match the requirements of t-statistics, it was eliminated from MIMIC 4-1-3. After verification of the latter model, variable  $X_{17}$  was eliminated due to its lower value of t-statistics in comparison to the corresponding values of variables  $X_1$ ,  $X_4$  and  $X_{12}$ .

Table 6. Output – LISREL - MIMIC model summary

Model	$X_1$ (t-stat.)	$X_4$	$X_8$	$X_{12}$	$X_{17}$	$X_2$	$X_3$	$\chi^2$ (p-value)	RMSEA (p-value)	GFI	NFI	AIC
MIMIC 5-1-3	-0.0460* (-5.75)	0.0152* (4.71)	0.1200 (1.56)	0.0174* (4.36)	-0.0038* (-3.31)	0.822 (0.31)	15.911 (5.13)	88.66 (0.00)	0.237 (0.00)	0.889	0.786	3470
MIMIC 4-1-3	-0.0459* (-5.96)	0.0147* (4.73)	-	0.0187* (4.52)	-0.0032* (-3.03)	-0.457 (-0.17)	15.429 (5.31)	87.30 (0.00)	0.266 (0.00)	0.877	0.753	3840
MIMIC 3-1-3	-0.0499 (-6.06)	0.0153 (4.71)	-	0.0179 (4.30)	-	-5.496 (-1.70)	15.349 (5.49)	77.34 (0.00)	0.291 (0.00)	0.871	0.751	2975
MIMIC 5-1-2	-0.0331 (-4.20)	0.0247 (4.20)	0.0461 (0.23)	0.0547 (6.38)	-0.0030 (-1.16)	-9.614 (-3.15)	-	15.32 (0.00)	0.192 (0.02)	0.971	0.926	3021
MIMIC 4-1-2	-0.0325 (-4.34)	0.0243 (4.27)	-	0.0548 (6.42)	-0.00277 (-1.16)	-9.675 (-3.16)	-	14.82 (0.00)	0.168 (0.01)	0.968	0.901	3388
MIMIC 3-1-2	-0.0345 (-4.91)	0.0241 (4.34)	-	0.0528 (6.25)	-	-10.659 (-3.30)	-	9.08 (0.01)	0.159 (0.03)	0.976	0.924	256

Notes: t-statistics are given in parentheses; \* – Means |t-statistic| > 1.96.

The same calculations were performed for MIMIC 5-1-2, 4-1-2 and 3-1-2 models.

As MIMIC 3-1-2 best fitted all the test values (RMSEA was equal to 0.159, GFI and NFI were close to 0.95, AIC value for MIMIC 3-1-2 was lowest), the mathematical equation of the model was expressed as follows:

$$\eta' = -0.0345X_1 + 0.0241X_4 + 0.0528X_{12} . \tag{7}$$

Lower *employment rate* in the Eurozone over the period 2005–2016 (the value of coefficient  $X_1$  was equal to -0.0345) raised the level of the shadow economy, i.e. when the value of  $X_1$  decreased by 1, the size of the shadow economy increased by 0.0345.

The number of the employed is commonly expressed as the number of hired employees. According to Bordignon and Zanardi (1997), a higher number of hired employees in economics determines a lower level of the shadow economy since hired employees are least inclined to evade taxes, whereas family business participants as well as the self-employed (working under a business licence or an individual activity certificate) are more inclined to illegal activities than hired employees with employment contracts. Family business participants and the self-employed are positioned closer to final consumers, with whom they can establish oral agreements and this way evade indirect taxes. In addition, small entrepreneurs are less frequently audited by public authorities. Aiming at evaluating the degree to which shadow work is conducted by the population in formal jobs, Williams and Horodnic (2016) reported the results of a 2007 survey involving 26,659 interviews in 27 EU member states. Their survey revealed that the population in formal jobs “undertake a disproportionate share of work in the shadow economy and benefit disproportionately from the shadow economy” (Williams & Horodnic, 2016, p. 411). The results of Williams and Horodnic’s (2016) research supported the thesis of reinforcement (proposing that the population in formal employment disproportionately participate in the shadow economy) rather than the thesis of marginalization (suggesting that shadow work is concentrated

among the marginalized population, such as the unemployed). The results of our research have confirmed that a large number of hired employees unambiguously negatively affects the size of the shadow economy.

Although the size and characteristics of *gender wage gap* may differ subject to the level of earnings (Rokicka & Ruzik, 2010), the hypothesis that the institutional framework of the labour market and gender wage gap are interrelated was confirmed by previous scientific studies (Tansel, 2001; Rokicka & Ruzik, 2010; Deininger, Jin, & Nagarajan, 2013; Yahmed, 2013 and others). Nevertheless, it must be noted that the findings of previous studies are rather contradictory. For instance, Tansel (2001), who researched the gender pay differential among the workers with and without social security coverage in Turkey, found that the adjusted gender wage gap is more substantial among formal (i.e. social security covered) than among informal (i.e. social security uncovered) workers. Rokicka and Ruzik's (2010) findings (the authors analysed the gender pay gap in the formal and informal labour markets in Poland) suggest the opposite: the informal economy is characterized by a larger gender gap than the formal economy. Yahmed's (2013) study on the inequalities of gender wage in Brazil revealed that the total average gender wage gap is positive and significant in both the formal and informal economies. The results of the empirical research introduced in this article propose that the increases in gender wage gap ( $X_4$ ) raise the size of the shadow economy in the Eurozone member states, and vice versa.

The effort to find the links between *income distribution* and the informal economy is not new. Most studies confirm the hypothesis that higher income inequality (often expressed as GINI coefficient) raises the level of the informal economy (Chong & Gradstein, 2004; Winkelried, 2005; Krstič & Sanfey, 2010 and others). On the other hand, Schneider and Enste (2000) argue that namely the development of the systems of social welfare aimed at flattening income inequalities raises the size of the informal economy because people with high income are deprived from a substantial part of this income (the systems of social welfare are commonly characterized by high taxes), which, in its turn, reduces the motivation to work formally.

The results of this research show that *GINI coefficient* is the third causal factor that affects the size of the shadow economy in 19 Eurozone member states. It refers to an economic derivative which reflects income disparities in domestic households. It should be noted that GINI coefficient shows distribution of income rather than distribution of assets. The MIMIC model developed for this research has disclosed that the growth in income inequality raises the size of the shadow economy in the Eurozone.

## Conclusions

Scientific literature, which is rich in qualitative and quantitative methodologies to estimate the size of the shadow economy, provides the opportunities to select the method best suited for fulfilment of particular scientific purposes and objectives. This research was aimed at identification of the determinants of the shadow economy in 19 Eurozone member states sharing one common feature – a single currency. The MIMIC model, selected as the main method for this research, helped to identify the following determinants of the shadow economy in

the Eurozone: employment rate, gender wage gap and income inequalities (expressed as GINI index). With reference to the results of the research, it can be concluded that all the determinants which affected the size of the shadow economy in the Eurozone over the period from 2005 to 2016 can be attributed to the group of *labour market determinants*. Hence, a better functioning of the labour market mechanism may significantly improve the statistics of the shadow economy in the euro area. On the other hand, it has been established that the level of the shadow economy determines a positive/negative degree of the public trust in the EU authorities. Therefore, the authorities responsible for combating the shadow economy in the Eurozone should pay special attention to the regulation of the labour market. The reduction of income inequalities, wage gaps and promotion of employment should become the priority measures when dealing with the issues of the shadow economy in the Eurozone.

The empirical research has disclosed the main features of the origin and variance of the shadow economy in the geographical area with a single currency. This study contributes to previous research in the informal (underground) economy through identification of the features and specificities that explain the origin and variance of this phenomenon (in the case under consideration, the determinants of the origin and variance of the shadow economy in the Eurozone mainly cover the labour market determinants). The results of this study have revealed that the MIMIC method allows to identify the determinants of the shadow economy and, therefore, is applicable for the conduct of the comprehensive research in this area (for instance, researchers could employ the surveys and pay the special attention to the labour market determinants grouped by the characteristics of the target areas). The limitations of this study are linked to the missing data values which reduce the number of the factors in the model.

## References

- Achim, M. V., Borlea, S. N., Gaban, L. V., & Cuceu, I. C. (2018). Rethinking the shadow economy in terms of happiness. Evidence for the European Union member states. *Technological and Economic Development of Economy*, 24(1), 199–228. <https://doi.org/10.3846/20294913.2016.1209250>
- AT Kearney. (2013). *The shadow economy in Europe*. Financial Institutions. Retrieved from [https://www.atkearney.com/financial-institutions/featured-article/-/asset\\_publisher/j8IucAqMqEhB/content/the-shadow-economy-in-europe-2013/10192](https://www.atkearney.com/financial-institutions/featured-article/-/asset_publisher/j8IucAqMqEhB/content/the-shadow-economy-in-europe-2013/10192)
- Bordignon, M., & Zanardii, A. (1997). Tax evasion in Italy. *Giornale Degli Economisti e Annali di Economia*, 56, 169–210.
- Bose, N., Capasso, S., & Wurm, M. A. (2012). The impact of banking development on the size of shadow economies. *Journal of Economic Studies*, 32(6), 620–638. <https://doi.org/10.1108/01443581211274584>
- Breusch, T. (2005a). The Canadian underground economy: an examination of Giles and Tedds. *Canadian Tax Journal*, 53, 367–391.
- Breusch, T. (2005b). *Estimating the underground economy using MIMIC models*. EconWPA. Retrieved from <https://econwpa.wustl.edu/eps/em/papers/0507/0507003.pdf>
- Buehn, A., & Schneider, F. (2012). Shadow economies around the world: novel insights, accepted knowledge, and new estimates. *International Tax and Public Finance*, 19, 139–171. <https://doi.org/10.1007/s10797-011-9187-7>
- Chong, A., & Gradstein, M. (2004). *Inequality, institutions and informality*. Inter-American Development Bank. <https://doi.org/10.2139/ssrn.1818716>



- Dey, C., Russell, S., & Thomson, I. (2011). Exploring the potential of shadow accounts in problematizing institutional conduct. In A. Ball, & A. S. Osbourne (Eds.), *Social accounting and public management: accountability for the common good* (Chapter 12). London: Routledge.
- Deininger, K., Jin, S., & Nagarajan, H. (2013). Wage discrimination in India's informal labor markets: exploring the impact of caste and gender. *Review of Development Economics*, 17(1), 130-147. <https://doi.org/10.1111/rode.12020>
- Dell'Anno, R. (2003). *Estimating the shadow economy in Italy: a structural equation approach*. University of Aarhus. Retrieved from [https://ftp.econ.au.dk/afn/wp/03/wp03\\_07.pdf](https://ftp.econ.au.dk/afn/wp/03/wp03_07.pdf)
- Dell'Anno, R., Davidescu, A. A., & Balele, N. P. (2018). Estimating shadow economy in Tanzania: an analysis with the MIMIC approach. *Journal of Economic Studies*, 45(1), 100-113. <https://doi.org/10.1108/JES-11-2016-0240>
- Elgin, C., & Oztunali, O. (2012). *Shadow economies around the world: model based estimates*. Bogazici University. Retrieved from [https://econpapers.repec.org/paper/bouwpaper/2012\\_2f05.htm](https://econpapers.repec.org/paper/bouwpaper/2012_2f05.htm)
- Elgin, C., & Schneider, F. (2013). *Shadow economies in OECD countries: DGE vs MIMIC approaches*. Bogazici University. Retrieved from [https://www.econ.boun.edu.tr/content/wp/EC2013\\_13.pdf](https://www.econ.boun.edu.tr/content/wp/EC2013_13.pdf)
- Elgin, C., & Schneider, F. (2016). Shadow economies in OECD countries: DGE vs MIMIC approaches. *Bogazici Journal Review of Social, Economic and Administrative Studies*, 30(1), 51-75. <https://doi.org/10.21773/boun.30.1.3>
- Fethi, M. D., Fethi, S., & Katircioglu, S. T. (2006). Estimating the size of the Cypriot underground economy: a comparison with European experience. *International Journal of Manpower*, 27(6), 515-534. <https://doi.org/10.1108/01437720610690464>
- Gasparėnienė, L., & Remeikienė, R. (2016). The methodologies of shadow economy estimation in the world and in Lithuania: whether the criterions fixing digital shadow are included? *Procedia Economics and Finance*, 39, 753-760. [https://doi.org/10.1016/S2212-5671\(16\)30277-5](https://doi.org/10.1016/S2212-5671(16)30277-5)
- Gasparėnienė, L., Remeikienė, R., & Schneider, F. G. (2017). Concept, motives and channels of digital shadow economy: consumers' attitude. *Journal of Business Economics and Management*, 18(2), 273-287.
- Georgiou, G. M. (2007). *Measuring the size of the informal economy: a critical review*. Central Bank of Cyprus. Retrieved from [https://www.centralbank.gov.cy/media/pdf/npwpe\\_no1\\_052007\\_\\_\\_pdf](https://www.centralbank.gov.cy/media/pdf/npwpe_no1_052007___pdf)
- Hassan, M., & Schneider, F. (2016a). *Size and development of the shadow economies of 157 countries worldwide: updated and new measures from 1999 to 2013*. IZA Institute of Labour Economics. Retrieved from <https://ftp.iza.org/dp10281.pdf>
- Hassan, M., & Schneider, F. (2016b). *Modelling the Egyptian Shadow Economy: a MIMIC model and currency demand approach*. Retrieved from [https://www.econ.jku.at/%5Cmembers%5CSchneider%5Cfiles%5Cpublications%5C2016%5CModellingtheESE\\_June2016.pdf](https://www.econ.jku.at/%5Cmembers%5CSchneider%5Cfiles%5Cpublications%5C2016%5CModellingtheESE_June2016.pdf)
- Hatipoglu, O., & Ozbek, G. (2011). On the political economy of the informal sector and income redistribution. *European Journal of Law and Economics*, 32(1), 69-87. <https://doi.org/10.1007/s10657-010-9179-6>
- Helberger, C., & Knepel, H. (1988). How big is the shadow economy? A re-analysis of the unobserved variable approach of B. S. Frey and H. Weck Hannemann. *European Economic Review*, 32(4), 965-976. [https://doi.org/10.1016/0014-2921\(88\)90055-4](https://doi.org/10.1016/0014-2921(88)90055-4)
- Herwartz, H., Tafenau, E., & Schneider, F. (2013). One share fits all? Regional variations in the extent of the shadow economy in Europe. *Regional Studies*, 49(9), 1575-1587. <https://doi.org/10.1080/00343404.2013.848034>
- Yahmed, S. B. (2013). *Gender wage gaps in formal and informal jobs, evidence from Brazil*. Editorial Express. Retrieved from [https://editorialexpress.com/cgi-bin/conference/download.cgi?db\\_name=CSAE2014&paper\\_id=537](https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=CSAE2014&paper_id=537)

- Juškienė, G. (2015). *Overview of the methods used to ensure exhaustiveness the national accounts of Lithuania*. Lithuanian Department of Statistics. Retrieved from <https://www.oecd.org/std/na/2069700.pdf>
- Kaufmann, D., & Kaliberda, A. (1996). *Integrating the unofficial economy into the dynamics of post socialist economies: a framework of analyses and evidence* (Policy Research Working Paper Series 1691). Washington, DC: World Bank.
- Krstič, G., & Sanfey, P. (2010). *Earnings inequality and the informal economy: evidence from Serbia*. European Bank for Reconstruction and Development. Retrieved from <https://www.ebrd.com/downloads/research/economics/workingpapers/wp0114.pdf>
- Lacko, M. (1998). The hidden economies of visegrad countries in international comparison: a household electricity approach. In L. Halpern & C. Wyplosz (Eds.), *Hungary: towards a market economy* (Chapter 6). Cambridge: Cambridge University Press.
- Lichard, T., Hanousek, J., & Filer, R. K. (2014). *Hidden in plain sight: using household data to measure the shadow economy* (No. 10483). CEPR Discussion Paper.
- Medina, L., & Schneider, F. (2017). *Shadow economies around the world: new results for 158 countries over 1991-2015*. Johannes Kepler University Linz. Retrieved from [https://www.econ.jku.at/members/Schneider/files/publications/2017/JointPaper\\_LeandroMedina\\_158countries.pdf](https://www.econ.jku.at/members/Schneider/files/publications/2017/JointPaper_LeandroMedina_158countries.pdf)
- Nchor, D., & Adamec, V. (2015). *Unofficial economy estimation by the MIMIC model: the case of Kenya, Namibia, Ghana and Nigeria*. ACTA. <https://doi.org/10.11118/actaun201563062043>
- Organization for Economic Co-Operation and Development. (2002). *Measuring the non-observed economy*. OECD Statistics. Retrieved from <https://www.oecd.org/std/na/1963116.pdf>
- Pocius, A. (2015). Šešėlinės ekonomikos ir neoficialaus užimtumo būklės bei tendencijų įvertinimas Lietuvoje [Evaluation of the shadow economy and informal employment trends in Lithuania]. *Lithuanian Journal of Statistics*, 54(1), 18-32.
- Putniņš, T. J., & Sauka, A. (2014). *Shadow economy index for the Baltic countries 2009-2013*. SSE Riga. Retrieved from <https://www.sseriga.edu/en/centres/csb/shadow-economy-index-for-baltics/>
- Putniņš, T. J., & Sauka, A. (2015). Measuring the shadow economy using company managers. *Journal of Comparative Economics*, 43, 471-490. <https://doi.org/10.1016/j.jce.2014.04.001>
- Remeikienė, R., Gasparėnienė, L., & Schneider, F. G. (2018). The definition of digital shadow economy. *Technological and Economic Development of Economy*, 24(2), 696-717. <https://doi.org/10.3846/20294913.2016.1266530>
- Rokicka, M., & Ruzik, A. (2010). *The gender pay gap in informal employment in Poland*. CERGE-EI and the Global Development Network. Retrieved from [https://www.cerge.cuni.cz/pdf/gdn/rrc/RRCIX\\_68\\_paper\\_01.pdf](https://www.cerge.cuni.cz/pdf/gdn/rrc/RRCIX_68_paper_01.pdf)
- Schneider, F. (2007). *Shadow economies and corruption all over the world: new estimates for 145 countries*. Economics E-Journal. Retrieved from [https://www.economicsejournal.org/economics/journalarticles/2007-9/version\\_1](https://www.economicsejournal.org/economics/journalarticles/2007-9/version_1)
- Schneider, F. (2017). *Implausible large differences in the sizes of the underground economies highly developed European countries? A comparison of different estimation*. CESifo Working Paper Series No. 6522. Retrieved from [https://www.econ.jku.at/members/Schneider/files/publications/2017/EstShadEc\\_OECDCountries.pdf](https://www.econ.jku.at/members/Schneider/files/publications/2017/EstShadEc_OECDCountries.pdf)
- Schneider, F., & Enste, D. (2000). Shadow economies: size, causes, and consequences. *Journal of Economic Literature*, 38, 77-114. <https://doi.org/10.1257/jel.38.1.77>
- Schneider, F., & Buehn, A. (2013). *Estimating the size of the shadow economy: methods, problems and open questions*. CESIFO. Retrieved from <https://www.economics.jku.at/papers/2013/wp1320.pdf>
- Schneider, F., & Buehn, A. (2016). *Estimating the size of the shadow economy: methods, problems and open questions*. The Institute for the Study of Labor. Retrieved from <https://ftp.iza.org/dp9820.pdf>

- Schneider, F., Raczkowski, K., & Mróz, B. (2015). Shadow economy and tax evasion in the EU. *Journal of Money Laundering Control*, 18(1), 34-51. <https://doi.org/10.1108/JMLC-09-2014-0027>
- Schneider, F., & Williams, C. (2016). *The shadow economy*. The Institute of Economics Affairs. Retrieved from <https://iea.org.uk/wp-content/uploads/2016/07/IEA%20Shadow%20Economy%20web%20rev%207.6.13.pdf>
- Sookram, S., Watson, P. K., & Schneider, F. (2009). Characteristics of households in the informal sector of an emerging economy. *Applied Economics*, 41, 3545-3559. <https://doi.org/10.1080/00036840701493733>
- Tansel, A. (2001). *Wage earners, self-employed and gender in the informal sector in Turkey*. The Economic Research Forum. <https://doi.org/10.2139/ssrn.263275>
- Trebicka, B. (2014). *Mimic model: a tool to estimate the shadow economy*. Mediterranean Center of Social and Educational Research. <https://doi.org/10.5901/ajis.2014.v3n6p295>
- Tregidga, H. (2017). Speaking truth to power: analysing shadow reporting as a form of shadow accounting. *Accounting, Auditing & Accountability Journal*, 30(3), 510-533. <https://doi.org/10.1108/AAAJ-01-2015-1942>
- Wang, D. H., Lin, J., & Hu, T. H. (2006). A MIMIC approach to modelling the underground economy in Taiwan. *Physica A: Statistical Mechanics and its Application*, 371(2), 536-542. <https://doi.org/10.1016/j.physa.2006.03.060>
- Williams, C. C. (2009). Rationales for outsourcing domestic services to off-the-books workers. *Journal of Economic Studies*, 36(4), 343-354. <https://doi.org/10.1108/01443580910973565>
- Williams, C. C. (2010). Out of the shadows: explaining the undeclared economy in Baltic countries. *Journal of Baltic Studies*, 41(1), 3-22. <https://doi.org/10.1080/01629770903525282>
- Williams, C. C., Bejakovic, P., Mikulic, D., Franic, J., Kadir, A., & Horodnic, I. A. (2017). *An evaluation of the scale of undeclared work in the European Union and its structural determinants: estimates using the Labour Input Method*. Luxembourg: Publications Office of the European Union.
- Williams, C. C., & Horodnic, I. A. (2016). Beyond the marginalization thesis: evaluating the participation of the formally employed in the shadow economy in the European Union. *Journal of Economic Studies*, 43(3), 400-417. <https://doi.org/10.1108/JES-06-2014-0105>
- Williams, C. C., & Nadin, S. (2012). Tackling entrepreneurship in the informal economy: evaluating the policy options. *Journal of Entrepreneurship and Public Policy*, 1(2), 111-124. <https://doi.org/10.1108/20452101211261408>
- Williams, C. C., & Nadin, S. (2013). Harnessing the hidden enterprise culture. Supporting the formalization of off-the-books business start-ups. *Journal of Small Business and Enterprise Development*, 20(2), 434-447. <https://doi.org/10.1108/14626001311326815>
- Williams, C. C., Round, J., & Rodgers, P. (2007). Beyond the formal/informal economy binary hierarchy. *International Journal of Social Economics*, 34(6), 402-414. <https://doi.org/10.1108/03068290710751812>
- Winkelried, D. (2005). *Income distribution and the size of the informal sector*. The Robinson Rojas Archive. <https://doi.org/10.2139/ssrn.777144>

## APPENDIX

### Description of calculating process

Table 2 depicts the statistics on the missing values. For instance, variable  $X_1$  has 8 missing values, which composes 3.8% of the total number of its values. It can be seen that variable  $X_{16}$  has the largest number of the missing values, which makes 50.2% of the total number of its values. Variables  $Y_4$  and  $X_3$  also possess a significant number of the missing values.

As the percentage of the missing values is comparatively high, we eliminate variables  $X_3$ ,  $Y_4$  and  $X_{16}$  from further research (in Table 2, these variables have been depicted as crossed out). This way, we get the total number of 188 entries. Then we analyse the missing values of the variables in the remaining dataset. Variable  $X_4$  possesses the largest number of the missing values, which amounts to 22% of the total number of its values. The percentage of the missing values estimated for variables  $X_{19}$  and  $Y_2$  amounts to 18.2%. In the latter case, the percentage of the missing values is not so high. Thus, the possibilities of the missing values insertion are considered. For establishment of whether insertion of the missing values is reasonable, MCAR (Missing Completely At Random) test is employed. This test reveals whether insertion of the missing values will not have any significant impact on final results or models in the analysis. The results of the test showed that insertion of the missing values is not reasonable as it might have a significant impact on the final results of the analysis. At last, the entries with the missing values are crossed out, and preconditions for the MIMIC model development are formed. The preconditions for verification include:

- Outliers in the dataset.
- Data normality.
- Multicollinearity.

As the dataset covers 19 different countries, it is obvious that dispersion of the values of particular variables is significant, and a number of outliers can be observed. A comparatively high number of outliers eliminates the opportunity to employ standard model parameter evaluation methods. Therefore, we eliminate the variables with a large number of outliers from our further analysis. Finally, the remaining dataset covers 9 causal factors and 3 indicators. The next step is to find out whether the variables are not multicollinear because the problem of multicollinearity can distort the trends of interdependency between the variables, which, in turn, may determine emergence of a multiplier with an opposite sign in the model. The correlation matrices developed for the causal variables and indicative variables in the remaining dataset have been depicted in Tables 3 and 4.

VIF index in Table 5 does not show any multicollinearities between the indicative variables. Then we verify normality of the variables. In practical applications, the requirement of variable normality is often ignored; the variables are required to be interval on condition that the dataset is not unreasonably small. Nevertheless, for the remaining variables we verify the hypothesis of normality. For this purpose, we employ two tests: Kolmogorov-Smirnov and Shapiro-Wilk. First of them is very sensitive to the amount of the data as the sample is not very large. For this reason, conclusions are made with reference to the results of the second test: two causal variables ( $X_1$  and  $X_8$ ) and two indicative variables ( $Y_1$  and  $Y_3$ ) do not satisfy the hypothesis of normality. These results will be considered while developing the MIMIC model in the further stage of this research.

Estimations of the parameter values do not reveal whether the model is applicable with the research data, i.e. whether the model and the data are aligned. The alignment between the model and the data is disclosed by criterion  $\chi^2$ : a low value of  $\chi^2$  is a positive indicator showing that the model is applicable with the research data. If the model is aligned with the data, value  $p$  (probability level) should be higher than 0.05.

The alignment between the model and the data is also disclosed by some other popular indicators:

- RMSEA (Root Mean Square Error of Approximation) – the result is positive when RMSEA is lower than 0.1.
- GFI (Goodness of Fit Index) – the model is well aligned with the data when GFI is close to 0.95.
- NFI (Normed Fit Index) – the result is positive when NFI is close to 0.95.
- AIC (Akaike) – this indicator is used for model comparison: the models with lower AIC values are considered to be better than the models with higher AIC values.