European Journal of Government and Economics

Vol. 13, No. 2 (2024), pages 189-205
https://doi.org/10.17979/ejge.2024.13.2.9940
Submitted: Sep 26, 2023 Accepted: Apr 30, 2024 Published: Dec 3, 2024

Article

Exploring determinants of VAT gaps using structural equation models: a MIMIC approach

Iva Hasíková [1,*]

[1] Mendel University, Czechia
*_Correspondence_: xhasikov@mendelu.cz

Abstract. This study explores the estimation of the VAT gap using the Multiple Indicators and Multiple Causes (MIMIC) model**,** a specific type of structural equation model. The VAT gap, a significant indicator of tax revenue loss and inefficiency in VAT collection, is treated as a latent variable driven by various causes and reflected through specific indicators. Drawing on data from European countries, the model incorporates variables such as economic openness, government spending, corruption perceptions, and the e-government development index, each serving as proxies for underlying VAT collection challenges. This study presents a first-time application of the MIMIC model to the Czech Republic, estimating the VAT gap of 2003 – 2020 and revealing fluctuations between 27 % and 35 % of total tax liability, with stabilization at 31 % since 2016. The findings underscore policy areas for potential improvements in VAT compliance and collection efficiency, particularly through enhanced digitalization efforts and governance quality in tax administration by highlighting key factors contributing to the VAT gap.

Keywords: VAT gap; determinants of VAT gap; efficiency of VAT collection; MIMIC model; structural equation model
JEL classification: C39 ; C51 ; H26

1. Introduction

Value-added tax (VAT) is one of the most significant sources of tax revenue in most European countries. According to Eurostat (2024), VAT revenues represent around 20 % of total tax revenues, including social contributions. However, VAT is frequently associated with the risk of tax evasion. According to estimates by the European Union (EU) in 2020, the total VAT evasion loss amounted to 93 billion euros, equating to approximately 3,000 euros lost every second throughout the year (Baert, 2023). Tax evasion poses significant threats to the economic environment, tax system, and society. Novysedlák and Palkovičová (2012) argue that tax evasion and avoidance undermine the economic environment, as certain entities gain advantages that may encourage others to adopt similar practices in the long term, even if tax evasion was not their initial intent. Furthermore, tax evasion is a significant source of inequality, regardless of a country's redistribution objectives (Argentiero et al., 2021).

The tax gap provides valuable insight into tax evasion and the efficiency of tax collection. In the context of VAT, the VAT gap has emerged as an indicator of VAT fraud (Moravec et al., 2021). The VAT gap is defined as the difference between the amount of VAT that should be collected under the legislation and the VAT that is actually collected (Carfora et al., 2020). It encompasses revenues lost due to tax fraud, bankruptcies, taxpayer insolvency, or tax liability miscalculations (Gajewski & Jonski, 2022).

Quantifying the VAT gap involves econometric methods that allow for the consideration of various factors, such as the quality of public sector institutions or the level of digitalisation of public administration. One such econometric method is the multiple indicators and multiple causes model (MIMIC). The MIMIC model, a variant of structural equation models, treats the VAT gap as a latent unmeasurable variable that manifests through multiple indicators and is the result of numerous causes (Frey & Weck-Hannemann, 1984; Schneider et al., 2010).

This paper, while developing the MIMIC model, considers variables such as the unemployment rate, the openness of the economy, the index of economic freedom, the corruption perception index (CPI), general government expenditure, final consumption expenditure, and the e-government development index (EGDI) as causes of the VAT gap. The VAT gap is evidenced by the growth of gross domestic product (GDP) per capita and VAT revenue relative to GDP.

The contribution of this paper is to construct a MIMIC model with data from selected European countries; this model will be applied to these countries. Specifically, the MIMIC model will be used to estimate the VAT gap in the Czech Republic. This method has not yet been used in the Czech Republic, providing a new perspective on the VAT gap, not only its extent but also its potential sources.

The Czech Republic, like other European countries, is not exempt from the challenges of tax evasion or the need for its elimination. The issue lies not only in the previously described disruptive effects of tax evasion but also in the loss of public funds, which are critically needed given the debt burden of Czech public finances. In an international comparison, the Czech Republic's debt remains relatively low. However, its debt dynamics have been above average among EU countries in recent years (The Czech Fiscal Council, 2022).

2. Literature review

The VAT gap can be quantified in monetary units and as a proportion of the tax collected relative to the total theoretical VAT liability (VTTL). Ideally, when tax compliance is maximised, and there are no tax credits or deductions, the VAT revenue as a proportion of the VTTL should be one (Cnossen, 2022).

To estimate the VAT gap and the tax gap in general, the literature identifies three main approaches: the bottom-up approach, the top-down approach, and methods based on econometric modelling (Alm, 2012; Kasnauskienė & Krimisieraitė, 2015; Poniatowski et al., 2020). The bottom-up approach relies on tax audits and direct interviews with taxpayers (Alm, 2012). According to Novysedlák and Palkovičová (2012), the bottom-up approach includes methods such as the selection of a random sample of taxpayers and targeted controls based on risk analysis. Gajewski and Jonski

(2022) caution that tax authorities using the latter method target taxpayers with the highest expected value of unreported tax liabilities, which could lead to an overestimation of the gap.

The top-down approach utilises macroeconomic and international accounts data to cover the entire national economy (Poniatowski et al., 2020). The data sources typically include statistical offices' reports on the production of GDP or supply and use tables that provide information on the output of individual industries and the consumption by these industries and sectors of the national economy.

One of the methods based on the top-down approach involves using supply-and-use tables (SUT) and input–output tables (IOT). These tables are designed so that all production and imports match their uses (intermediate consumption, final consumption, gross capital formation, and exports). Valderas-Jaramillo et al. (2019) highlight concerns about the delay between the reference year of SUT and IOT and their publication year; official SUT and IOT are often published too late to be useful for policy-oriented research. Nonetheless, the tables are detailed enough to allow for the direct assignment of a legislated rate to each item. The Center for Social and Economic Research (CASE), a consulting firm, used an effective VAT rate in its estimates of the VAT gap for the EU (Poniatowski et al., 2020).

In contrast to the other two approaches, the approach based on econometric modelling not only provides information about the size of tax evasion but also identifies the factors and determinants that influence its magnitude. However, the econometric model can only determine the development (year-on-year changes) of tax evasion over time; its value in the base period must be established by another method (Kasnauskienė & Krimisieraitė, 2015; Schneider, 2005). The choice of the base variable is crucial as it sets the level.

The MIMIC model is another method derived from econometric modelling. This model was pioneered by Frey and Weck-Hanneman (1984), who used it to estimate the extent of the shadow economy in member states of the Organisation for Economic Cooperation and Development. Schneider et al. (2010) and Tedds (2005) also applied it to determine the proportion of the shadow economy in the GDP. Frey and Weck-Hanneman (1984) adopted this method in response to the fact that all approaches used until that time assessed the extent of the shadow economy based on a single indicator, depending on the method employed (the currency in circulation – the demand for currency approach). Furthermore, they scarcely considered any other causes of the shadow economy. The MIMIC model was also utilised to investigate the determinants of the VAT gap in Lithuania (Kasnauskienė & Krimisieraitė, 2015).

Using the MIMIC model, it is possible to incorporate variables such as EGDI or CPI in the calculation of the VAT gap; however, these variables have not yet been considered in relation to VAT gap calculations in the Czech Republic.

3. Methodology

3.1 Variable selection

This section introduces the variables for the MIMIC model. Two types of variables are required for this model: causes and indicators.

- Causes: unemployment rate, the openness of the economy, index of economic freedom, CPI, general government expenditure, final consumption expenditure, EGDI.
- Indicators: growth of GDP per capita, VAT revenue on GDP.

3.1.1 Unemployment rate

The CASE included the unemployment rate among the potential determinants of the VAT gap as an index of taxpayers' liquidity difficulties (Poniatowski et al., 2018). The unemployment rate also reflects income inequality or poverty (Reckon, 2009). In this research, the unemployment rate will be considered an index of the economic cycle.

3.1.2 Openness of economy

The openness of an economy is expressed as the proportion of the sum of imports and exports to a country's GDP. This variable was examined in relation to the VAT gap by Aizenman and Jinjarak (2008) and, more recently, by Carfora et al. (2020). According to these authors, an economy's openness positively influences the efficiency of VAT collection (i.e. it reduces the VAT gap). Additionally, research by Pluskota (2022) suggests that the share of foreign trade (exports and imports) in GDP is significant across the EU.

The openness of the economy also presents an opportunity for missing trade, intra-community fraud, and carousel fraud (Frunza, 2019). In this context, the openness of the economy has a negative impact on the efficiency of VAT collection.

3.1.3 Index of Economic Freedom

The efficiency of the tax system is strongly influenced by the quality of the government, which primarily involves the formulation and implementation of various regulations and the degree of independence of tax administration from political pressure (Godin & Hindriks, 2015). According to Chan and Ramly (2018), the redistributive effect of the VAT system also depends on the quality of the government structure; with a low-quality structure, the VAT system can be highly regressive and exacerbate income inequality.

In Czechia, Slovakia, Hungary, and Poland, substantial evidence suggests that taxpayers avoid paying taxes if they believe that the public services provided do not justify the taxes collected (Hanousek & Palda, 2004).

The index of economic freedom encompasses 12 principles for sustained progress and prosperity, many of which relate to the quality of government as described above. These principles are organised into four broad categories:

- The rule of law (property rights, government integrity, judicial effectiveness);
- Government size (government spending, tax burden, fiscal health);
- Regulatory efficiency (business freedom, labour freedom, monetary freedom);
- Open markets (trade freedom, investment freedom, financial freedom); (The Heritage Foundation, 2023).

3.1.4 Corruption perception index

The CPI is another variable incorporated into the model. It measures the perceived level of corruption in each country's public sector, according to experts and businesspeople. A higher CPI indicates lower perceived corruption. Alongside the index of economic freedom, the CPI serves as another indicator of the quality of public institutions. Reckon (2009) included CPI in its econometric analysis of the VAT gap, finding it to be the variable most strongly associated with the size of the VAT gap; lower perceived corruption corresponds to a smaller VAT gap.

The Center for Social and Economic Research also included CPI in its regression analysis of the determinants of VAT gaps, but their results showed a positive, albeit insignificant, relationship between the VAT gap and CPI; this indicates that an improvement in the perception of corruption within a country is associated with a higher VAT gap (Barbone et al., 2013). These consulting firms, CASE and Reckon, have reported contrasting results regarding the relationship between CPI and the VAT gap.

3.1.5 General government expenditure

For the model, government expenditure is expressed as a share of the GDP in each country. Reckon (2009) included government expenditure in his research into the causes of the VAT gap as it reflects the total tax burden and the size of the public sector, encompassing tax audits and other regulatory types. Similarly, Zídková and Pavel (2016) factored in government spending as a share of GDP in their study on the causes of the VAT gap, arguing that a larger public sector size leads to a reduction in the VAT gap.

3.1.6 Final consumption expenditure

Final consumption expenditure refers to the portion of expenditure primarily spent by households on goods and services intended to directly satisfy individual needs (Eurostat, 2016). According to Zídková and Pavel (2016), final consumption poses more challenges for VAT collection compared to intermediate consumption by manufacturing businesses, which can claim VAT deductions on their purchases. These purchases can be made in cash and, as such, may evade the scrutiny of tax authorities. A study by Immordino and Russo (2018) demonstrates that cashless payments negatively impact the VAT gap.

3.1.7 E-government development index

The EGDI is an index produced by the United Nations for its member states to reflect the utilisation of information technology. The index comprises three sub-indices: provision of online services (online service index), human participation (human capital index), and telecommunication

connectivity (telecommunication infrastructure index) (United Nations, 2023).

Utilising this index aids in considering the advance of digitalisation, which aims to eliminate tax evasion. CASE incorporated information technology expenditures related to GDP to gauge the effect of implementing innovative processes into tax administration; their research indicates a statistically significant negative impact of information technology expenditures on the VAT gap (Poniatowski et al., 2020).

Digitalisation will enable governments to access and analyse the necessary information, thus increasing tax collection efficiency (Alm, 2021). However, Alm (2021) also cautions that technological changes will make tax evasion increasingly difficult for entities whose transactions leave an electronic trail or are subject to third-party information reporting. These entities comprise most taxpayers in developing and developed countries. Conversely, digitalisation appears to facilitate tax evasion for multinational corporations through profit-shifting, high-income individuals via tax havens and money laundering, and independent contractors who operate without intermediaries, clearing centres, or banks, leaving no electronic trail of their transactions.

Table 1. Candidate causes and indicators of the VAT gap.

Variable	Explanation	Author	Predicted effect on the VAT gap	Source
Unemployment rate	Poverty, income inequality, index of the economic cycle	Poniatowski et al. (2018); Reckon (2009)	+	Eurostat
Gross capital formation	Investments	Kasnauskienė and Krimisieraitė (2015)	−	Eurostat
Openness of economy	Risk of carousel fraud, the openness of the economy	Zídková and Pavel (2016)	+/−	Eurostat
Index of economic freedom	Government quality, tax burden, open market	Godin and Hindriks (2015); Hanousek a Palda (2004)	−	The Heritage Foundation
CPI	Government quality, corruption perceived	Reckon (2009)	+/−	Transparency International
General government expenditure	Size of the public sector	Reckon (2009); Zídková and Pavel (2016)	+/−	Eurostat
Final consumption	Purchases of final consumers, potential VAT base	Zídková and Pavel (2016)	+	Eurostat
EGDI	Information technologies in government and tax offices	Poniatowski et al. (2020)	−	United Nations
Growth of GDP per capita	Reflection of tax evasion between taxpayers	Kasnauskienė and Krimisieraitė (2015)	+	Eurostat
VAT revenue on GDP	Level of VAT revenue in each country	Kasnauskienė and Krimisieraitė (2015)	−	Eurostat

3.1.8 Growth of GDP per capita

Gross domestic product per capita will be used as the primary indicator to reveal VAT collection inefficiencies, particularly VAT evasion. Schneider et al. (2013) utilised this variable as an indicator when examining the shadow economy, arguing that the informal economy must necessarily be reflected in the formal economy captured by statistical offices. Kasnauskienė and Krimisieraitė (2015) examined the determinants of the VAT gap using the MIMIC model, employing real GDP per capita as one of the indicators to negate the effect of inflation.

3.1.9 VAT revenue on GDP

Tax noncompliance is expected to manifest as a decrease in VAT revenue; thus, VAT revenue on GDP was selected as a second indicator (Kasnauskienė & Krimisieraitė, 2015).

Table 1 summarises each cause and indicator, providing a brief description of its inclusion in the model, the authors who have worked with these variables, the predicted effect on the VAT gap, and the source of data.

3.2 Methodology – the MIMIC model

The MIMIC model is founded on the statistical theory of a latent (unobserved) variable, which is ascertained using multiple measurable causes and indicators. Multiple causes lead to the existence of a latent variable, while simultaneously, multiple indicators of its presence can be observed (Schneider et al., 2010).

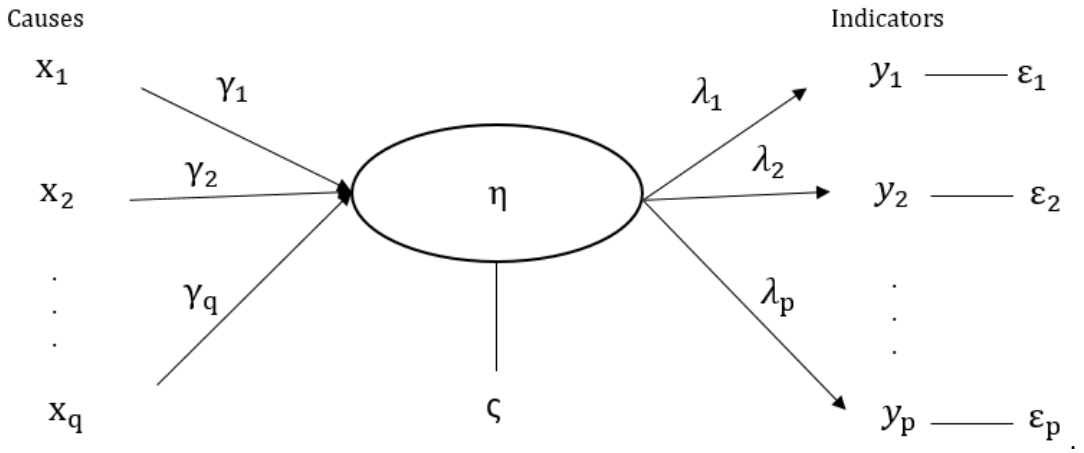

Figure 1. Mimic model – general structure. Source: Own elaboration from Schneider et al. (2010).

Figure 1 illustrates the general structure of the MIMIC model. This model is a specific type of structural equations model comprising two components: a structural model and a measurement model (Schneider et al., 2010). To estimate the variance of a latent variable, the MIMIC model utilises unstandardised estimates, meaning the first indicator is permanently fixed at level 1 and is termed the reference indicator. All other estimates vary by a specified coefficient if the reference indicator changes by 1 (Acock, 2013).

Additionally, the two models encompass several measurable (observed) variables and a latent variable, each assuming a distinct role. In the structural model, the latent variable serves as the dependent variable, influenced by the measurable variables entering the model. The equation can be expressed as

$$\eta_t = \gamma' x_t + \varsigma_t, \quad [1]$$

where x_t' is a (1*q) vector of time series x_{it}, i=1,...,q, containing potential causes of the hidden variable η_t, and γ' is a vector of coefficients describing the relationship between the hidden variable and its causes. ς_t represents the error term.

In the measurement model, the latent variable is independent, whereas the measurable variables entering the model are dependent on it.

$$y_t = \lambda \eta_t + \varepsilon_t \quad [2]$$

where y_t' is a (1*p) vector of a time series of indicators of the hidden variable, λ is a vector of regression coefficients, and ε' is a vector of white noise. In this scenario, the latent unobserved variable, the VAT gap, is initially linked to the observed indicator variables within the measurement model. Subsequently, the relationships between the latent unobserved variable and the observed explanatory variables (causes) in the structural model are examined.

By employing Equation 1 in Equation 2, a multiple regression model is derived where the explanatory endogenous variables y_{jt}, j = 1,...,p, are indicators of the latent variable η, and the explanatory exogenous variables x_{it}, i = 1,...,q, are causes. The model can be represented by the following equation:

$$y_t = \Pi x_t + z, \quad [3]$$

where $\Pi = \lambda\gamma'$ is the matrix and $z = \lambda\varsigma + \varepsilon$ represents the error term, a (p*1) vector of linear combinations of white noise ς and ε from the structural model and the measurement model.

The final MIMIC model will retain only those variables that are significant at a minimum 5% level of statistical significance, using the p-value as an indicator. The MIMIC index is computed using the structural model equation, Equation 1. Equation 4 outlines this calculation; it is a modified version of Equation 1.

$$\tilde{\eta}_t = \gamma x_{1t} + \gamma x_{2t} .. + \gamma x_{qt}, \quad [4]$$

where x_{1t} to x_{qt} represent the variables of causes at a level of at least 5%. The MIMIC index merely indicates the relative development of the latent variable, the VAT gap in this instance. To translate the relative values into absolute terms, a baseline variable obtained by another method must be employed.

For conversion, the following equation is used:

$$\hat{\eta}_{\mathrm{t}} = \frac{\tilde{\eta}_t}{\tilde{\eta}_{base}} * \dot{\eta}_{base}, \quad [5]$$

where $\tilde{\eta}_t$ denotes the value of the MIMIC index at time t according to Equation 4, $\tilde{\eta}_{base}$ is the value of the MIMIC index in the base period, and $\dot{\eta}_{base}$ is an estimate of the latent variable obtained by another method.

3.3 Data

The MIMIC model is constructed using panel data from selected European countries (a total of 26 countries) spanning the years 2002 to 2020. Due to the availability of data, the maximum time series was selected to provide an adequate database for estimating the MIMIC model. Concerning the availability of the EGDI, which represents digitalisation, the time series could not commence earlier than 2002, with the objective of capturing an overview of VAT gap development up to the most recent year feasible. Regarding other variables, the list of selected countries is presented in Appendix 1. The group also includes non-European countries, as the study does not consider EU membership.

Table 2 contains descriptive statistics. The total dataset comprises 26 panels and 20 time periods (years), with 520 observations collected for each variable. For the MIMIC model, the data must be stationary. The data were tested for the presence of a unit root using the Levin–Lin–Chu test, designed for panel data. This test is suitable for data where the number of panels does not exceed 100 and the ratio of the number of panels to the number of time periods is close to zero (Levin et al., 2002; STATA, 2023). The result of the test confirmed the presence of a unit root, indicating that the data are not stationary and require differencing.

Table 2. Descriptive statistic*s*

Variable	Number of observations	Mean	Standard deviation	Minimum	Maximum
Unemployment rate	520	8.03	4.32	2.00	27.50
Openness of economy	520	59.87	30.42	22.80	176.70
Economic freedom index	520	69.47	6.25	48.70	82.60
CPI	520	0.65	0.16	0.26	0.95
General government expenditure	520	0.44	0.07	0.24	0.65
Final consumption	520	0.53	0.09	0.23	0.70
EGDI	520	0.73	0.11	0.47	0.98
Growth of GDP per capita	520	1.88	3.95	−14.50	23.20
VAT revenue on GDP	520	20.19	4.07	11.20	30.50

Particularly noteworthy is the maximum unemployment rate of 27.5% recorded in Greece in 2013. This is an exceptional case, as the standard deviation is approximately 4.62. The data for GDP per capita growth present another extreme value. The minimum for this variable is -14.5%, recorded in Estonia in 2009, while the maximum represents a year-on-year increase of 23.2%, occurring in Ireland in 2015.

In terms of data variability, the standard deviation provides the most informative measure. If the standard deviation is around zero, the data exhibit low variability despite the presence of outliers. Data on the openness of the economy show the highest standard deviation. Luxembourg consistently records very high values of economic openness, often exceeding 100%. In contrast, openness is low in Spain, Greece, and Sweden.

4. Results

4.1 MIMIC model

Table 3 presents the MIMIC models, detailing both parts: the structural model and the measurement model. In Model 1, all variables are included with the goal of retaining only those that are statistically significant in the structural model at a minimum 5% level of significance. A stepwise selection process is employed to systematically eliminate statistically insignificant variables from the structural model; this process is fully detailed in Table 3. In the measurement model, both variables – **growth of GDP per capi**ta and **VAT revenue on GDP** – are statistically significant at the 1% level, allowing them to remain in all subsequent models.

The variable **index of economic freedom** was the first to be removed from the structural model due to its high p-value. The structural model of Model 2 consists of six variables. The **unemployment rate** also had to be removed due to its high p-value, resulting in Model 3, the final model, which consists of only four variables.

Information criteria are useful when determining which model is optimal. The Akaike information criterion (AIC) was applied in this research. The most appropriate model is identified by the lowest value of this criterion, which is the final model with the lowest AIC. Other indicators of model quality are also optimal in the last model. The comparative fit index (CFI) measures how close a given model is to a perfect fit with the data used. It ranges from 0 to 1; a higher value indicates a better model. The final model, with a CFI of 0.86, shows the highest value.

The interpretation of coefficients in a structural model of the MIMIC model closely resembles that in regression analysis. Their value indicates the resultant change in the VAT gap for a unit change in the causal variable under the ceteris paribus condition. In the following paragraphs, the coefficients from the structural model are interpreted according to Model 1, which includes all variables, significant and insignificant. Final consumption emerges as the most crucial driver of the VAT gap. According to Model 1, if **final consumption** increases by 1 %, then the VAT gap also increases by about 5.71 %. **Final consumption** is significant at the 5% level; hence, it is retained in

Model 2 and the final model, with the value of its coefficient fluctuating around 6 % to 7 %. A higher **unemployment rate** increases the VAT gap; according to Model 1, if the **unemployment rate** increases by 1 %, the VAT gap increases by 0.01 %. In Model 2, the coefficient of the unemployment rate remains stable. However, this variable was removed from the final model due to an excessively high p-value. Greater **openness of the economy** also raises the VAT gap; if the **openness of the economy** increases by 1 %, then the VAT gap increases by 0.02 %. **The openness of the economy** is significant at the required level, with the value of its coefficient remaining stable in Model 2 and the final model. There is no unified conclusion about the effect of the openness of the economy on the VAT gap; Aizenman and Jinjarak (2008) suggest a negative impact, while Frunza (2019) argues that an open economy provides opportunities for carousel frauds and other VAT frauds, which have a positive effect on the VAT gap. This research supports the latter view regarding the positive impact of economic openness on the VAT gap.

Table 3. Model MIMIC for Europe-23 (author calculations using STATA) – VAT revenue on GDP is used as a reference indicator.

	Model 1	**Model 2**	**Model 3** ***(final model)***
Structural	Coefficient (P-value)	Coefficient (P-value)	Coefficient (P-value)
Unemployment rate	0.01 (0.31)	0.01 (0.26)	
Openness of economy	**0.02 (0.00) *****	**0.02 (0.00) *****	**0.02 (0.00) *****
Index of economic freedom	−0.01 (0.36)		
CPI	**−0.27 (0.00) *****	**−0.32 (0.00) *****	**−0.14 (0.00) *****
General government expenditure	**−4,44 (0.00) *****	**−5.3 (0.00) *****	**−4.53 (0.00) *****
Final consumption	**5.71 (0.00) *****	**7.08 (0.00) *****	**6.05 (0.00) *****
EGDI	**−2.01 (0.00) *****	**−2.44 (0.00) *****	**−2.20 (0.00) *****
Measurement			
VAT revenue on GDP	1	1	1
Growth of GDP per capita	13.01 (0.00) ***	10.89 (0.00) ***	12.97 (0.00) ***
Statistics			
Chi-square	50.9 (0.00) ***	60.30 (0.00) ***	42.83 (0.00) ***
Degrees of Freedom	6	5	4
CFI	0.851	0.80	0.86
SRMR	0.042	0.056	0.05
RMSEA	0.12	0.15	0.140
AIC	850.57	−838.84	−2517.93

The index of economic freedom has a negative effect on the VAT gap; if **the index of economic freedom** increases by one percentage point, then the VAT gap decreases by 0.01 %. This variable is not statistically significant at the required level. Due to an excessively high p-value, it was removed from Model 2 and the final model.

The **CPI** also has a negative effect on the VAT gap. If the **CPI** increases by one percentage

point – indicating a lower perception of corruption – the VAT gap decreases by 0.27 %. **The CPI** is significant at the required level of significance, so it remains in Model 2 and the final model. The value of its coefficient remains stable. The CASE (Barbone et al., 2013) and Reckon (2009) examined the influence of CPI on the VAT gap and reached opposite conclusions. This research supports the findings of Reckon (2009) regarding the negative effect of CPI on the VAT gap.

General government expenditure also negatively impacts the VAT gap, but this variable is more substantial. If the share of **general government expenditure** in GDP increases by 1 %, then the VAT gap decreases by 4.44 %. This variable is significant at the required level of significance, so it remains in Model 2 and the final model, with the value of its coefficients fluctuating between 4 and 5.5 %. Similar to the case of **economic openness**, there is no unified conclusion about the effect of general government expenditure on the VAT gap; however, this research supports the findings of Zídková and Pavel (2016) regarding the negative impact of **general government expenditure** on the VAT gap.

The **EGDI** also decreases the VAT gap, although it is not as strong a variable. If the **EGDI** increases by one percentage point, then the VAT gap decreases by 2.01 %. **EGDI** is significant at the required level, so it remains in Model 2 and the final model. The value of its coefficient remains stable.

Table 4. Robustness test: MIMIC model for EU-17.

	Model MIMIC for the EU-17
Structural	Coefficient (P-value)
Unemployment rate	0.01 (0.12)
Openness of economy	**0.01 (0.00) *****
Index of economic freedom	−0.01 (0.38)
CPI	**0.75 (0.00) *****
General government expenditure	**−1.71 (0.00) *****
Final consumption	**6.05 (0.00) *****
EGDI	**−0.86 (0.00) *****
Measurement	
VAT revenue on GDP	1
Growth of GDP per capita	15.71 (0.00) ***
Statistics	
Chi-square	20.65 (0.00) ***
Degrees of Freedom	6
CFI	0.85
SRMR	0.04
RMSEA	0.09
AIC	−2517.93

4.2 Test for the robustness of coefficients

As outlined in the methodology section, the dataset exhibits high variability, with the openness of the economy introducing the most variation. It is thus prudent to test the robustness of the coefficients of the MIMIC model. To this end, another MIMIC model was estimated, including only developed EU countries, while nine transit countries were excluded from the dataset. Table 4 presents the results of the MIMIC model for Europe-17. Despite the dataset's limitations, the coefficients of the variables remain consistent, with the same signs and levels of significance.

4.3 Application of the MIMIC model in the Czech Republic

In this subsection, the MIMIC model is applied to the Czech Republic. In Equation 4, each coefficient of a significant variable is multiplied by the corresponding variable (for example, the coefficient of the **openness of the economy** is multiplied by the level of openness of the economy in the Czech Republic), producing the MIMIC index. The MIMIC index expresses the relative size of the VAT gap. To convert this into absolute values, it is necessary to use a base variable obtained by another method. For this research, an estimate by CASE (Barbone et al., 2013) is used as the baseline variable, setting the VAT gap in the Czech Republic in the baseline year of 2002 at 29% of VTTL. Values of the VAT gap for the period 2003–2020 are determined according to the MIMIC index, influenced by significant causes of the VAT gap.

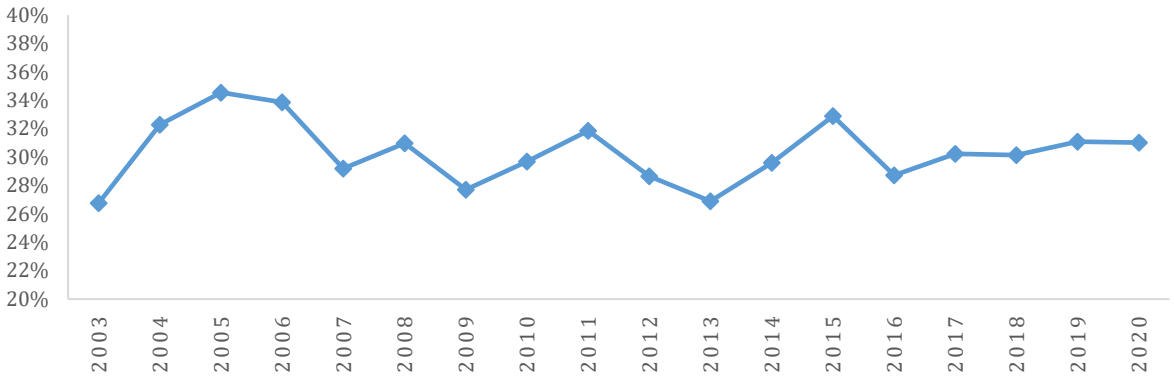

Figure 2. The size of the VAT gap in the Czech Republic using the MIMIC model. Source: author's calculations

Figure 2 displays estimates of the VAT gap in the Czech Republic from 2003 to 2020, expressed as a percentage of VTTL. It peaked in 2005 at almost 35%. The second highest point was in 2015, at 33 %. Conversely, the lowest estimated values were around 27 % and occurred in 2003 and 2013, when the VAT gap fell to a minimum of 27 %. The VAT gap level has been rising steeply since 2013, reaching one of its highest values in 2015 and then decreasing again in 2016. Since 2017, the VAT gap has stabilised at around 31 % of VTTL.

5. Conclusions

The paper provides a different insight into the VAT gap through econometric modelling. The MIMIC model does not process statistical data from National Accounts; instead, it evaluates the relationship between inputted variables and the VAT gap. The contribution of the MIMIC model lies in the evaluation and identification of significant variables that can proxy potential sources of VAT collection inefficiencies. Such information allows for the development of policy-related recommendations. However, this contribution also conceals a significant limitation of the research conducted: the omission of a vital input variable. This omission could lead to biased results from the MIMIC model. To avoid such an omission, a literature search was conducted on studies concerning the causes or determinants of the VAT gap. The search included studies by both foreign and domestic authors, as well as those under the auspices of the EU, such as CASE or Reckon.

Significant causes identified include final consumption, general government expenditure, the openness of the economy, the CPI, and the EGDI. In contrast, the unemployment rate and the index of economic freedom are not statistically significant variables. In the MIMIC model, the unemployment rate represents the economic cycle; according to this analysis, the economic cycle does not significantly influence the extent of the VAT gap. The index of economic freedom and CPI were included in the model to express the quality of government, among other factors. A key difference between these two variables is that the CPI is based on the perceptions of businesspeople and experts about corruption in their home country, while the index of economic freedom is based on data from the World Bank. This underscores the importance of a country not only maintaining high-quality governance but also effectively conveying this to its citizens and taxpayers.

Recommendations for further development of indirect tax policy should be based on the statistically significant causes of VAT evasion. A key recommendation from this research is to focus on the digitalisation of tax offices, which can increase the efficiency of VAT collection. Digitalisation can also facilitate cooperation between tax offices at home and abroad, potentially mitigating the impact of economic openness.

Acknowledgements

This research was supported by the Internal Grant Agency of FBE, Mendel University in Brno, registration no.: IGA-PEF-DP-23-012.

References

Acock, A. C. (2013). *Discovering structural equation modeling using Stata* (Vol. 1). Stata Press.

Aizenman, J., & Jinjarak, Y. (2008). The collection efficiency of the value added tax: Theory and international evidence. *Journal of International Trade and Economic Development, 17*(3), 391-410. https://doi.org/10.1080/09638190802137059

Alm, J. (2012). Measuring, explaining, and controlling tax evasion: Lessons from theory, experiments, and field studies. *International Tax and Public Finance, 19*(1), 54-77. https://doi.org/10.1007/s10797-011-9171-2

Alm, J. (2021). Tax evasion, technology, and inequality. *Economics of Governance, 22*(4), 321-343. https://doi.org/10.1007/s10101-021-00247-w

Argentiero, A., Casal, S., Mittone, L., & Morreale, A. (2021). Tax evasion and inequality: Some theoretical and empirical insights. *Economics of Governance, 22*(4), 309-320. https://doi.org/10.1007/s10101-021-00261-y

Baert, P. (2023). *VAT in the digital age*. European Parliamentary Research Service. Retrieved from https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739371/EPRS_BRI(2023)739371_EN.pdf

Barbone, L., Belkindas, M., Bettendorf, L., Bird, R. M., Bonch-Osmolovskiy, M., & Smart, M. (2013). *Study to quantify and analyse the VAT gap in the EU-27 member states*. CASE Network Reports, No. 116, Center for Social and Economic Research (CASE). https://doi.org/10.2139/ssrn.2365189

Carfora, A., Dongiovanni, S., Marabucci, A., & Pisani, S. (2020). The impact of domestic factors and spillover effects on EU countries VAT gap. *Argomenti di Discussione, 1*, 1-35.

Cnossen, S. (2022). The C-inefficiency of the EU-VAT and what can be done about it. *International Tax and Public Finance, 29*(1), 215-236. https://doi.org/10.1007/s10797-021-09683-0

Eurostat. (2016). Glossary: Final consumption expenditure. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Final_consumption_expenditure

Eurostat. (2024). Main national tax aggregates. Retrieved from https://ec.europa.eu/eurostat/web/products-datasets/-/gov_10a_taxag

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

Frunza, M. C. (2019). *Value added tax fraud*. Routledge. https://doi.org/10.4324/9781315098722

Gajewski, D. J., & Jonski, K. (2022). 'VAT Gap' estimation: Distinguishing between informality and fraud. *EC Tax Review, 31*(3), 124-130. https://doi.org/10.54648/ECTA2022012

Godin, M., & Hindriks, J. (2015). A review of critical issues on tax design and tax administration in a global economy and developing countries. *Core Discussion Paper, 2015*(28), 2-76.

Hanousek, J., & Palda, F. (2004). Quality of government services and the civic duty to pay taxes in the Czech and Slovak Republics, and other transition countries. *Kyklos, 57*(2), 237-252. https://doi.org/10.1111/j.0023-5962.2004.00252.x

Chan, S. G., & Ramly, Z. (2018). The role of country governance on value-added tax and inequality. *E & M Ekonomie a Management, 21*(4), 79-93. https://doi.org/10.15240/tul/001/2018-4-006

Immordino, G., & Russo, F. F. (2018). Cashless payments and tax evasion. *European Journal of Political Economy, 55*, 36-43. https://doi.org/10.1016/j.ejpoleco.2017.11.001

Kasnauskienė, G., & Krimisieraitė, J. (2015). Using MIMIC models to examine determinants of VAT gap in Lithuania. *Organisations and Markets in Emerging Economies, 6*(1), 107-126. https://doi.org/10.15388/omee.2015.6.1.14230

Levin, A., Lin, C. F., & Chu, C. S. J. (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics, 108*(1), 1-24. https://doi.org/10.1016/S0304-4076(01)00098-7

Moravec, L., Hinke, J., & Borsiczká, M. (2021). VAT gap estimation and influence of selected methods of financial administration on its reduction. *DANUBE, 12*(1), 45-60. https://doi.org/10.2478/danb-2021-0004

Novysedlák, V., & Palkovičová, J. (2012). *The estimate of the value added tax revenue loss*. Institute for Financial Policy, The Ministry of Finance of the Slovak Republic. Retrieved from https://www.mfsr.sk/en/finance/institute-financial-policy/working-papers/the-estimate-value-added-tax-revenue-loss/

Pluskota, A. (2022). VAT gap determinants in the European Union and Poland. *Ekonomia i Prawo-Economics and Law, 21*(3), 623-633. https://doi.org/10.12775/EiP.2022.033

Poniatowski, G., Bonch-Osmolovskiy, M., Duran-Cabré, J. M., Esteller-Moré, A., & Śmietanka, A. (2018). *Study and reports on the VAT gap in the EU-28 member states: 2018 final report*. CASE Reports, No. 496, Center for Social and Economic Research (CASE). https://doi.org/10.2139/ssrn.3272816

Poniatowski, G., Śmietanka, A., & Bonch-Osmolovskiy, M. (2020). *Study and reports on the VAT gap in the EU-28 member states: 2020 final report*. CASE Reports, No. 503, Center for Social and Economic Research (CASE). https://doi.org/10.2139/ssrn.3744157

Reckon. (2009). *Study to quantify and analyse the VAT gap in the EU-25 member states*. Report for DG Taxation and Customs Union.

Schneider, F. (2005). Shadow economies around the world: What do we really know? *European Journal of Political Economy, 21*(3), 598-642. https://doi.org/10.1016/j.ejpoleco.2004.10.002

Schneider, F., Buehn, A., & Montenegro, C. E. (2010). Shadow economies all over the world: New estimates for 162 countries from 1999 to 2007. *World Bank Policy Research Working Paper (5356)*, 3-52. https://doi.org/10.4337/9780857930880.00007

STATA. (2023). *Panel-data unit-root tests*. Retrieved from https://www.stata.com/features/overview/panel-data-unit-root-tests/

Tedds, L. M. (2005). Revisiting the Latent Variable/MIMIC model approach to measuring the size of the Canadian underground economy. In Schneider, F., & Bajada, C. (Eds.), *Size, causes and consequences of the underground economy* (pp. 159-176). Ashgate Publishing.

The Czech Fiscal Council. (2022). *Development of the national debt of the Czech Republic*. Retrieved from https://www.rozpoctovarada.cz/publikace/informacni-studie-vyvoj-statniho-dluhu-ceske-republiky/

The Heritage Foundation. (2023). *2023 Index of economic freedom: About the index*. Retrieved from https://www.heritage.org/index/about

United Nations. (2023). *E-Government development index (EGDI)*. Retrieved from https://publicadministration.un.org/egovkb/en-us/About/Overview/-E-Government-Development-Index

Valderas-Jaramillo, J. M., Rueda-Cantuche, J. M., Olmedo, E., & Beutel, J. (2019). Projecting supply and use tables: New variants and fair comparisons. *Economic Systems Research, 31*(3), 423-444. https://doi.org/10.1080/09535314.2018.1545221

Zídková, H., & Pavel, J. (2016). What causes the VAT gap? *Ekonomicky Casopis, 64*(9), 811-826.

Appendix 1

Selected countries:

Austria
Belgium
Cyprus
Czech Republic
Denmark
Estonia
France
Germany
Greece
Hungary
Ireland
Italy
Latvia
Lithuania
Luxembourg
Malta
Netherlands
Poland
Portugal
Romania
Slovenia
Spain
Slovakia
Sweden
Non-EU countries:
Norway
Switzerland