Size and Causes of the Shadow Economy in Iran: An Analysis with the MIMIC Approach

Soroush Hashemi¹, Mohammad Mahdi Lotfi Heravi², Friedrich Schneider³

ARTICLE INFO

Article history:
Date of submission: 01-06-2022
Date of acceptance: 01-10-2022

JEL Classification:
C39
O17
O53
J69

Keywords:
Shadow Economy
Undocumented workers
Iran
MIMIC

ABSTRACT

This study aims to estimate the size of the Iranian Shadow Economy (SE) using the multiple indicators multiple causes (MIMIC) approach, a variant of the structural equation modeling (SEM) approach. The application of the MIMIC approach allows for the consideration of established drivers of the SE as well as potential causes that might be of particular importance to the Iranian SE, namely inflation and the size of the construction sector in the economy, the latter of which is considered in this study for the first time. The significant determinants of the Iranian SE are found to be inflation, unemployment, trade openness, and the size of the construction sector. The SE is also found to be positively related to cash usage and negatively related to the labor force participation rate and the Iranian GDP. Using annual data for the 1979-2019 period, our empirical results show that after an initial surge in SE size during the Iran-Iraq war (1980-1988), the Iranian SE generally decreased, although fluctuations were also present. The average SE size during the entire studied period of 1979-2019 is found to be 31.83% of GDP.

1. Introduction

The definition, size, and key characteristics of the shadow economy (SE), as well as informal or undeclared employment, have been sources of intensive debates in both policy and academic circles. Recent global developments, such as the worldwide pandemic, migration waves, climate

¹. Department of Management, Science and Technology, Amirkabir University of Technology, Iran, Email: s.hashemi@aut.ac.ir
². Department of Management, Science and Technology, Amirkabir University of Technology, Iran. (Corresponding author), Email: Mahdi.lotfi@aut.ac.ir
³. Research Institute of Finance and Banking, Johannes Kepler University, A-4040 Linz, Austria. Email: friedrich.schneider@jku.at
change, and trade tensions, have triggered a renewed interest in the fields surrounding the SE and informal or undeclared employment. In particular, recent research has highlighted the importance of the role that the SE plays in various societal, economic, and environmental issues. For instance, the SE has been found to promote income inequality (Berdiev & Saunoris, 2019), have a complementary relationship with corruption (Buehn & Schneider, 2012b), increase financial instability and public indebtedness (Elgin & Uras, 2013), hinder entrepreneurial entry (Estrin & Mickiewicz, 2012), negatively influence economic and sustainable development (Hoinaru et al., 2020), and cause pollution (Biswas et al., 2012; Pang et al., 2021). Furthermore, existence of a thriving SE will result in official economic statistics becoming increasingly unreliable. Policies based on unreliable and erroneous data might result in ineffective policies (Enste & Schneider, 2000). The existence of the SE can potentially have major implications for budgetary policies. For example, if the burden of taxation is a prominent cause of the SE in an economy, then raising the tax rates by the authorities will result in even less revenues for the government (Giles, 1999). Therefore, in order for effective policies to be put in place, the policymaker requires a sound understanding of the causes, size, and trends of the SE.

In light of these, our paper aims to provide new estimations for the size of the Iranian SE between 1979 and 2019 using SE determinants that are commonly used within the literature as well as variables that have the potential to be of importance in the specific context of the Iranian SE such as size of the construction sector. To the authors’ best knowledge, no previous research has considered the size of the construction sector as a potential driver of the Iranian SE before us. Estimations are done by applying the MIMIC approach.

The rest of this paper is organized in the following order: section 2 provides a discussion on the definition of the SE, section 3 presents the literature review, section 4 discusses the causes and indicators of the SE and

---

1. Compare here Medina and Schneider (2019, 2021), where the global aspects are handled in much more detail.
the MIMIC approach, section 5 details our strategy and presents and discusses our results, and section 6 provides the concluding remarks and some policy recommendations to reduce the Iranian SE size.

2. Literature Review

Authors are often faced with the difficulty of how to define the SE when attempting to measure it. In economics, separate definitions have been put forward and used by different researchers\(^1\). For instance, one succinct definition comes from Smith (1994), who defines the SE as “Market-based production of goods and services, whether legal or illegal, that escapes detection in the official estimates of GDP”. In this paper, however, we use Medina & Schneider’s (2019, 2021) definition who define the SE as follows:

“All economic activities which are hidden from official authorities for monetary, regulatory, and institutional reasons. Monetary reasons include avoiding paying taxes and all social security contributions, regulatory reasons include avoiding governmental bureaucracy or the burden of regulatory framework, while institutional reasons include corruption law, the quality of political institutions and weak rule of law.” The SE, based on the aforementioned definition, therefore, mostly includes productive and legal economic activities that would contribute to the GDP had they been recorded, rather than criminal, do-it-yourself, or household activities.

The first empirical contribution to the Iranian SE literature comes from research done by Khalatbari (1990), in which she provides definitions and estimations for the SE. Using the Gutmann (1977) model, she estimates the size of the Iranian SE in 1986 to be approximately 20% of GDP after applying some adjustments. Bagheri Garmarudy (1998) uses the cash demand approach to estimate the size of the Iranian SE between 1971 and 1995. The findings indicate that the Iranian SE was, on average, 23% of the GDP in the studied period. Esfandiari & Mehrabani (2007) find that Iranian

\(^1\) See Fleming et al. (2000), Thomas (2001) and Dell’Anno (2021) for more detailed discussions on the definition of the SE.
SE had a generally upward trend with an average of 19.64% of the GDP over the period 1996-2003 with the maximum SE size being in 1999 with 32.07% of the GDP.

Arab Mazar Yazdi (2001) is the first to employ the MIMIC approach in order to estimate size of the Iranian SE individually for a three-decade period between 1968 and 1998. Arab Mazar Yazdi’s estimates show that the Iranian SE continued to mostly fluctuate around 8% of the GDP between 1968 and 1988, but experienced a rapid increase during the last studied decade and reached approximately 23% of the GDP in 1998. One important and perhaps surprising outcome of his research is that although tax burdens are usually considered to be the principal factor affecting the SE size, they appear to play a substantially less important role in the Iranian SE dynamics compared to other causes. Sameti et al. (2010) apply the MIMIC approach in order to estimate the size of the Iranian SE for over forty years between 1965 and 2005. In their chosen MIMIC model, trade openness, government expenditures, and unemployment rate are the three factors that are found to be significantly affecting the SE size. In addition, their empirical results show a negative relationship between the Iranian SE and the GDP growth. The MIMIC approach results show that the Iranian SE increased over the studied period, from 6.2% of the GDP in 1965 to 26.1% of the GDP in 2005. Using the MIMIC approach for the time period between 1974 and 2013, Pirae & Rajaee (2015) find that unemployment rate, inflation, tax burden, and trade openness are the significant drivers of the Iranian SE and that the size of the Iranian SE rose continuously, from 7% of the GDP in 1974 to 38.5% of the GDP in 2013. Shahabadi et al. (2020) estimate SE sizes for a number of developing countries with natural resources abundance between 2004 and 2015 by using the MIMIC approach. They find that unemployment rate has the highest positive impact on the SE while economy openness has the most negative impact on it. They also find that Iranian SE size averaged around 22% of GDP during this period and that the SE sizes in selected countries were generally increasing. Similarly, using the CDA approach and
data from Iranian provinces, Karbor et al. (2019) find that during the period 2001-2015 the SE size increased in every Iranian province. To our knowledge, other than Medina & Schneider (2019) no other research has provided estimations for the Iranian SE size for the years following 2015.

Schneider & Savaşan (2007) use the MIMIC method to estimate the size of the SE of Turkey and her neighboring countries\(^1\) for the 1999-2005 period. Over this period, Iran’s SE is estimated to be on average 19.81% of the GDP, which is the lowest of any studied country. Schneider (2017) estimates the SE of 143 countries from 1996 to 2014, and results reveal that the average Iranian amounted to 18.2% of the GDP in the studied period. Medina & Schneider (2017), who estimate the size of SE for 158 countries across the world between 1991 and 2015, find that the Iranian SE size had an average of 17.9% of the GDP in this period. Estimations of Medina & Schneider (2019) indicate that the Iranian SE was 14.3% of the GDP in 1991 and that it underwent a short expansion to reach 20.5% of the GDP in 1995 but then declined, constituting only 15.9% of the GDP in 2017. According to Medina & Schneider (2019), the average Iranian SE size between 1991 and 2017 was 17.1% of the GDP.

3. The MIMIC Approach and Data

Structural Equation Modeling (SEM) is a technique used to determine the statistical relationship between both the observed and unobserved (latent) variables. The main advantage of this technique as stated by Venturini & Mehmetoglu (2017) is that unlike the “single equation techniques” (such as linear regression and logistic regression), SEMs allow for the simultaneous use of both more than one independent variable and more than one dependent variable. The procedure consists of a comparison of the model-implied covariance matrix and the empirical or “databased” covariance matrix. After estimation, if the two matrices are consistent with each other,

---

1. The group of countries considered in this paper are: Armenia, Bulgaria, Georgia, Greece, Iran, Lebanon, Syria and Turkey.
then we can consider the model to be a likely explanation for the examined variables (Dell’Anno, 2007). The MIMIC (Multiple Indicators Multiple Causes) model is a special case of SEM which links a set of observed variables (causes) to a latent construct (the SE) on one side and the latent variable (the SE) itself to another set of observed variables (indicators) on the other side. Formally, the MIMIC model comprises two parts: the structural model and the measurement model (Hassan & Schneider, 2016b). The structural model (Eq.1) explains the relationship between the causal variables (causes of the SE) and the latent variable (the SE):

\[ \eta = \gamma'X + \xi \]  

(1)

Where \( \eta \) is the latent variable (Index of SE), \( X \) is the vector of the causes, \( \gamma \) is the vector of structural coefficients, and \( \xi \) is the vector of error terms.

The second part the MIMIC model, called the measurement model (Eq.2), is specified by the following equation and connects the latent variable (the SE) to the indicator variables (which track the movements of the SE):

\[ Y = \lambda\eta + \epsilon \]  

(2)

In this equation \( Y \), is the vector of indicator variables, \( \lambda \) is the vector of estimated regression coefficients, and \( \epsilon \) is the vector of error terms. However, it is not possible to obtain unique solutions to \( \lambda \) and \( \gamma \) by running an estimation for the aforementioned model. The solution is to choose an indicator, called the “variable of scale” or alternatively the “reference indicator”, and to assign a constant value to it (either +1 or -1 depending on the variable of scale) in the \( \lambda \) matrix. This is both sufficient for the model to have a unique solution and convenient for economic interpretation (Dell’Anno & Solomon, 2008). Figure 1 shows a general MIMIC model.

---

1. More details regarding applying the MIMIC approach are present in Dell’Anno (2003) and Dell’Anno (2007).
After designating $\Sigma(\theta)$ as the covariance matrix of the MIMIC model, the main purpose of an SEM estimation becomes obtaining parameters and covariances that produce the closest estimation of $\Sigma(\theta)$ to the sample covariance matrix of the observed causal and indicator variables. The method that computes how close to each other the covariance matrices are, is called the fitting function (Schneider et al., 2010). Although a number of different options are available\(^1\), the Maximum Likelihood (ML) method has been most commonly used. Once the estimation of the MIMIC model is complete, the estimated coefficients of the variables are then used to approximate the size of the SE. However, the MIMIC estimation gives only an index of the size of the SE. In a process called “benchmarking” or “calibration”, this relative index is converted into absolute values for the size of the SE using a previously-known SE size belonging to a particular year. This exogenous value usually comes from either the author’s own estimation for the SE using another approach or from another research altogether.

As with every other method of estimating the SE, the MIMIC approach has been subject to criticism, and there have been concerns and discussions about the accuracy of the estimation produced by this approach\(^2\). Several of

---

\(^1\) See Hayashi et al. (2011) for a detailed technical explanation for different estimation methods.

\(^2\) See Breusch (2005) and Dell’Anno & Schneider (2006).
the shortcomings of the MIMIC approach are laid out in Dell’Anno et al. (2007) and Williams & Schneider (2016, p. 33). Among others, criticism include the lack of confidence that the latent variable is in fact the SE and not another similar concept, the instability of the results with respect to changes in the model specifications, the benchmarking procedure, and the reliance of the estimations on an exogenous value. However, in the absence of better alternatives, the MIMIC approach remains a widely used and fruitful tool in estimation of the shadow economy of different countries.

The data used in the Iranian SE estimation in this paper are annual time series that cover the 1979-2019 period. Table A.1 describes the variables, the sources, and the data transformations. Figure 2 contains information on selected Iranian economic variables. The GDP per capita (Figure 2a) undergoes a decline starting at 1979, gradually climbing to just shy of its initial value towards the end of the period. Another dip in its value occurred during the years 1983-1988 caused by the Iran-Iraq war. In this period, a large part of the fluctuating Iranian GDP can be attributed to factors such as over-reliance on oil and the economic sanctions levied upon the nation. The total tax burden (Figure 2b), albeit going through turbulent changes during 1995-2004, stays relatively the same. The year 1999 saw the highest value of the total tax burden with around 8.5%. The unemployment rate (Figure 2c) goes through major dips and rises over the years. The value of this variable started at 10% in 1979, before soaring to an all-time high of 15% in 1986. Afterwards, a precipitous downward trend started until the year 1996, marking the least unemployment rate during the studied period. In the years following, the values experience constant dips and rises, while maintaining an average of ~13%. Government expenditure as a percentage of GDP (Figure 2d) almost consistently decreased from its starting value of 20% to 11% in 2019, while trade openness (Figure 2e) saw a major drop in 1986 and two less significant ones in 1998 and 2015, but generally increased in value from 42% in 1979 to 62% in 2019. Construction sector value added % GDP (Figure 2f) peaked in 1983 with ~12%, before dropping to 6.5% in 1990. The trend then plateaued for a while, then fluctuating between 6% and 9% during the years after. The value stabilized to some extent after the year
2016. The inflation rate (Figure 2g), however, is the most erratic variable, undergoing major fluctuations throughout all the years studied. The peak of this figure happened in 1995 with ~49% and bottoming out at 7% in 2016.

Figure 2. Selected Iranian economic variables (Data sources: The World Bank, Central Bank of Iran, Statistical Centre of Iran, Plan and Budget Organization of the Islamic Republic of Iran)
Based on previous research on the shadow economies of Iran and other countries, this paper considers the following causes and indicators in the estimation of the Iranian SE:

3.1. Causes of the SE

*Tax Burdens:* Tax burdens might prompt an individual or business to move from the official economy to the shadow one. The greater the difference between income and after-tax income, the more tempted individuals and businesses will be to evade these taxes and enter the SE. Hence, higher tax rates and burdens should theoretically result in bigger underground productions and, consequently, a larger SE. Several studies have confirmed this positive connection empirically (e.g. Almenar et al., 2020; Hassan & Schneider, 2016b; Tan et al., 2017).

Therefore, we hypothesize a positive and significant relationship between taxes and the SE in our model, ceteris paribus.

*Unemployment Rate:* As unemployment rises in a country, more people will turn towards working in the SE, thus increasing its size. This positive relationship has been found to be significant in different studies (e.g. Buehn & Schneider, 2012a; Dell’Anno et al., 2007; Dell’Anno & Davidescu, 2019). Mauleón & Sardà (2016) find the link between unemployment rate and the SE to be statistically significant in countries with high unemployment rates, such as Spain and Greece, but they do not find this to be the case in countries with low unemployment rates, such as Germany and Italy. Based on the established work on how unemployment and the SE are related, we hypothesize a positive impact of unemployment on the SE in our model, ceteris paribus.

*Government Size:* Increasing involvement of the government in the economy can be considered as a contributing factor to increases in the size of SE since more state involvement usually equals more regulation, more bribery and corruption, and distortion in the allocation of resources between private and public businesses (Dell’Anno, 2007). Empirical works have
shown the positive relationship between government’s involvement and SE size for different countries (e.g. Dell’Anno et al., 2018; Medina & Schneider, 2017; Sharifkarimi et al., 2017). Government expenditures to GDP ratio has been commonly used as a proxy variable for government’s involvement in the economy in literature. We hypothesize a positive relationship between government size and the SE in the model, ceteris paribus.

Trade Openness: Trade openness has been defined as the sum of export and import measured as a percentage of the GDP. According to Medina & Schneider (2017), with the growth of economies, the relocation of economic activities from the official economy to the unofficial one would likely be more difficult. They also add that it would be more challenging to conceal trade from the regulators in the presence of increasing international trade. Among others, studies by Schneider et al. (2010) and Medina et al. (2019) have shown this negative link empirically. As such, other things equal, we expect an increasing degree of trade openness to negatively impact SE in our model.

Size of the Construction Sector: In addition to Iranian workers who engage in informal activities, there are a large number of foreign workers who are forced to work in the shadow economy since they don’t have the legal permits to work in the official economy. This is particularly the case for Afghan immigrants whom Iran has been hosting for decades. Latest official reports from the Statistical Centre of Iran (SCI) put the number of foreign nationals in 2016 in Iran around 1.6 million individuals, out of which more than 1.5 million (above 95%) were Afghans\(^1\). However, during the same year there were less than 250 thousand work permits owned by foreign individuals in Iran\(^2\), indicating that a large number of foreign individuals,

---

1. See the official report (in Persian).
2. See the official report (in Persian) from Iran’s Ministry of Cooperatives Labour and Social Welfare for details.
74.6% of which were between 10 and 64 years old\textsuperscript{1}, did not have work permits. As a result of the fact that the majority of these individuals do not have high educations, they are forced to accept demanding, low-wage, informal jobs to earn a living. The majority of these workers find jobs in the construction sector of the Iranian economy (Eisazadeh & Mehranfar, 2014). These workers are among the most vulnerable of the workforce and do not receive the essential benefits, such as insurances, despite the hazardous work environment. The employers’ reluctance in disclosing their employment of workers with no work permit to the authorities, combined with the persistent daily use of cash to settle payments for these workers, might produce a large shadow section in the construction sector.

It should be noted that since 2016, due to the decreasing living standards in Afghanistan there has been a larger influx of new immigrations to Iran, so much that according to an official report published by the UNHCR (United Nations High Commissioner for Refugees) in February 2021, there were more than 3 million Afghan immigrants living in Iran, out of which between 2.1 and 2.25 million were undocumented and 780 thousand were refugees\textsuperscript{2}. Therefore, it’s likely that the number of workers without permits working in the Iranian sectors, including the construction sector, has increased even more.

We hypothesize that an increase in the size of the construction sector should result in an increase in the size of the SE, ceteris paribus.

Inflation: When nominal disposable income is constant, and the real purchasing power is falling due to inflation, attempts will be made to restore the real value of disposable income. In order to achieve this, one possible strategy is to avoid tax payments, raising the available disposable income (Fishburn, 1981). Greater inflation also would likely result in increased production costs for the firms; therefore, the affected firms might opt to

\textsuperscript{1} See the official report (in Persian) from Iran’s Ministry of Cooperatives Labour and Social Welfare for details.

\textsuperscript{2} See the original report for more information.
operate in the shadow sector of the economy to dodge other costs, such as tax payments. Increased inflation also might raise discount rates, thereby prompting underground activities (Goel & Nelson, 2016). The positive relationship between inflation and the SE size has been empirically shown to exist in studies such as Maddah (2014).

Accordingly, we hypothesize the inflation rate to be positively related to the shadow activities ceteris paribus.

3.2. Indicators of the SE

Currency usage (Currency Outside the Banks to M2 ratio): In an effort to dodge detection and remain hidden, actors who participate in the shadow activities prefer to use cash in their transactions, as opposed to cheques or bank transactions. As a result, all else being equal, a bigger SE should translate to a bigger cash holding and usage in the economy. The positive link between SE and currency held by the public has been empirically shown to be significantly positive by different studies such as Hassan and Schneider (2016b).

Therefore, we expect a positive relationship between the SE and cash usage in our model, ceteris paribus.

Real GDP per capita index (Real GDP per capita / Real GDP per capita 2004): As the SE grows, productive factors and resources are absorbed by the SE, thereby depressing official economic productivity and growth (Dell’Anno & Solomon, 2008). Results of studies by Baklouti & Boujelbene (2019) for 33 developed and 14 developing countries and Schneider (2012) for 21 OECD countries indicate that the SE is negatively related to the official economy. As such, we hypothesize a negative relationship between SE growth and GDP of Iran, ceteris paribus.

Labor Force Participation Rate: A growth in the shadow activities also may impact the official labor force participation rate as the human resources move from the official economy into the shadow one. Studies by Schneider et al. (2010) and Buehn & Schneider (2012a) have found a negative and
significant relationship between the SE and the labor force participation rate. Based on the results of the usage of this indicator in the SE literature, we expect a negative relationship between the shadow economy and the labor force participation rate.

4. Empirical Results
The causes included in our MIMIC models due to their potential importance are the inflation rate, trade openness, unemployment rate, size of the construction sector, government size, and tax burdens\(^1\). The considered indicators, which reflect the movements of the SE size, are the currency outside the banks to M2 ratio, the index of GDP per capita, and the labor force participation rate.

We begin our MIMIC estimation process by first testing the characteristics of our time series. These tests include the unit root tests and tests for multivariate normality.

To check the stationarity of our variables, we use the Augmented-Dickey Fuller (ADF) and Philip-Peron (PP) tests. Test outcomes are available in table A.2. Results show that all the variables, both the causes and the indicators, are non-stationary at level. Therefore, all variables used in the model have been transformed into first difference to achieve stationarity\(^2\).

According to Hassan & Schneider (2016a), two of the requirements of MIMIC estimations are the multivariate normality of variables and the largeness of sample size (\(N>50\)). The assumption of multivariate normality is essential in ensuring that the statistical properties of estimators and the Chi-square calculated are preserved (Dell’Anno, 2003). We test for the multivariate normality of our variables using the Doornik-Hansen test.

---

1. Taxes on wealth and goods and services were also tested and were found to be statistically insignificant. However, they are included in the “total tax % GDP” variable.
2. Naturally, co-integration issues and application of ECM models should be considered, however, limited research has been done regarding the steps of this procedure in MIMIC estimations and, as a result, in this study we opt to follow Dell’Anno et al. (2007) and use differenced variables in our models which results in the removal of unit roots and better goodness of fit statistics.
results indicate that our variables violate the assumption of multivariate normality and are non-normal\(^1\). In the face of non-normal and skewed data, we use the Satorra-Bentler (SB) option in our SEM estimation that produces improved goodness of fit statistics when dealing with non-normal data (Satorra & Bentler, 1994). Our data also suffers from a rather small sample (\(N=41\)). Consequently, another assumption is violated. To solve for this issue, we also include in our estimations the Swain correction that adjusts and improves the chi-square for small and/or complex models\(^2\).

Table 1 reports the results of our estimations and their goodness of fit indices. While government size is statistically significant in some models and insignificant in others, none of the different types of tax burdens are found to be statistically significant. Based on the goodness of fit statistics, we choose the 4-1-3 model as the optimal model in describing the SE of Iran.

<table>
<thead>
<tr>
<th>Effect of causes on the latent variable</th>
<th>Dependent variable: Shadow economy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6-1-3A</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔInflation Rate</td>
<td>0.003*</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔUnemployment Rate</td>
<td>0.021**</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔTrade Openness</td>
<td>-0.003**</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔGovernment Expenditures to GDP</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔConstruction Sector’s Size</td>
<td>0.024**</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔTotal Tax burden</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔIncome Tax burden</td>
<td></td>
</tr>
</tbody>
</table>

1. The test rejects multivariate normality at 1% significance level for all of our MIMIC models.
2. Here we use the “swain” module developed for this purpose in Stata by Antonakis & Bastardoz (2020).
<table>
<thead>
<tr>
<th>Effect of the latent variable on the indicators</th>
<th>Dependent variable: Shadow Economy</th>
<th>Dependent variable: Currency Outside of Banks / Money (M2)</th>
<th>Dependent variable: Labor Force Participation Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corporate Tax burden</td>
<td>$\Delta$-0.001</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>Import Tax to burden</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** *, **, *** indicate significance at the 10%, 5% and 1% respectively. Chi-square p-value=1 corresponds to perfect fitting. Smaller AICs and BICs indicate better model fit. RMSEA<0.06 and SRMR<0.08 are accepted. CFI & TLI should be above 0.95 for acceptance (Schreiber et al., 2006). Closer values of CD to 1 are considered better. Source: Research findings
In model 4-1-3 (Figure 3), inflation, unemployment rate, and size of the construction sector are found to be significant and positively impact the SE size, while trade openness is also significant but maintains a negative effect on the SE. The SE, in turn, has a negative relationship with the official economy and labor force participation rate as expected and a positive link to cash usage. The results of our chosen MIMIC model are in line with the SE literature as all of our significant variables, in both the structural and the measurement model, have the theoretically expected sign.

After choosing the optimal MIMIC model, the index of the size of the SE can be calculated for each year through a linear prediction of the structural model (Eq.3).

\[ \eta_t = 0.0028 \Delta \text{(inflation}_t) - 0.0029 \Delta \text{(trade openness}_t) + 0.0203 \Delta \text{(unemployment rate}_t) + 0.0208 \Delta \text{(construction sector’s size}_t) \]  

\[ (3) \]
The latent variable index ($\eta$) has the same unit of measure as the variable of scale. Since the variable of scale (reference indicator) is the first difference of real GDP per capita to real GDP per capita of 2004 ($\frac{\Delta(\text{real GDP per capita})}{\text{real GDP per capita}_{2004}}$), the index of the SE is also calculated as the first difference of the real SE per capita to real GDP per capita of 2004 ($\frac{\Delta(\text{real SE per capita})}{\text{real GDP per capita}_{2004}}$).

In order to find the ratio of the SE to current GDP, we take a two-step procedure:

The first step includes transforming the latent variable, which is in first difference, back to level. By definition, first difference of $X_t$ is:

$$\Delta X_t = X_t - X_{t-1}. \quad (4)$$

As such, we can calculate the first difference of real SE per capita index ($\frac{\Delta(\text{real SE per capita})}{\text{real GDP per capita}_{2004}}$) as

$$\frac{\text{real SE per capita}_t}{\text{real GDP per capita}_{2004}} = \frac{\Delta(\text{real SE per capita})}{\text{real GDP per capita}_{2004}} + \frac{\text{real SE per capita}_{t-1}}{\text{real GDP per capita}_{2004}} \quad (5)$$

and

$$\frac{\text{real SE per capita}_{t+1}}{\text{real GDP per capita}_{2004}} = \frac{\Delta(\text{real SE per capita})}{\text{real GDP per capita}_{2004}} + \frac{\text{real SE per capita}_t}{\text{real GDP per capita}_{2004}}. \quad (6)$$

In order to use the above equations, we need an exogenous estimation for the size of the SE as a percentage of the official GDP in the year 2004. This value is taken be 18.45% of the GDP in 2004, which is the average of several estimates. The reason why this year was chosen is the fact that the estimation for this year’s SE size is available from several studies and, what’s more, the estimated values for this year are all within the 16-21% range and close to each other. The sources and the estimates for this year are given in Table 2.
Table 2. The estimated value for SE size of Iran in 2004

<table>
<thead>
<tr>
<th>Source</th>
<th>Estimated Value</th>
<th>Method of estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Sameti et al., 2010)</td>
<td>20.20%</td>
<td>MIMIC</td>
</tr>
<tr>
<td>(Schneider &amp; Savaşan, 2007)</td>
<td>20.50%</td>
<td>MIMIC</td>
</tr>
<tr>
<td>(Medina &amp; Schneider, 2018)</td>
<td>16.01%</td>
<td>MIMIC</td>
</tr>
<tr>
<td>(Medina &amp; Schneider, 2019)</td>
<td>17.10%</td>
<td>MIMIC</td>
</tr>
<tr>
<td>Average Value</td>
<td>18.45%</td>
<td>-</td>
</tr>
</tbody>
</table>

Source: Respective authors.

By inserting this exogenous value (18.45%) into Eq.5 & Eq.6, our estimated index is transformed into real SE per capita as a percentage of the real GDP per capita of the year 2004 \( \left( \frac{\text{real SE per capita}_t}{\text{real GDP per capita}_{2004}} \right) \). We then proceed to apply the benchmarking procedure employed by Dell’Anno & Schneider (2006) in the second step.

The second step includes the usage of the following equation:

\[
\left( \frac{SE_t \cdot pop_{2004}^{t}}{pop_t \cdot GDP_{2004}} \right) \cdot \left( \frac{SE_{2004}}{GDP_{2004}} \right) \cdot \left( \frac{1}{\left( \frac{SE_{2004} \cdot pop_{2004}^{2004}}{pop_{2004} \cdot GDP_{2004}} \right)} \right) \cdot \left( \frac{GDP_{2004} \cdot pop_t}{pop_{2004} \cdot GDP_t} \right) = \frac{SE_t}{GDP_t} \quad (7)
\]

In Eq.7, \( \left( \frac{SE_{2004} \cdot pop_{2004}^{2004}}{GDP_{2004}} \right) \) is the outcome of the first step for the year 2004, \( \left( \frac{SE_{2004} \cdot pop_{2004}^{2004}}{pop_{2004} \cdot GDP_{2004}} \right) \) is the outcome of the first step for the year 2004, \( \left( \frac{GDP_{2004} \cdot pop_t}{pop_{2004} \cdot GDP_t} \right) \) is there to convert the size of the SE from being in respect to the year 2004 to the year \( t \), and \( \frac{SE_t}{GDP_t} \) is the size of the SE for the year \( t \) as a percentage of the GDP in year \( t \), the final result of our estimation.

Figure 4 depicts the size and trends of the Iranian SE with respect to GDP based on the results of the 4-1-3 MIMIC model.
According to our estimation based on the MIMIC model of choice, the SE size has ranged from a minimum of 13.98% of the GDP to a maximum of 64.01% of the GDP between 1979 and 2019. Based on Figure 4, it is evident that after an initial and considerable expansion of the SE between 1980 and 1986, taking place during the eight-year-long Iran-Iraq War (1980-1988), the Iranian SE experienced a generally downward trend to reach levels of as low as 16.83% of GDP in 2006. Following 2006, the Iranian SE saw an expansion over the 2007-2012 period and a subsequent abatement over the 2013-2017 period. Based on the MIMIC estimation, the average SE size for the full studied period, that is 1979 to 2019, is 31.83% of the GDP and for the post-war period (1989 onward) the SE averages 27.73% of the GDP.

The main drivers of Iranian SE are inferred from the MIMIC model 4-1-3 to be inflation, unemployment rate, trade openness, and the size of the construction sector. As a result of the simultaneous effect of these major causes on the SE, which frequently put pressure on the SE in opposing directions, it is difficult to say with certainty which factor was the major contributor to the changes in the SE for each year. However, examining the changes in the values of SE’s significant causes for specific periods can be
of use. Table 3 reports (in percentage points) the averages of annual changes in the major determinants of the SE during time periods in which the SE underwent significant and sudden changes. Based on Table 3, it is evident that the increases in the size of SE over the periods 1980-1986 and 2007-2012 were due to rises in unemployment, inflation, and size of the construction sector as well as decreases in trade openness. Similarly, it can be seen that the SE plummeted during the periods 1989-1990 and 2013-2017 as a consequence of drops in inflation, unemployment, and size of the construction sector, while the economy became more open to trade. Over the period 1996-2004, the fall in inflation rate and unemployment rate and rise in trade openness caused a decrease in the SE even though during this period size of the construction sector increased. Finally, for the two year period of 2018-2019, a large increase in the inflation rate (+13.15 percentage points on average) accompanied by a growth in the size of the construction sector induced an increase in the size of the SE, offsetting the shrinking effects that a fall in the unemployment rate and rise in trade openness have on the SE.

Table 3. The changes in the major drivers of the SE

<table>
<thead>
<tr>
<th>Period</th>
<th>Unemployment Rate</th>
<th>Inflation Rate</th>
<th>Trade Openness</th>
<th>Construction Sector's Size</th>
<th>SE%GDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980-1986</td>
<td>0.63</td>
<td>1.76</td>
<td>-4.09</td>
<td>0.03</td>
<td>6.18</td>
</tr>
<tr>
<td>1989-1990</td>
<td>-1.00</td>
<td>-9.95</td>
<td>7.75</td>
<td>-0.27</td>
<td>-14.36</td>
</tr>
<tr>
<td>1996-2004</td>
<td>-0.04</td>
<td>-3.87</td>
<td>1.80</td>
<td>0.04</td>
<td>-2.63</td>
</tr>
<tr>
<td>2007-2012</td>
<td>0.15</td>
<td>2.48</td>
<td>-0.97</td>
<td>0.58</td>
<td>2.25</td>
</tr>
<tr>
<td>2013-2017</td>
<td>-0.06</td>
<td>-4.10</td>
<td>0.28</td>
<td>-0.84</td>
<td>-3.26</td>
</tr>
<tr>
<td>2018-2019</td>
<td>-0.65</td>
<td>13.15</td>
<td>3.70</td>
<td>0.37</td>
<td>3.25</td>
</tr>
</tbody>
</table>

Note: values reported are averages of the annual changes (in percentage point).

Source: Research findings

5. Conclusion

Researches that aim to investigate and assess the Iranian shadow economy (SE) individually have been very limited and almost exclusively published
within Iran and in Persian. This paper investigates the causes of the Iranian SE and provides an up-to-date estimation for the Iranian SE through the usage of the MIMIC (Multiple Indicators, Multiple Causes) approach using annual data from 1979 to 2019.

Application of the MIMIC approach permits us to consider various variables that are commonly viewed as the main drivers of the SE (causes) alongside variables through which the movements of the SE are reflected (indicators). We include the size of the construction sector as a potential cause of the Iranian SE in this study for the first time. We find that inflation, unemployment rate, and size of the construction sector have a statistically significant positive effect on the SE whilst the degree of the openness of the economy to trade is also statistically significant but has an inverse relationship with the size of the SE. The effects of the other causes are found to be insignificant. Concerning the indicators, our results show that increases in the SE are accompanied by decreases in GDP per capita and labor force participation rate but increases in cash usage.

Analysis of the estimations produced by our best MIMIC model (4-1-3) shows that the Iranian SE has experienced rapid fluctuations and heavy inflations and deflations in the past. It appears that after an initial expansion during Iran-Iraq War, the Iranian SE declined overall, fluctuations notwithstanding. More recently, however, there was an uptick in SE between 2018 and 2019. The average SE/GDP value over the entire 1979-2019 period is found to be 31.83%.

Naturally, if the authorities wish to reduce the Iranian SE size, their policy formulation should be directed toward reducing it through the channel of its major drivers. In other words, decreases in SE size can be achieved via initiatives that result in lower inflation and unemployment rates as well as higher trade levels and smaller amounts of shadow economy activities in the construction sector. While Iranian high inflation rates have been generally accepted to be chiefly a result of substantial money supply growth, the importance of other factors such as budget deficits and Iranian economic
sanctions have been highlighted by the literature as well\(^1\); therefore, not only is there a need for both monetary and fiscal policy reform but also for international negotiations that would result in the abatement of sanctions. The lifting of sanctions is doubly important as they are also a significant barrier to trade, thereby lowering trade openness and increasing the SE. In addition, a more diligent monitoring of the construction sector can result in less informal activity and safer work environments for the undocumented workers. Lastly, policies that provide incentives for entrepreneurs to create employment opportunities, such as tax cuts or lessening of business regulations, should be put in place as means of lowering the unemployment rate and, as a consequence, the SE size.

Estimating the size of the SE, which is by nature hidden, is a very challenging task. This is further complicated by shortcomings of the various estimation approaches which virtually all have been subject to criticism. In light of these limitations, estimations of the SE should always be regarded as rough approximations and not exact measurements. Therefore, any interpretation and subsequent policymaking based on any estimation of the SE, should be done with extreme caution. Nevertheless, our results provide a better view into the nature of the SE of Iran, its causes, its indicators, and its movements which can help the authorities implement measures to move the shadow activities toward the official economy.

**Competing Interests**
No competing interests.

**Funding**
This research received no funding.

---

\(^1\) For a more detailed discussion about the determinants of inflation in Iran see Khandan and Hosseini (2016), Hemmati, Niakan, and Varahrami (2018) and Dastgerdi, Yusof, and Shabazz (2018).
References


Informal Workforce: Priorities for Inclusive Growth, International Monetary Fund, 10–69.


Appendix:

Table A.1. Variables description & sources of data

<table>
<thead>
<tr>
<th>Sym</th>
<th>Variable</th>
<th>Unit Root [Transformation]</th>
<th>Sources [observations; mean; minimum; maximum; standard deviation]</th>
<th>Description &amp; notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>Inflation rate</td>
<td>$I(1){Δ(X1)}$</td>
<td>CBI - SCI [41; 19.42; 6.80; 49.40; 8.87]</td>
<td>Base year = 2016, annual</td>
</tr>
<tr>
<td>X2</td>
<td>Unemployment rate</td>
<td>$I(1){Δ(X2)}$</td>
<td>CBI - MPORG [41; 11.91; 9.10; 14.50; 1.35]</td>
<td>Defined as: unemployed population (10 years or more)/total active population (10 years or more)*100</td>
</tr>
<tr>
<td>X3</td>
<td>Trade Openness</td>
<td>$I(1){Δ(X3)}$</td>
<td>WDI [41; 40.84; 14.14; 65.05; 10.39]</td>
<td>Defined as: (volume of import + export)/GDP*100</td>
</tr>
<tr>
<td>X4</td>
<td>Government size</td>
<td>$I(1){Δ(X4)}$</td>
<td>WDI [41; 13.05; 9.24; 21.45; 2.98]</td>
<td>Defined as: government expenditure/GDP*100</td>
</tr>
<tr>
<td>X5</td>
<td>Tax burden - Total</td>
<td>$I(1){Δ(X5)}$</td>
<td>CBI [40; 5.75; 3.71; 8.48; 1.67]</td>
<td>Defined as: sum of all taxes/GDP*100 - data is missing for 2019, we replace the missing value with the average amount of 2011-2018</td>
</tr>
<tr>
<td>X6</td>
<td>Size of the Construction sector</td>
<td>$I(1){Δ(X6)}$</td>
<td>CBI [41; 7.52; 4.97; 11.86; 1.80]</td>
<td>Defined as: construction sector value added/GDP*100</td>
</tr>
<tr>
<td>X7</td>
<td>Tax burden - Corporate</td>
<td>$I(1){Δ(X7)}$</td>
<td>CBI [40; 2.08; 0.68; 4.07; 0.62]</td>
<td>Defined as: corporate tax/GDP*100 - data is missing for 2019, we replace the missing value with the average amount of 2011-2018</td>
</tr>
<tr>
<td>X8</td>
<td>Tax burden - Income</td>
<td>$I(1){Δ(X8)}$</td>
<td>CBI [40; 0.89; 0.66; 1.18; 0.14]</td>
<td>Defined as: income tax/GDP*100 - data is missing for 2019, we replace the missing value with the average amount of 2011-2018</td>
</tr>
<tr>
<td>X9</td>
<td>Tax burden - Import</td>
<td>$I(1){Δ(X9)}$</td>
<td>CBI [40; 1.46; 0.63; 2.54; 0.44]</td>
<td>Defined as: import tax/GDP*100 - data is missing for 2019, we replace the missing value with the average amount of 2011-2018</td>
</tr>
<tr>
<td>Y1</td>
<td>Index of GDP per capita</td>
<td>$I(1){Δ(Y1)}$</td>
<td>WDI [41; 0.94; 0.65; 1.24; 0.17]</td>
<td>Defined as: real GDP per capita / real GDP per capita 2004, constant 2010 dollars</td>
</tr>
<tr>
<td>Y2</td>
<td>Currency Usage</td>
<td>$I(1){Δ(Y2)}$</td>
<td>CBI [41; 12.86; 2.47; 27.31; 8.10]</td>
<td>Defined as: Notes &amp; Coins with the Public/M2*100</td>
</tr>
<tr>
<td>Y3</td>
<td>Labor Force Participation Rate</td>
<td>$I(1){Δ(Y3)}$</td>
<td>MPORG-SCI [41; 38.45; 34.70; 41.40; 1.62]</td>
<td>Defined as: total active population (10 years or more)/total population (10 years or more)*100</td>
</tr>
</tbody>
</table>

Table A.1. WDI indicates World Development Indicators (from World Bank) – CBI stands for Central Bank of Iran – SCI stands for “Statistical Center of Iran” and MPORG refers to “Plan and Budget Organization of the Islamic Republic of Iran”.

Source: As stated in the table.
Table A.2. The stationary tests

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit root [Tranf.]</th>
<th>At level</th>
<th>First Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( I^T )</td>
<td>( I^T &amp; I )</td>
<td>( N )</td>
</tr>
<tr>
<td>Inflation rate</td>
<td>-4.32***</td>
<td>-4.20***</td>
<td>-0.91</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>-3.22**</td>
<td>-3.30*</td>
<td>-0.20</td>
</tr>
<tr>
<td>Trade % GDP</td>
<td>-1.37</td>
<td>-2.51</td>
<td>0.03</td>
</tr>
<tr>
<td>Government Expenditures % of GDP</td>
<td>-2.41</td>
<td>-2.32</td>
<td>-1.39</td>
</tr>
<tr>
<td>Income Tax % GDP</td>
<td>-3.01**</td>
<td>-3.00</td>
<td>-0.73</td>
</tr>
<tr>
<td>Corporate Tax % GDP</td>
<td>-3.80***</td>
<td>-4.52***</td>
<td>-0.94</td>
</tr>
<tr>
<td>Total Tax % GDP</td>
<td>-2.91*</td>
<td>-3.17</td>
<td>-0.38</td>
</tr>
<tr>
<td>Import Tax % GDP</td>
<td>-2.37</td>
<td>-2.68</td>
<td>-1.61</td>
</tr>
<tr>
<td>Construction Value Added % GDP</td>
<td>-2.73*</td>
<td>-3.03</td>
<td>-1.00</td>
</tr>
<tr>
<td>Currency Outside of Banks to M2 Ratio</td>
<td>-1.08</td>
<td>-0.98</td>
<td>-3.82***</td>
</tr>
<tr>
<td>Real GDP Per Capita Index [2004]</td>
<td>-1.99</td>
<td>-2.21</td>
<td>-0.64</td>
</tr>
<tr>
<td>Labor Force Participation Rate</td>
<td>-1.99</td>
<td>-1.72</td>
<td>-0.38</td>
</tr>
</tbody>
</table>

Note: ***, ** and * mean stationary at 1%, 5% and 10%, respectively. Lag length for the ADF (Augmented Dickey-Fuller) test was chosen using Schwarz Information Criterion. For the PP (Philips-Perron) test the Bartlett kernel method with Newey-West bandwidth was used. “I” refers to a model with an intercept. “T&I” refers to a model specification with a trend and an intercept and “N” refers to a model with neither an intercept (constant) nor a trend.

Source: Research findings