

2017. 12. 13(vol. 4)

# KIPF BRIEF

## Estimation of Income Tax Gap and Underground Economy in Korea

---

An, Jongseok Senior Fellow  
Kang, Sung Hoon Associate Fellow  
Oh, Jonghyeon Associate Fellow



---

## Estimation of Income Tax Gap and Underground Economy in Korea

An, Jongseok Senior Fellow (044-414-2210)

Kang, Sung Hoon Associate Fellow (044-414-2289)

Oh, Jonghyeon Associate Fellow (044-414-2289)

---

I. Introduction .....	02
II. Estimation of Income Tax Gap .....	03
III. Estimation of Underground Economy .....	15



## I Introduction

- ● Over the past several years, the legalization of the underground economy and the reduction of tax evasion have become important tasks of the National Tax Service.
  - Reduction of tax evasion has the effect of expanding the tax base and is also expected to contribute to improving the fairness of taxation so that the credibility of governmental policies such as tax policies can be enhanced and social integration can be achieved.
- ● To reduce tax evasion, the effectiveness of tax investigation should be improved through the use of scientific methods. To this end, it is necessary to grasp the current state of the size and structure of tax evasion per tax item. Based thereon, the investigation system can be reorganized, including the establishment of the tax investigation plan and personnel reshuffling.
  - Tax gap refers to the gap between the theoretical tax burden and the actual tax burden.
    - The theoretical burden refers to the tax amount that must be paid when tax law is applied correctly.
    - Tax gap is a concept that covers both tax evasion and tax avoidance.
  - As part of the purpose, the aggregate income tax gap is estimated in this study.
    - Tax gap can be classified into the underreported gap and the underpayment gap. In this study, the underreported gap is divided into two parts: the observed underreported gap, and the unobserved underreported gap.
    - The observed underreported gap refers to the underreported amount that the National Tax Service has found through its investigation for taxable income, etc., after the due date, and the unobserved underreported gap refers to the underreported amount that the National Tax Service has not found.
- ● In addition, this study estimates the scale of the underground economy in Korea.

- Either micro or macro data sets can be used when performing underground economy estimation. This study utilizes macro data.
- Among the estimation methods using macro data, we employ the most widely used one that is based on the function of the demand for cash currency.
  - Taking into consideration the fact that estimation results may differ depending on model type and variables used, the scale of underground economy is estimated by using multiple models and variables. Implications will be derived based on a comparison of estimation results..

## II Estimation of Income Tax Gap

- The main purpose of this study is to estimate the underreported gap of aggregate income tax.
  - Aggregate income tax payers can be classified into self-employed entrepreneurs and non-entrepreneurs. The most important part of the underreported gap estimation is the aggregate income tax of self-employed entrepreneurs.
  - The aggregate income tax of business income earners is mostly composed of business income tax, which is levied based on tax returns submitted by entrepreneurs. There are too many self-employed entrepreneurs, and a considerable portion of them are targeting for final consumption with a high portion of cash transactions. Therefore, the likelihood of the underreporting and the size of underreported tax amount seem to be relatively large.
    - According to 3-1-2 Filing of Total Revenue and Taxable Income in the Statistics Yearbook of National Tax, business and real estate incomes accounted for 93.59% of the total aggregate income reporting as of 2011 (attribution).

- The Income of non-entrepreneurs consists of wage income, interest and dividend income, pension income, and other incomes. In this case, it is easy for the National Tax Service to get informations needed for taxation—i.e. for evaluating the accuracy of tax returns. Thus, it is expected that the underreported gap among non-entrepreneurs would not be large.
  - Still, there are cases in which the National Tax Service tracks down underreported gaps through amended tax returns or reassessment by investigation.
- ● This study estimates the underreported gap of the aggregate income tax by dividing it into the observed underreported gap and the unobserved underreported gap.
  - The observed underreported gap refers to the underreported tax amount that the National Tax Service has found in tax returns in which taxpayers underreported intentionally or by negligence, and the unobserved underreported gap refers to the underreported tax amount that the National Tax Service has not found.
  - The observed underreported gap includes underreported gaps captured by the National Tax Service in both the entrepreneur and non-entrepreneur groups by means of amended tax returns or investigations.
  - The unobserved underreported gap is estimated only for the entrepreneur group.
    - Since the entrepreneur group is likely to commit an act of tax evasion, there may be the underreported gap that the National Tax Service has not found. On the other hand, since the non-entrepreneur group is unlikely to commit an act of tax evasion, the underreported gap is likely to be insignificant.
- ● The unobserved underreported gap of the entrepreneur's aggregate income tax is estimated by investigating randomly selected self-employed entrepreneurs.
  - The National Tax Service conducted random sampling of self-employed entrepreneurs who had filed aggregate income tax returns on income earned in the years 2012 and 2013, and conducted an investigation for the purpose of estimating the underreported income tax gap.

- This study estimates the underreported gap of the aggregate income tax of 2011, using this data.
  - An Jongseok et al. (2016) estimated the tax gap of other tax items, excluding aggregate income tax. To be consistent with the previous estimation, we estimate the underreported gap by taking the year 2011 as the base year.
  - However, since there is no data from random samples for the year of 2011, the underreported gap mentioned above is estimated by applying to the population of 2011, the analysis results that are estimated using investigations of taxable income for 2012 through 2013.

## 1. Estimation of Unobserved Underreported Gap

### 1) The Method of Estimation

- In this section, we estimate the unobserved underreported gap of the aggregate income tax of self-employed entrepreneurs using random sample data and yield the total underreported gap. And in the next section, we estimate the total unobserved income tax gap by adding the unobserved underreported gap to the observed underreported gap.
  - The unobserved underreported gap refers to the underreported tax amount that the National Tax Service has failed to collect since it could not find errors, omissions, or tax evasions in income tax returns in which taxpayers underreported intentionally or by negligence.
  - Data on tax evasion is collected based on the investigation for taxable income of the randomly chosen samples.
  - The samples were extracted from taxpayer groups that did not file amended tax returns or have not been investigated.
    - Since the National Tax Service did not conduct additional reviews of tax returns filed by this group, the underreported gap would not have been observed.

- ● The easiest way to estimate the underreported gap of a population using results from random sample survey is to estimate the underreported gap of a population by reflecting the underreporting ratio in each group within the random sample in the population without modification.
  - If the sample mean is highly reliable thanks to a large number of samples, this method would be sufficient.
  - However, we have only 500 samples in 2012 and 497 in 2013—i.e. 997 samples in total for the two years. Thus, the number is relatively small.
  
- ● To address this problem, we tried two different methods. One is using the sample raw data as described above. The other is an estimation method using an econometric model. The brief description of each method is as follows:
  
- ● (Sample raw data analysis) We calculate the ratio of the underreported tax amount to the tax amount reported in the tax return, and estimate the underreported gap by applying the ratio to the population.
  - (Step 1) For the random samples, we calculate both the reported liability of income tax from their tax returns and the underreported liability of income tax from the random sample investigation by groups divided by industry and income level.
  - (Step 2) We calculate the ratio of the underreported tax amount to the reported tax, using the result calculated in (Step 1) for each group.
  - (Step 3) We apply the ratio calculated in (Step 2) to the reported tax liability of each group of population to estimate the unobserved underreported gap.
  - (Step 4) We calculate the total underreported gap by adding all underreported gaps of groups estimated in (Step 3).

- (Econometric model analysis) (1) We estimate a regression equation for the probability of underreporting and the ratio of underreport by each business according to the characteristics of the business through an econometric model, and (2) estimate the underreported gap by reflecting taxpayers' information of the population (the ratio of tax base to total income, total income, industry) on the estimated regression equation.
  - (Step 1) Assuming that the difference in the ratio of tax base to total income, total income, and industry affects the probability of underreporting and the ratio of underreport, we construct a regression equation that can describe the probability of underreporting and the ratio of underreport.
    - The ratio of underreport refers to  $[\text{underreported tax amount} / (\text{reported tax liability} + \text{underreported tax amount}) \times 100]$ .
  - (Step 2) We estimate the regression equations constructed in (Step 1), using Tobit and Truncated Normal Hurdle estimation methods.
  - (Step 3) We reflect informations in individual tax returns of the population (the ratio of tax base to total income, total income, and a dummy variable for industry) on the regression equation estimated in (Step 2), and thereby estimate the underreported tax amount of individuals in the population.
  - (step 4) We estimate the total unobserved underreported tax amount by adding all tax amounts underreported by individual taxpayers.
- The advantage of econometric model analysis is to estimate an individual taxpayer's underreporting by estimating factors of a taxpayer's underreporting through the econometric model, and applying an individual taxpayer's characteristics to the result thereof.
  - In other words, the advantage is that we can make direct use of estimation results on underreporting of the population in identifying what causes underreporting and developing a policy to reduce underreporting since the model allows us to identify the characteristics of each business that cause underreporting and estimate underreporting

by the population based on such characteristics.

- ● On the other hand, the biggest disadvantage of econometric model analysis is that, even though we considered the ratio of tax base to total income, income amount, and industry-specific characteristics as factors affecting the behavior of underreporting, other possible factors cannot be effectively counted in a regression analysis.
  - Besides major factors affecting tax evasion taken into consideration in a regression analysis, the effect of other factors is not observed; so, it is included in the error term.
  - The more variables are taken into consideration in a model, the better the explanatory power of the model. However, among the currently available data, there are no other indicators we find useful for improving the explanatory power of our model.
- ● In this study, we conduct sample raw data analysis and econometric model analysis in parallel, and thereby compare the underreported amount that can be estimated by the econometric model and the underreported amount that cannot be obtained by the econometric model.

## 2) Unobserved Underreported Gap

- ● The ratio of unobserved underreporting estimated by sample raw data analysis is about 31%, which is six to nine percentage points larger than that estimated using the econometric model.
  - The result shows that other factors than the ratio of tax base to total income, income amount, and industry-specific characteristics may affect an act of evasion.
    - About 80% of the underreported gap that is revealed in sample raw data analysis can be explained in the econometric model; however, about 20% can be interpreted as being

caused by other factors not included in the model or simple errors.

〈Table 1〉 Unobserved underreported gap of income earned in 2011

(unit: one million, %)

Classification	Sample raw data analysis	Tobit Model analysis	Truncated Normal Hurdle Model analysis
Underreported gap	3,508,873.8	2,504,073.0	2,202,706.1
Underreporting ratio <sup>1)</sup>	31.39	24.62	22.32

Note: 1) The underreporting ratio refers to the percentage of the unobserved underreported gap to the sum of the reported tax liability and the underreported tax amount. The observed underreported gap is not included.

Source: prepared by the authors

- The difference between the result of sample raw data analysis and the result of econometric model analysis is too large to consider it to be caused by a simple error. This indicates that it is necessary to develop a model to explain the underreported gap accurately.
  - In particular, it is necessary to construct data on the probability of underreporting among the characteristics of individual business and factors affecting the size of the underreported gap where underreporting occurs.
    - In this study, we only used the ratio of the reported tax base to total income, total income, and dummy variable for industry as a result of reviewing available data among the random sample survey data. It is necessary to find additional methodologies that are capable of generating other data.
- Among econometric models, the underreported gap estimated using the Tobit Model is higher than in the case of the Truncated Normal Hurdle Model.
  - The reason for this difference is likely that the Tobit Model explains the decisions on whether or not underestimation occurred, and the underreported tax amount with one mechanism, while the Truncated Normal Hurdle Model distinguishes and explains them with two different mechanisms.

- Since the Truncated Normal Hurdle Model is based on the assumption that the decisions on whether or not to underreport and underreporting amount are independent of one another, unlike the Tobit Model, the ratio of reported tax base to total income, income amount, and industry-specific characteristics may have different impacts on the probability of underreporting and the underreport ratio.
- Comparing the two results, we could find that both the sizes and signs of estimated coefficients for explanatory variables are not be the same.
- The Tobit Model has a limitation in estimating the coefficients of explanatory variables, compared to the Truncated Normal Hurdle Model. This seems to cause the difference of the estimated underreported gaps of the two models.

## 2. Estimation of Observed Underreported Gap and Total Underreported Gap

### 1) Estimation of Observed Underreported Gap

- ● In this study, we calculated the observed underreported gap of the aggregate income tax imposed on entrepreneurs and non-entrepreneurs—i.e. the underreported amount captured through reassessment after the due date—by using the National Tax Service’s reassessment data on income earned in 2011.
  - The observed underreported gap can be classified into the underreported gap of self-employed entrepreneurs and the underreported gap of non-entrepreneurs.
- ● The observed underreported gap is calculated by deducting the reported tax liability in the initial tax return from the reassessed tax liability. Penalty tax is not included.
  - The observed underreported gap = [Reassessed tax liability – reported tax liability in the initial report]
- ● The observed underreported gap of 2011 (underreported amount captured through reassessment) is 1 trillion 3,856 billion won, and the observed underreporting ratio is 22.59%.

- The observed underreported gaps of the entrepreneur group and the non-entrepreneur group are 1 trillion 3,350 billion won, and 50.7 billion won, respectively; the observed underreporting ratios thereof are 23.67% and 11.38%.

〈Table 2〉 Observed underreported gaps and underreporting ratios of 2011 of entrepreneur group and non-entrepreneur group

(unit: million won, %)

Classification	Reported tax liability	Observed underreported gap <sup>1)</sup>	Observed underreporting ratio <sup>2)</sup>
	A	B	$(B/(A+B)) \times 100$
Entrepreneur	4,304,065.6	1,334,975.7	23.67
Non-entrepreneur	445,311.6	50,664.6	11.38
Total	4,749,377.2	1,385,640.3	22.59

Note: 1) Penalty tax is not included in reported and reassessed tax liabilities.

2) The observed underreporting ratio refers to the percentage of the observed underreported gap to the sum of the reported tax liability and the observed underreported tax amount. The unobserved underreported gaps are excluded.

Source: The National Tax Revenue

## 2) Total Underreported Gap

- The total underreported gap is the sum of the observed underreported gap and unobserved underreported gap.
  - The unobserved underreported gap refers to the unobserved underreported gap not captured by the National Tax Service, which is estimated through sample raw data analysis and econometric model analysis.
  - The observed underreported gap refers to the underreported gap captured by the National Tax Service. Underreporting in the initial tax return that is filed by the statutory due date yet modified by the amended tax return or reassessment by the investigation after the due date.

- The total underreported gap and the underreporting ratio of the aggregate income tax estimated based on the result of sample raw data analysis are 4 trillion 894.5 billion won and 24.45%, respectively, which are higher than the result obtained from the econometric model analysis.
  - The underreporting ratio refers to the ratio of the total underreported gap to the theoretical tax burden.
    - Theoretical tax burden = reported tax liability in initial tax return + Unobserved underreported gap + Observed underreported gap
- Based on the results of the Tobit and Two-part model (Truncated Normal Hurdle Model) analysis, the estimated total underreported gaps are 3 trillion 889.7 billion won and 3 trillion 588.4 billion won, respectively, and total underreporting ratios are 20.46% and 19.13%, respectively.
  - The total underreported gap and the total underreporting ratio estimated based on the result of the Tobit Model analysis are higher than those estimated based on the Truncated Normal Hurdle Model analysis. This is, as explained earlier, because the Tobit Model is more restricted than the Truncated Normal Hurdle Model.

〈Table 3〉 Total underreported gap in aggregate tax income and underreporting ratio of income earned in 2011

(unit: million won, %)

Analysis method	Theoretical tax burden	Reported tax liability in initial tax return	Total underreported gap	Total underreporting ratio
	A=B+C	B	C	C/A×100
Sample raw data analysis	200,652	151,707	48,945	24.45
Tobit Model analysis	190,604	151,707	38,897	20.46
Truncated Normal Hurdle Model analysis	187,591	151,707	35,884	19.13

Source: prepared by the authors

- ● <Table 4> shows the total underreporting ratios that are divided into the unobserved underreporting ratio and the observed underreporting ratio.
- ● The unobserved underreporting ratio of 2011 ranges from 11.74 to 17.53%, depending on the analysis method.
  - The result of sample raw data analysis shows that the observed underreporting ratio is 6.92%, and unobserved underreporting ratio is 17.53%.
  - The results of the Tobit Model shows that observed underreporting ratio is 7.29%, and unobserved underreporting ratio is 13.17%.
  - The results of the Truncated Normal Hurdle Model shows that the observed underreporting ratio is 7.39%, and unobserved underreporting ratio is 11.74%.

<Table 4> Observed underreporting ratio and unobserved underreporting ratio of 2011 by analysis method

(unit: million won, %)

Analysis method	Observed underreporting ratio	Unobserved underreporting ratio	Total underreporting ratio
	A	B	C=A+B
Sample raw data analysis	6.92	17.53	24.45
Tobit Model analysis	7.29	13.17	20.46
Truncated Normal Hurdle Model analysis	7.39	11.74	19.13

Source: prepared by the authors

### 3) Limitations in Underreported Gap Estimation

- ● The most important limitation that this study has is insufficient number of random samples, which prevents a more accurate estimation.
  - Since the number of the sample is too small, the random sample of this study may not represent the population well, and thereby may lead instability in statistics and estimates.
  - In subsequent studies to estimate the underreported gap of the aggregate income tax, it is necessary to gradually increase the number of random samples.
  
- ● This study shows that tax evasion is mainly affected by the ratio of tax base to total income, income amount, and industry-specific characteristics based on econometric model analysis; however, the comparison of the estimation results of sample raw data analysis and the econometric model analysis shows that tax evasion could be affected by other factors than those main factors.
  - This implies that the underreported gap estimated by the econometric model may be biased. Therefore, it is necessary to enhance the explanatory power of a econometric model by developing a better econometric model and constructing available data.
  
- ● In addition, since there is a possibility that when the investigation for taxable income is conducted on random samples, an investigator may not be able to detect tax evasion, so the actual unobserved underreported gap is likely to be larger than the estimate in this study.

### III Estimation of Underground Economy

- In this study, we estimate the scale of Korea's underground economy using data from 1975 through 2015 based on the demand function for cash currency, which is a macroscopic approach.
  - In addition to a demand-for-cash-currency function, other macroscopic approaches include a method based on a function of the demand for electricity and one that uses a Multiple Indicators Multiple Causes model.
  - The other methods can measure the scale of the underground economy relative to a base year; however, they have a limitation in estimating underground economy ratio in absolute terms.
- Based on the demand-for-cash-currency function, the estimation of the scale of the underground economy assumes that the underground economy exists to avoid the tax burden and that all transactions are carried out in cash.
  - The demand for cash currency in the underground economy is defined as the difference between the demand for cash currency in a target year estimated using a demand-for-cash-currency function, in which the tax burden is added to explanatory variables, and the demand for cash currency estimated using the same function yet with an assumption that the tax burden is zero.
  - The scale of the underground economy can be estimated by multiplying the estimated demand for cash currency in the underground economy by the velocity of money.

## 1. Demand-for-Cash-Currency Function and Underground Economy Estimation Method

- The function of the demand for cash currency function is estimated on the assumptions of the log series such as Equation (1) below and the original series such as Equation (2) below.
  - As we will see later, there is a significant difference in the estimation results, because the way other control variables than main variables affect the estimation of underground economy may change, depending on which assumptions we choose between the log series and the original series.
  - $C_t$  refers to cash currency,  $M2_t$  refers to broad money,  $TAX_t$  refers to tax rate variable, and  $X_t$  refers to other control variables in the following equations.
  - The control variables include the ratio of 50,000 won notes to total outstanding cash currencies, real GDP per capita, nominal interest rate, wage relative to national income, etc.

$$\ln\left(\frac{C_t}{M2_t}\right) = \beta \ln(1 + TAX_t) + X_t\gamma + \epsilon_t \quad \text{equation (1)}$$

$$\left(\frac{C_t}{M2_t}\right) = \beta TAX_t + X_t\gamma + \epsilon_t \quad \text{equation (2)}$$

- In previous studies, there are analyses based on a Log-Log type regression equation and those based on a Level-Level type regression equation.
  - Tanzi (1983), Lee (2001), Bae (2005), etc. are the previously studies using the log series.
  - Noh and Yoon (2007), Kang et al. (2015), etc. are the ones using the original series.

- For the estimation of Equations (1) and (2), the OLS estimation method as well as the Prais-Winsten estimation method are used for considering serial correlation problems.
  - To consider the serial correlation problem, Tanzi (1983) estimates a demand-for-cash-currency function by using the Cochrane-Orcutt estimation method.
  - However, the Cochrane-Orcutt estimation method has a problem in that the first observation value disappears.
  - The Prais-Winsten estimation method is the one that redresses the problem arising from the Cochrane-Orcutt estimation method.
  - In addition, the Prais-Winsten estimation method is a special case of the Feasible Generalized Least Squared (FGLS) estimation method, which estimates the scale of the underground economy in the same way as the OLS estimation method once a demand-for-cash-currency function is estimated.
  - However, the Prais-Winsten estimation method is not superior to the OLS estimation method because it should additionally estimate the serial correlation parameter ( $\rho$ ) of residuals.
- From a demand for cash currency function, which is estimated through the method above, the size of the cash currency representing the underground economy can be estimated. The scale of the underground economy can be estimated by multiplying it by the velocity of money.
  - The velocity of money in the underground economy is assumed to be equal to that in the formal economy.
  - Money in circulation in the formal economy is calculated by deducting cash in circulation in the underground economy from the total money supply. The official velocity of money can be calculated by dividing the GDP by money in circulation in the formal economy.

- ● When calculating the velocity of money in the formal economy, estimating the scale of the underground economy may vary depending on how the total money supply is defined.
  - The scale of the underground economy may vary substantially depending on which one of the monetary indicators described below is used for calculating the velocity of money: narrow money (M1) is defined as the sum of cash and demand deposit; or broad money (M2) is defined as narrow money (M1) plus saving deposit of less than two-year tenor and deposit of a similar nature
  
- ● Most studies use narrow money (M1); however, several previous studies used broad money (M2).
  - Tanzi (1983), Kang et al. (2015), etc. assume the total money supply as narrow money (M1).
  - On the contrary, Noh and Yoon (2007) assume the total money supply as broad money (M2).
  
- ● The reason why most previous studies applied the velocity of narrow money (M1) to estimate the underground economy seems to be because they wanted to apply the velocity of money of a nature most similar to cash, basically assuming the means of payment in the underground economy is cash.
  
- ● On the other hand, if the following aspects are taken into consideration, it would be also reasonable to apply the velocity of broad money (M2).
  - Financial markets generally use the velocity of broad money (M2).
  - In the sense that the means of transactions in the underground economy is likely to be cash, the velocity of narrow money (M1) seems to be more appropriate. However, since it is considered that the velocity of money in the underground economy may be lower than that of the formal economy, the velocity of broad money (M2) could be

more appropriate.

- In the underground economy, since there is a restriction that parties engaged in transactions must meet and exchange goods with currency, the velocity of money is likely to be lower than that of the formal economy.

〈Table 6〉 Detailed methodology of underground economy scale estimation

Classification	Dependent variable	Estimation method	Total money supply	Previous study
Method 1	The log series	OLS	Narrow money(M1)	Lee (2001), Bae (2005)
Method 2			Broad money(M2)	
Method 3		Prais-Winsten	Narrow money(M1)	Tanzi(1983)
Method 4			Broad money(M2)	
Method 5	The original series	OLS	Narrow money(M1)	Kang et al. (2015)
Method 6			Broad money(M2)	
Method 7		Prais-Winsten	Narrow money(M1)	
Method 8			Broad money(M2)	

Note: 1) Lee (2001) and Bae (2005) do not define the total money supply.

2) Although Lee (2001) stated that he estimated the cointegration relation, it is unclear whether it was estimated based on an error-correction model, or whether a long-term relation equation was estimated using the OLS method.

3) Bae (2005) does not explain the estimation method in detail.

4) Tanzi (1983) uses the Cochrane-Orcutt method, which is similar to the Prais-Winsten estimation method.

Source: prepared by the authors

- Taking into consideration of these points, we estimate the scale of the underground economy based on the velocity of narrow money (M1), and in a sensitivity analysis, we apply the velocity of broad money (M2) and compare the results with those obtained by using the velocity of narrow money (M1).

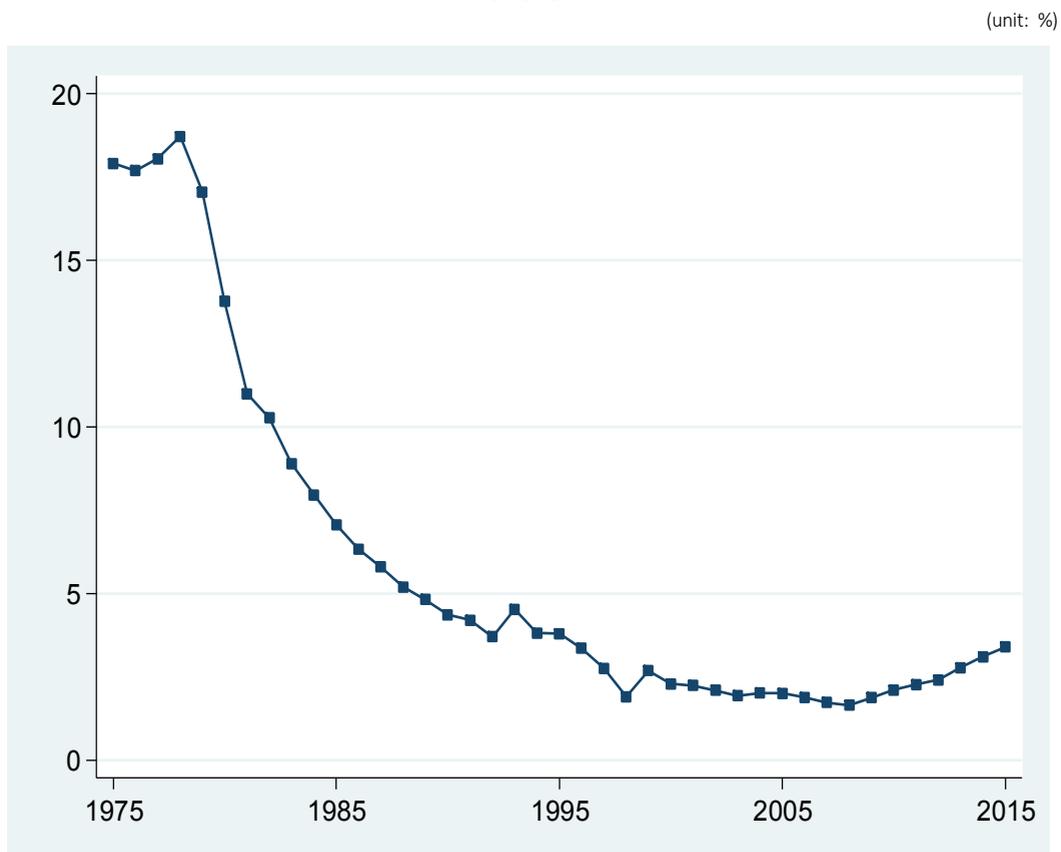
## 2. Choice of Tax Rate Variable

- When the function of a demand for cash currency is estimated, it is very important to determine which indicator is used for tax burden. The following indices could be used as a tax burden variable: the simple average of nominal rates, the highest statutory income tax rate, the tax burden ratio that means total tax revenue divided by GDP, and the total burden ratio that means the sum of total tax revenue and the social security contribution by GDP.
  - The average income tax rate has an advantage as an ex ante tax burden indicator, but is problematic in that a simple average of nominal tax rates lacks a the logical basis, and the tax burden by other tax items is not considered.
  - Although the highest statutory income tax rate is also an ex ante tax burden indicator like the average income tax rate, it has a limitation because it represents only the tax burden on high-income classes while low-income classes also form a part of the underground economy through tax avoidance.
  - Both the tax burden ratio and the total burden ratio have an advantage in that they are tax burden indices covering all tax items; however, they are also problematic in the sense that they are ex post indices—not ex ante indices that can be recognized by taxpayers in advance.
- As a result of the estimation of the function of a demand for cash currency, both the tax burden ratio and the total burden ratio did not show a significant positive (+) relation with the demand for cash currency. So, we did not use those indices for estimating the scale of the underground economy.
  - The ratio of cash currency to broad money (M2), which is a dependent variable in the

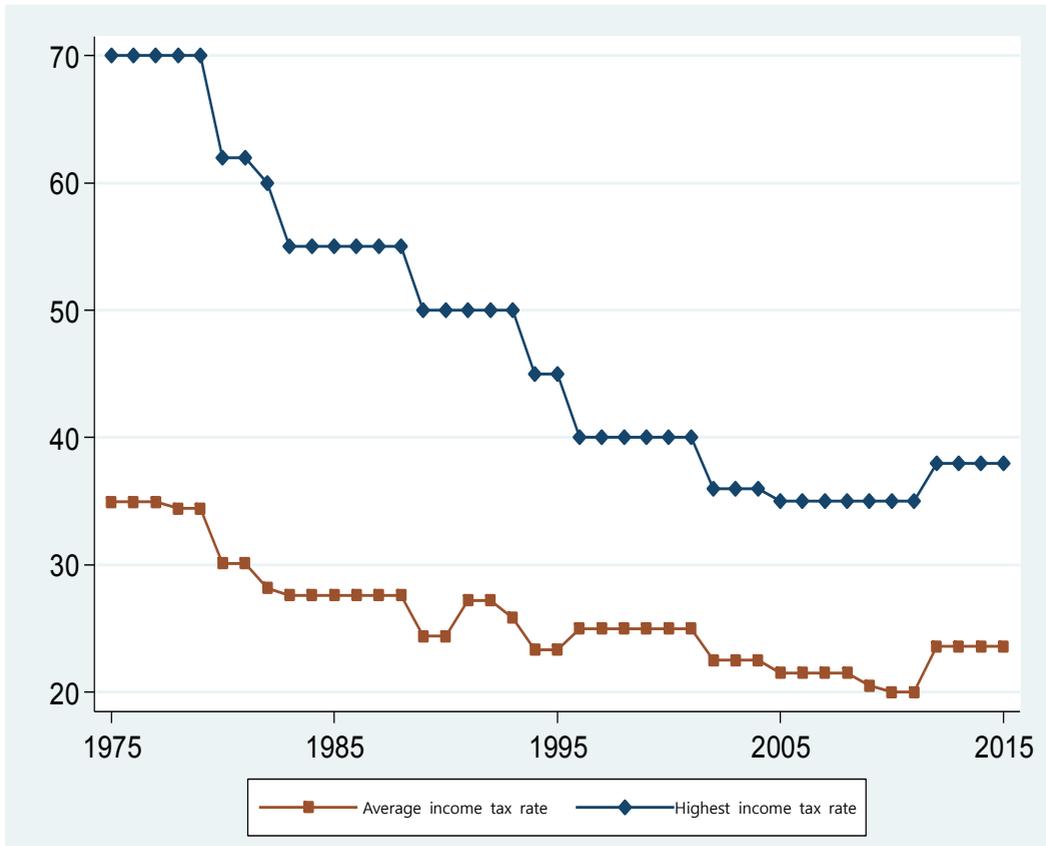
demand-for-cash-currency function, has been steadily declining except for the most recent years (see [Figure 1]).

- Ex ante tax rate variables in [Figure 2] are showing a similar trend to [Figure 1].
- On the other hand, the ex post tax rate variables in [Figure 3] are in an overall upward trend so it is highly possible that they are inversely correlated with a dependent variable in the demand for cash currency function.

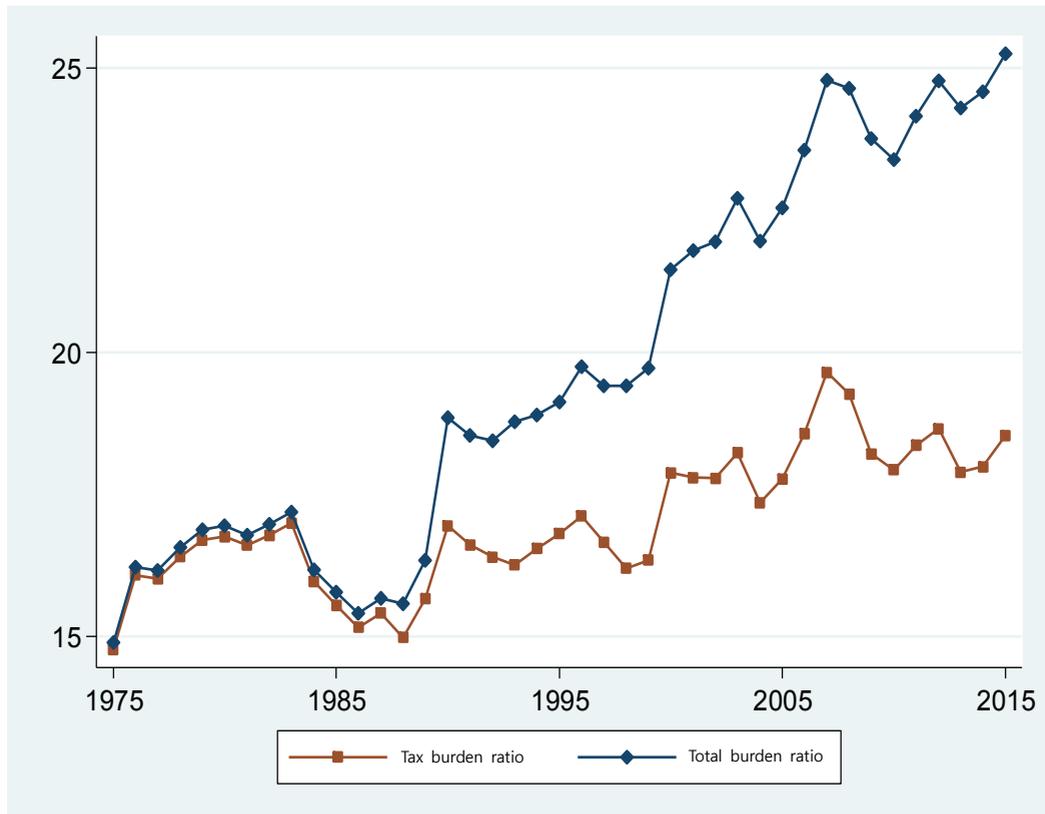
[Figure 1] Trend in (C/M2) as dependent variable of demand-for-cash-currency function



Source: The Bank of Korea, Economic Statistics System (ECOS)



Source: 1. Average income tax rate prepared by the authors using income tax information for each year provided by the National Law Information Center  
 2. The highest income tax rate prepared by the authors using income tax information for each year provided by the National Law Information Center



Source: OECD, Revenue Statistics

- There can be three interpretations of the phenomenon that both the tax burden ratio and the total burden ratio do not show a significant positive (+) relation with the demand for cash currency as follows:
  - The first interpretation is that both the tax burden ratio and the total burden ratio are not appropriate variables for estimating the scale of the underground economy.
  - The second interpretation is that the ratio of cash currency to the total money supply shows a declining trend due to factors other than tax; however, the estimated demand for cash currency function model used in this study does not explain the influence of those other factors.

- In other words, if the factors other than tax were able to sufficiently explain the decline in demand for cash currency, there would be a positive (+) correlation between the tax burden and cash currency.
- The third interpretation is that although the tax burden ratio or the total burden ratio is an appropriate variable for estimation of the scale of the underground economy, it cannot be assumed as a statistically significant value because the scale of the underground economy caused by the tax burden is very small.
- ● Among those three interpretations, the third one is almost unlikely while it is considered that both the first and second interpretations are likely. It is highly possible that the results mentioned above are obtained because these two work hand in hand.
- ● On the other hand, although the tax burden ratio or the total burden ratio is not an appropriate variable for estimation of the scale of the underground economy, it has advantages in comparison with other variables such as average income tax rate.
  - Average income tax rate has an advantage that it is an ex ante tax burden indicator, but also has a limitation that it can only reflect the changes in tax rate, but not changes in tax burden caused by changes in tax allowances and tax brackets.
    - Another problem is that the average income tax is calculated by simply averaging progressive tax rates.
  - On the other hand, the tax burden ratio or the total burden ratio has a disadvantage as an ex post tax burden indicator; however, it also has an advantage that it reflects changes in various taxation systems.
- ● Therefore, it is hard to say which one is overwhelmingly superior, but empirically, the average income tax rate shows a statistically significant positive (+) correlation with demand for cash currency, which is stable regardless of changes in model.

- However, the positive (+) correlation between the average income tax rate and the demand for cash currency does not mean that the average income tax rate is a more appropriate indicator.
  - As in the second interpretation above, we cannot rule out the possibility that such factors other than tax may not adequately explain changes in the demand for cash currency caused thereby.
  - In this case, if an analysis using the same model shows a positive (+) correlation between the average income tax rate and the demand for cash currency, then it may be interpreted as an accidental statistical correlation rather than an explanation based on a causal relationship in which the average income tax rate affects the demand for cash currency.
  
- In conclusion, under the currently used demand-for-cash-currency model, it is difficult to estimate the scale of the underground economy using the tax burden ratio or the total burden ratio.
  - It is because it cannot be statistically proved that an increase in the tax burden increases the demand for cash currency.
  - We speculate that this is because the two indices contain elements that are inappropriate as tax burden indices and also because other factors affecting the demand for cash currency than the tax burden have not been adequately controlled for.

### 3. Estimation Results and Implications

- If we estimate the function of demand for cash currency function using the average income tax rate as a tax burden indicator and estimate the scale of the underground economy based thereon, there is a significant difference in the ratio of the underground economy to the formal economy, depending on the type of dependent variable used in estimating the function of demand for cash currency.

- ● The scale of the current underground economy becomes significantly underestimated in the case of the log series-based estimation that uses the logarithm of the ratio of cash currency to the total money supply as a dependent variable than in the case of the original series-based estimation that uses the ratio of cash currency to the total money supply as a dependent variable.
  - In 2015, the ratio of the underground economy to the formal economy is 5.33% to 7.96% for the log series. The figure becomes five to six times higher for the original series.
  
- ● If we make a compare between the estimated scale of the underground economy using OLS without taking into consideration of the autocorrelation of the time series and the estimated scale of the underground economy using the Prais-Winsten method while taking into consideration of the autocorrelation of the time series, the estimated scale of the latter is smaller.
  - In the case of the log series, the OLS estimate ranges from 6.65% to 7.96% and the Prais-Winsten estimate ranges from 5.33% to 6.43%.
  - In the case of the original series, the OLS estimate is five or six times higher the log series estimate, and the Prais-Winsten estimate is about four times higher the log series estimate in 2015.
  
- ● Estimation models have advantages and disadvantages. We find that the log series-based estimation is more appropriate than the original series-based estimation when comprehensively considering explanatory power, significance of coefficient estimate, a concordance with theoretical prediction, etc.
  - The explanatory power of a demand-for-cash-currency function is higher in the case of the log series.
    - According to estimation results of a demand-for-cash-currency function using the log series, the coefficient of determination ( $R^2$ ) using the OLS estimation ranges from 0.97 to 0.99,

- and the coefficient of determination using the Prais-Winsten estimation ranges from 0.90 to 0.93.
- On the other hand, in the case of the original series, the coefficient of determination using the OLS estimation ranges from 0.93 to 0.97, which seems to be high, whereas the coefficient of determination using the Prais-Winsten estimation with serial correlation removed is as low as 0.67 to 0.81.
  - Even when a comparison is made on estimates of correlation coefficients of individual explanatory variables, the signs of coefficients in the case of the log series are more consistent with intuition, and statistical significances are also high.
- According to the sensitivity analysis, the scale of the underground economy is smaller when narrow money (M1) is used as the total money stock indicator than when broad money (M2) is used.
    - In the case of the log series, if narrow money (M1) is used as the total money stock with dummy variables excluded, the scale of the underground economy in 2015 is estimated to be 7.96% of the formal economy; if broad money (M2) is used as the total money stock, the estimate goes down to 2.38%.
  - If the dependent variable of the demand for cash currency is assumed to be the log series and the total money supply is defined as narrow money (M1), the result indicates that the ratio of the underground economy to the formal economy has been generally declining since 1975.
    - This change is consistent even if the total money supply is assumed to be broad money (M2).
  - In particular, the ratio of the underground economy to the formal economy has been gradually decreasing from 8.72%, 8.49% to 7.96%, respectively, from 2013 through 2015.

- This is the result of analysis where the demand-for-cash-currency function is estimated using OLS without dummy variables such as foreign exchange risks.
  - In the sensitivity analysis, if the effect of the introduction of 50,000 won notes is reflected under the same conditions, the ratios of the underground economy to the formal economy in the most recent three years post 10.08%, 9.88% and 9.30%, respectively, which are somewhat higher than the above figures.
    - As in the case of the tax burden variables, these are estimates on the scale of the underground economy with a premise that additional demand for cash currency by the introduction of 50,000 won notes was caused by the underground economy.
  - If the ratio of the underground economy to the formal economy in 2015—i.e. 7.96% (9.30%)—is converted to an amount, it would amount to about 124 trillion won (145 trillion won).
- ● Also when the scale of the underground economy is estimated using the highest income tax rate—instead of the average income tax rate—as the tax burden indicator, the ratio of the underground economy to the formal economy is estimated to vary significantly depending on the type of dependent variable used in estimating the demand-for-cash-currency function.
  - If the dependent variable is assumed to be the log series, the ratios of the underground economy for 2013 through 2015 are estimated to be 10.06%, 9.83%, and 9.28%.
    - This is the case where the demand-for-cash-currency function is estimated using OLS without dummy variables such as foreign exchange risks.
  - On the other hand, if the dependent variable is assumed to be the original series, the ratios of the underground economy for the same period are about four times higher, compared to the case where the dependent variable is assumed to be the log series.
  - As in the case of the average income tax rate, the estimation of the demand-for-cash-currency function is estimated using the highest income tax rate suggests that the log series fares better in terms of the explanatory power of a model, statistical significance of coefficient estimates, and concordance with theoretical

predictions than the original series. Therefore, it is judged that estimation on the assumption of the log series is more appropriate.

- In spite of these estimation results, the underground economy estimation result based on the demand-for-cash-currency function varies depending on the estimation method, the choice of variables, and the method of application of variables. Therefore, the scale of Korea's underground economy cannot be determined based only on existing research results including this study.
  - As shown in this study, there is a room for controversy regarding such key variables as tax burden and money supply, and the estimation result may vary considerably depending on which one is used.
  - The scale of the underground economy varies widely depending on whether the estimation model uses the original series or the log series.
  - In addition, there is doubt that the widely used current model sufficiently explains changes in the demand for cash currency generated by factors other than tax burden.
    - In particular, in Korea, a rapid decline of the ratio of the demand for cash currency to the total money supply has been seen in the course of rapid economic development. Yet, it is not clear that the reason behind the phenomenon is sufficiently explained by this model.

## References

- Kang Moon-soo, Noh Gi-seong, Kwon Seon-joo, Bae Deuk-jong, Seong Myeong Jae, and Jeon Byeong Mok, "Studies on Estimation Methods of the Shadow Economy and International Comparison of Estimates of the Shadow Economy," Future Society Joint Research Series 15-15-01, KDI School of Public Policy and Management / National Research Council for Economics, Humanities and Social Sciences, 2015.
- Noh Gi-seong, and Yoon Yeo-pil, "IT Advancement and Tax Base Broadening," Policy Research Series 2007-11 KDI, 2007.
- Bae Min-geun, "Underground Economy is Declining in Korea," LG Weekly Economy Report, LG Economic Research Institute, 12/28/2005 pp. 34-38.
- An Jongseok, Kim Hak-soo, Park Myeon-ho, Kang Sung Hoon, Shin Sanghwa, and Yi Dong Gyu, "Estimation of Tax Gap by Tax Item and Estimation of Underground Economy Scale Related to Tax," Korea Institute of Public Finance, 2016.
- Lee Woo-seong, "How Big is Underground Economy in Korea?," LG Weekly Economy Report, LG Economic Research Institute, 4/25/2001, pp. 4-25.
- Tanzi, Vito, "The underground economy in the United States: annual estimates, 1930-80," Staff Papers 30 (2), 1983, pp. 283-305.

## Websites

- National Law Information Center, <http://www.law.go.kr/main.html>, Accessed 12/15/2016
- Bank of Korea, Economic Statistics System(ECOS), <http://ecos.bok.or.kr/>, Accessed 11/21/2016
- Statistics Korea, Korean Statistical Information Service, <http://kosis.kr/>, Accessed 11/21/2016
- OECD, Revenue Statistics, <http://stats.oecd.org/>, Accessed 12/13/2016



---

## Estimation of Income Tax Gap and Underground Economy in Korea

2017. 12. 13(vol. 4)

---

- Publisher : Park, Hyung-Soo
  - Issued by : Korea Institute of Public Finance
  - Address : 336, Sicheong-daero, Sejong-si, Korea 31047
  - Phone Number : 82-44-414-2123
  - Fax Number : 82-44-414-2129
  - Designs & Printing : AMIGO DESIGN
-

