Int. J. Advance Soft Compu. Appl, Vol. 9, No. 1, March 2017
ISSN 2074-8523

Grey Exponential Smoothing for Forecasting Indonesian Income Tax

Arum Handini Primandari[1], Tuti Purwaningsih[1], Randi Putra Oetama[1]

[1]Islamic University of Indonesia,
e-mail: primandari.arum@uii.ac.id, tuti.purwaningsih@uii.ac.id,
randiputraoetama@gmail.com

Abstract

Income tax is one of Indonesian financing sources to fund its expenditure. It is used to defray national development for improving public welfare. Therefore, the government sets the target for tax revenue so that it can be managed effectively. Predicting the target would help the government to thoroughly prepare the tax management. Because of either trend or seasonal pattern on its data, the basic method commonly used is holt-winter exponential smoothing. However, due to various schedule of tax payment, data have random shock. We employed grey method which was applied on exponential smoothing to reduce this random, hence we have smoothed data. In optimizing the parameters of the exponential smoothing, we used levenberg-marquardt algorithm. The result shows us that grey-holt exponential smoothing method is not always outperforming basic holt-winter exponential smoothing for forecasting Indonesian income tax.

Keywords: *forecasting, exponential smoothing, levenberg-marquardt, income tax.*

1 Introduction

As it is already known that tax is an obligatory contribution to state imposed upon both individuals and legal entity. Taxation is regulated by law and is used to fund the government. The money provided by taxation has been used to defray national development, such as roads, health, education, public services, etc.

There are several types of taxes in Indonesia; one of them is income tax. The government through the Directorate General of Taxation will set the amount of income tax revenue targets. This target is used to make the tax allocation plan so the tax is managed effectively. Thus, tax target value is a crucial factor to predict taxation.

The income tax data we used have both slight trend and seasonal pattern. It can be seen from its graphics. Moreover, there are some shocks in the specific data due to payment schedule.

Holt exponential smoothing (HES) has high accuracy prediction on the data with trend. While holt-winter exponential smoothing (HW) works well on both seasonal and trend data. Over the time, there are a lot of researches on exponential smoothing to improve its ability in forecasting [3][4].

Since it seems that there are more series which have multiplicative trend than additive trend, some researches tend to develop this scope. The parameter to dampening the multiplicative trend added to exponential smoothing improves the forecast accuracy [2][6]. Furthermore, to deal with the random interference on data, grey method generates an operator to smooth it. It has been proved that applying grey method on double exponential smoothing could outperform the traditional one [7].

In this research, we apply grey method on both HES and HW for forecasting the income tax. The Levenberg-Marquardt optimization is used to optimize the parameters on those exponential smoothing. By comparing the Mean Squared Error (MSE), we gain the best method to predict Indonesian income tax.

2 Problem Formulations

Since the payment of tax is already scheduled, the tax revenue will be significantly increased on the due date. However, because of some reasons, the revenue will be unpredictable. These patterns are the characteristic of Indonesian income tax.

The traditional exponential smoothing with double smoothing parameters is suitable to predict the data with trend which either has multiplicative or additive trend. The optimal value of parameters will give us the best forecasting since it minimizes error. Beside the parameters itself, prediction accuracy depends on how we take the initial value.

Data with random interference need a specific treatment. The sudden move which is happened on the data sometimes causes a problem. Therefore, we need to handle this problem first.

In the tax data which have both seasonal and trend pattern, we employ a method named HW. However, because of the random trend which is occurred in some data, apart from the peak of the seasonal data, we also need another method that can solve this problem.

One of the methods which provide an operator that can smooth the random interference is grey method. We apply this method on exponential smoothing method so it can elevate the precision of forecasting.

3 Method

3.1 Grey Exponential Smoothing

HES is method for handling linear trend. The most prominent advantage of this method is to vary the deterministic trend on time series data [3]. It has two basic smoothing formulas for level term and growth term, which are

Level : $\ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} - b_{t-1})$

Growth : $b_t = \beta(\ell_t - \ell_{t-1}) + (1-\beta)b_{t-1}$

The prediction formula defines as follow:

Forecast : $\hat{y}_{t+h} = \ell_t + b_t h$ (1)

Where α and β (with value between 0 and 1) are respectively for the level smoothing parameter and growth smoothing parameter, *h* is the number of prediction period.

Adding the seasonal parameter, HW has three parameters for smoothing. The formula of multiplicative HW method is expressed as:

Level : $\ell_t = \alpha \dfrac{y_t}{s_{t-m}} + (1-\alpha)(\ell_{t-1} + b_{t-1})$

Growth : $b_t = \beta(\ell_t - \ell_{t-1}) + (1-\beta)b_{t-1}$

Seasonal : $s_t = \gamma \dfrac{y_t}{\ell_{t-1} - b_{t-1}} + (1-\gamma)s_{t-m}$

The prediction formula defines as follow:

Forecast : $\hat{y}_{t+h} = (\ell_t + b_t h)s_{t-m+h}$ (2)

Where *m* is the length of seasonality (e.g. the number of months or quarters in a year).

Define 3.1 [7] For the original time series $Y^{(0)} = \{y^{(0)}(1), y^{(0)}(2), ..., y^{(0)}(n)\}$, an r-order accumulated generating operator (AGO) sequence $Y^{(r)} = \{y^{(r)}(1), y^{(r)}(2), ..., y^{(r)}(n)\}$, $r \in R_+$ can be generated by r-AGO as follows:

$$y^{(r)}(k) = \sum_{i=1}^{k} \binom{k-i+r-1}{k-i} y^{(0)}(i); k = 1, 2, 3, ..., n, \tag{3}$$

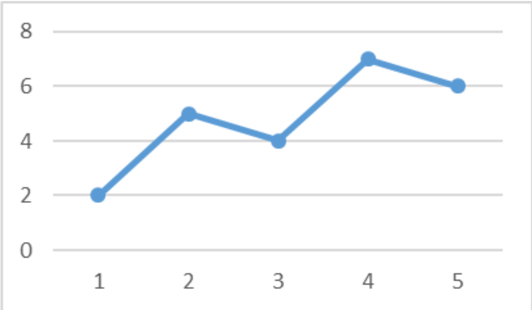
set $\binom{r-1}{0} = 1, \binom{k-1}{k} = 0, \binom{k-i+r-1}{k-i} = \dfrac{(k-i+r-1)(k-i+r-2)...(r+1)r}{(k-i)!}$.

The inverse of *r*-order of AGO is called inverse accumulated generating operator (IAGO). When we take 0<*r*<1, the r-order of IAGO of $Y^{(r)}$ is expressed as follows:

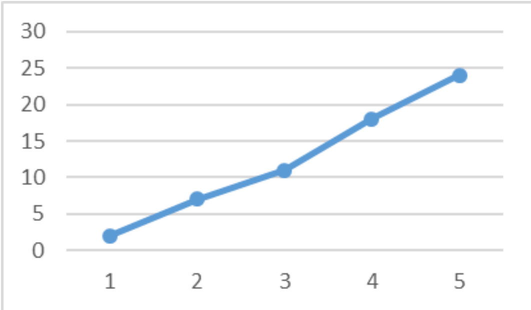
$$Y^{(-r)} = {}^{(1)}Y^{(1-r)} = \left\{ {}^{(1)}y^{(1-r)}(1), {}^{(1)}y^{(1-r)}(2), \dots, {}^{(1)}y^{(1-r)}(n) \right\} \tag{4}$$

where $\lceil r \rceil = \min\left\{ n \in Z \middle| r \le n \right\}$, ${}^{(1)}y^{(1-r)}(k) = y^{(1-r)}(k) - y^{(1-r)}(k-1)$.

As an example, given the sequence $Y^{(0)} = \{2, 5, 4, 7, 6\}$. We take the increasing order $r = 0.4$. By definition 3.1., we obtain the sequence $Y^{(0.4)} = \{2, 5.8, 6.56, 10.45, 11.42\}$.



(a)



(b)

Fig. 1: (a) The original sequence; (b) The sequence by 0.4-AGO

The process of calculating grey method on exponential smoothing is defined as follow:

Step 1 : Set the number of *r* and obtain the r-AGO sequence using equation (3).

Step 2 : Calculate the optimal parameter for exponential smoothing.

Step 3 : Compute the predictive value using equation (1) or (2).

Step 4 : Transform the prediction value to the original sequence by means of IAGO using equation (4)

3.2 Levenberg-Marquardt

In order to fitting a function $\hat{y}(t; p)$ of an independent variable *t* and a vector of *n* parameters *p* to set of m data points (t_i, y_i), it is convenient to minimize the weighted squared of residual [1]. The residual define as follows:

$$\chi^2(p) = \sum_{i=1}^{m} \left[\frac{y(t_i) - \hat{y}(t_i; p)}{w_i} \right]^2$$
$$= y^T W y - 2 y^T W \hat{y} + \hat{y}^T W \hat{y}$$

where the value of w is a measure of the error in measurement $y(t_i)$. The weighting matrix W is diagonal with $W_{ii} = 1/w_i^2$.

The Levenberg-Marquardt algorithm parameter updates define as follows [1]:

$$\left[J^T WJ + \lambda diag\left(J^T WJ\right)\right] h_{lm} = J^T W\left(y - \hat{y}\right),$$

where h is the parameter update, J is the jacobian matrix $\left[\partial \hat{y} / \partial p\right]$.

3 Numerical Result

Income tax is one of the most contributing taxes to tax revenue levied by General of Taxation Directorate of Indonesia. There are several kinds of income taxes; the top three income taxes which give the greatest contribution are 21st article of income tax, 25th article of corporate income tax, 25th article of individual income tax. The assessable incomes include salaries, honoraria, allowances, and other payments. While taxpayers have more than earning source, they pay installment written on article 25.

The data are monthly income tax from 2013 to 2015. Since monthly data in 2016 has not been published, the recent data we could manage to obtain is data in December 2015. We transformed the data using *ln* to shrink the scale because the unit data amount is in billion.

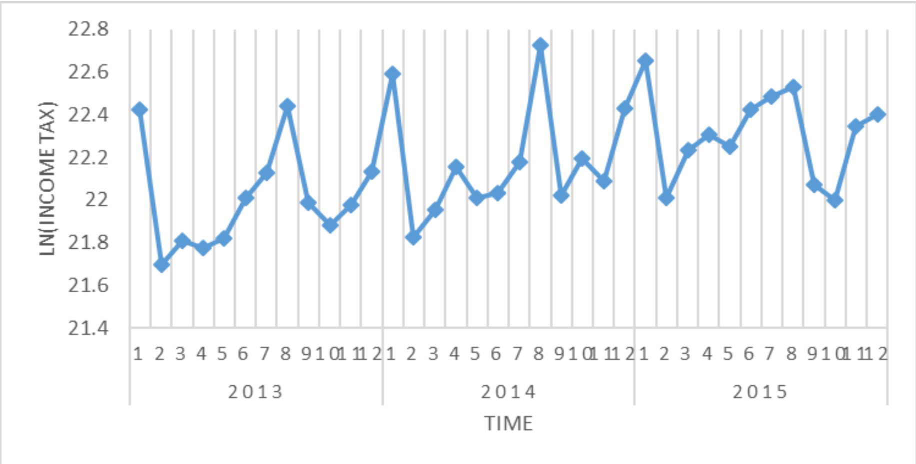


Fig. 2: The ln transformed of 21st article income tax

The data in Fig.2., in every year, have their peak in the first and the eighth month. Therefore, we considered it as seasonal data. Even though in others months, the cycle was not always the same, for example the fourth month in 2014 and 2015 became local peak, but in 2013 it did not.

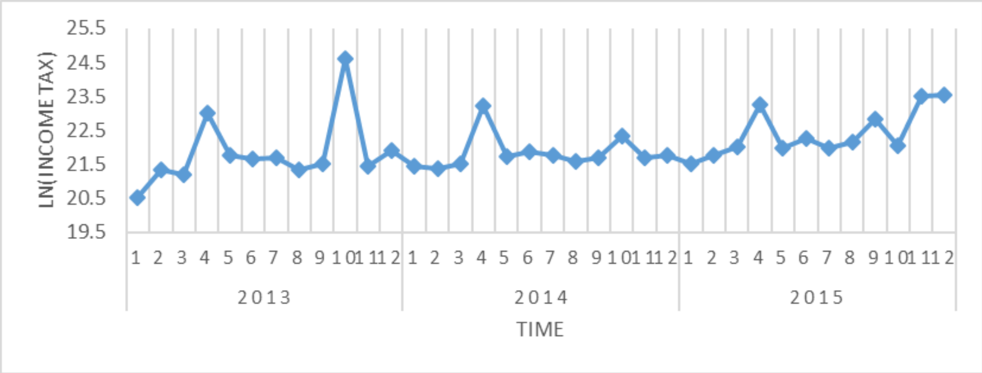


Fig. 3: The ln transformed of 25th article corporate income tax

The data in Fig. 3. has different highest peak in every year. We did not acknowledge it to has seasonal pattern. Althogh in 2013 and 2014 the spike occurred at fourth and tenth, but in 2012 it did not.

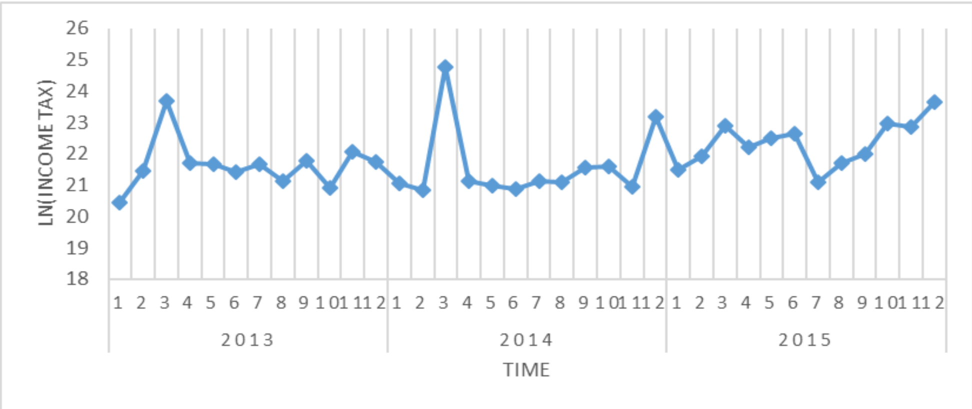


Fig. 4: The ln transformed of 25th article of individuals income tax

Comparing to others income tax data, the 25th article individual income tax is less random. It only has a slight trend, both upward and downward.

3.1 Analysis Result

We made comparison of grey method applied in exponential smoothing and basic exponential smoothing. There are three cases according to the data.

Case 1: 21st article income tax

Take the series in Fig.1., we analyze it using both HW and GHW with levenberg-marquardt optimization. We used r = 0.01 as increasing factor to GHW. The result is showed in Table 1.

Table 1: The result of 21st article income tax

Method	α	β	γ	MSE
HW	0.99767537	1.53e-05	0.033502	0.027763

GHW	0.99575795	2.55E-14	0.184287	0.108152

Case 2: 25th article corporate income tax

We worked both HES and GHE to data in Fig. 2. Since it showed low growth, we used r = 0.001 to GHE. The result is presented bellow.

Table 2: The result of 25th article corporate income tax

Method	α	β	MSE
HES	0.2875	0.5565	0.914669
GHE	0.2924	0.5538	0.888442

Case 3: 25th article of individuals income tax

In the third case, we take r = 0.0011 as the consequence of slight increment. The result of HES and GHE are presented in Table 3.

Table 3: The result of 25th article of individuals income tax

Method	α	β	MSE
HES	0.5028	0.3103	1.47466
GHE	0.5035	0.3167	1.48157

3.2 Forecasting

Since the data is not up to date, we left the method as proposed method for taxation forecasting. The forecasting as shown in the next three tables was the simultaneously result of Table 1., Table 2., and Table 3.

Table 4: The fitted values 21st article income tax

	Actual data	HW	GHW
August	22.528	22.771	22.938
September	22.071	22.025	22.063
October	22.002	21.945	22.134
November	22.347	22.041	22.074
December	22.403	22.483	22.644

Table 5: The fitted values 25th article corporate income tax

	Actual data	HES	GHE
August	22.148	22.449	22.434
September	22.821	22.370	22.352
October	22.039	22.573	22.566
November	23.499	22.417	22.405
December	23.556	22.286	22.894

Table 6: The fitted values 25th article of individuals income tax

	Actual data	HES	GHE

August	21.692	21.779	21.771
September	21.986	21.575	21.566
October	22.972	21.686	21.679
November	22.867	22.438	22.438
December	23.638	22.826	22.830

Using *exp()* to transforming back, we would have the original data. For example, the forecasting of December 2015 of 21st article income tax is IDR 5,812,635,968 by HW. The graphs of the three cases fitted values are presented in the appendix.

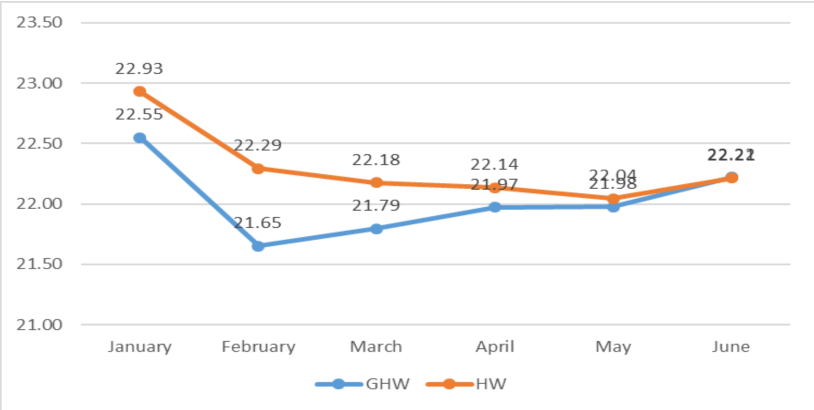


Fig. 5: The Forecasting values of 21st article income tax

The forecasting value of 21st article income tax can be seen in Fig. 5. According to Fig. 5., traditional HW always have higher value of forecasting than GHW. This is a good sign because Holt-Winter method always encounter overforcast [7].

4 Conclusion

If 0<*r*<1, then *r* becomes increasing order to smooth linear trend. It provides more weight to the recent changes. However, not all cases have the same result which is GHES outperforms HES method. Even though employs *r* as the increasing order, in some cases it does not significantly improve the accuracy of prediction. Moreover, applying grey method on both holt and holt-winter exponential smoothing does not always show more precise forecasting.

5 Open Problem

Besides optimizing the parameter of exponential smoothing, we also need to optimize the value of *r* so that we can obtain the smallest error. We can employ non-linear optimization method, such as Levenberg-Marquardt. Moreover, studying multiple seasonal allows us to analyze the taxation data better since the payment period has been arranged.

ACKNOWLEDGEMENTS.
Academic Development Directorate of Islamic University of Indonesia (BPA UII) for funding the conference.

References

[1] Gavin, Henri P. 2015. The Levenberg-Marquardt Method for Nonlinier Least Square Curve-Fitting Problems. *Structural Dynamics and the Seismic Response Control Laboratory: Experimental System.* September 29. Accessed June 9, 2016. http://people.duke.edu/~hpgavin/ce281/.

[2] Hyndman, Rob J, Anne B. Koehler, J. Keith Ord, and Ralph D. Snyder. 2008. *Forecasting with Exponential Smoothing: The State Approach,* Springer, Germany.

[3] Jalil, Nur Adilah Abd, Maizah Hura Ahmad, and Norizan Mohamed. 2013. Electricity Load Demand Forecasting Using Exponential Smoothing Methods. *World Applied Science Journal*, Vol. 22, No. 11, 1540-1543.

[4] Li, Xiaochen. 2013. Comparison and Analysis between Holt Exponential Smoothing and Brown Exponential Smoothing Used for Freight Turnover Forecast. *Third International Conference on Intelligent System Design and Engineering Applications*, IEEE, 453-456.

[5] Taylor, James W. 2012. Density Forecasting on Intraday Call Center Arrivals Using Models Based on Exponential Smoothing. *Management Science*, Vol. 53, No. 3, 534-549.

[6] Taylor, James W. 2003. The Exponential Smoothing with a Damped Multiplicative Trend. *International Journal of Forecasting*, Vol. 19, 715-725.

[7] Wu, Lifeng, Sifeng Liu, and Yingjie Yang. 2016. Grey Double Exponential Smoothing Model and Its Application on Pig Price Forecasting China. *Applied Soft Computing*, Elsevier, Vol. 39, 117-123.

Appendix

Case 1: 21st article income tax

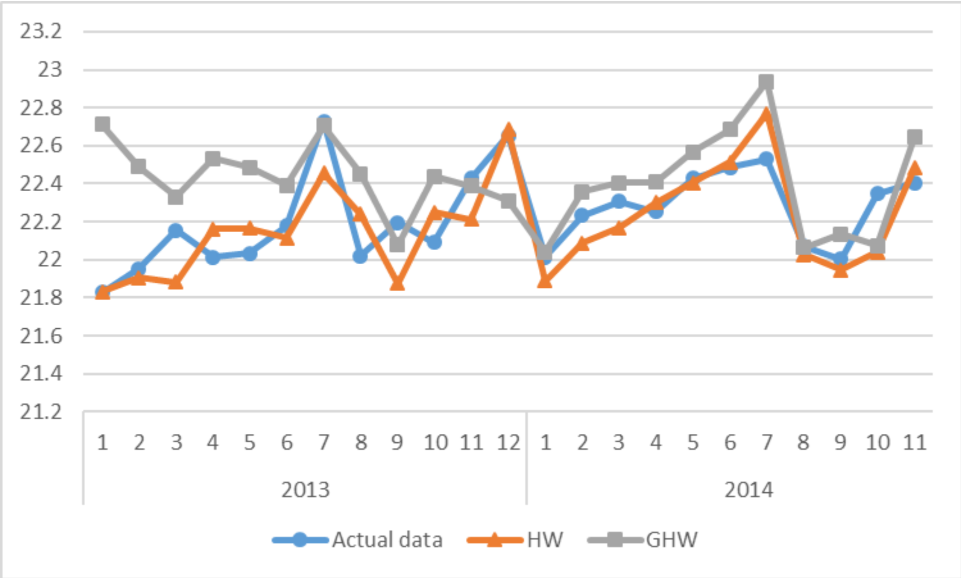


Fig. 6: The fitting graphs of HW and GHW

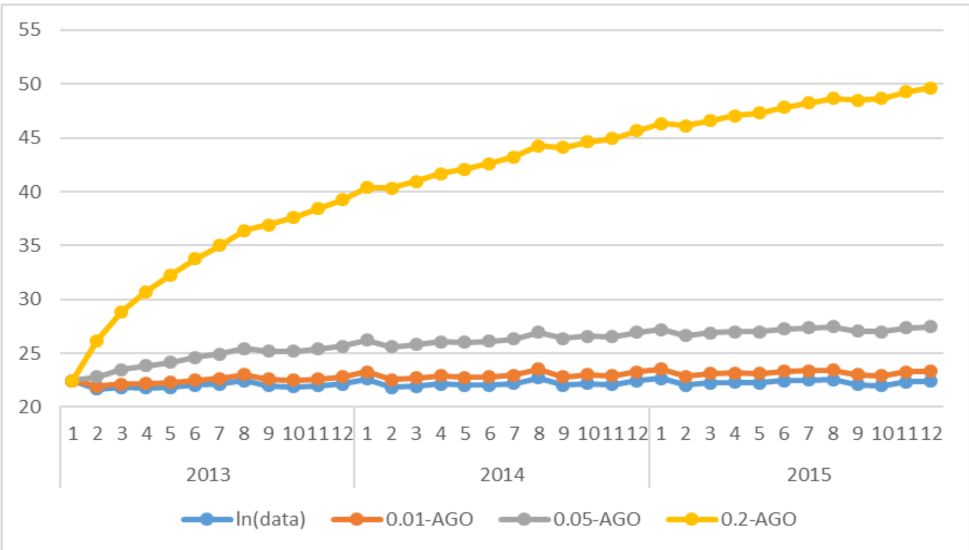


Fig 7: The Comparison of Different Value of *r*

Case 2: 25th article corporate income tax

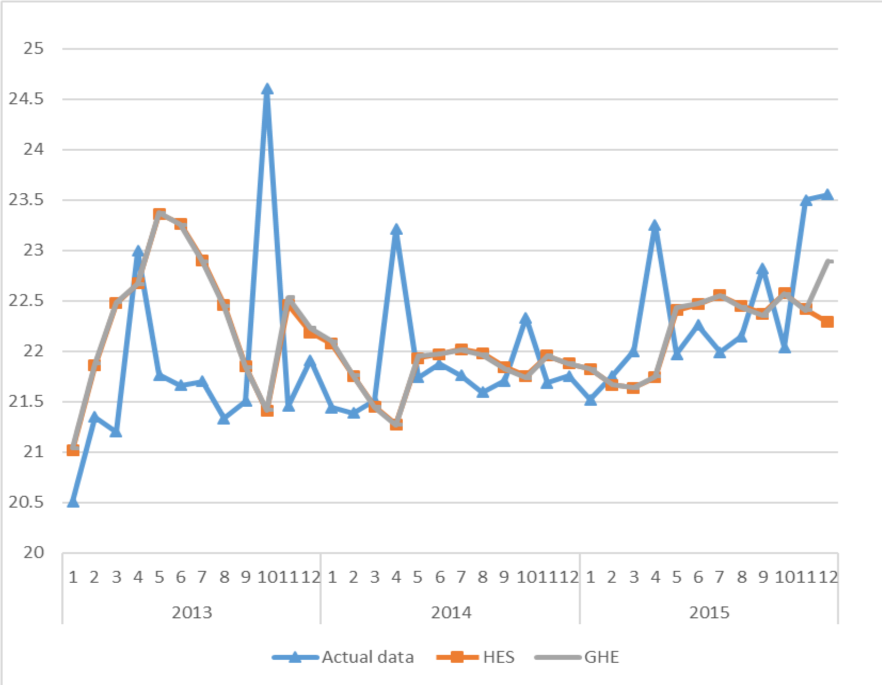


Fig. 8: The fitting graphs of HES and GHE

Case 3: 25th article of individuals income tax

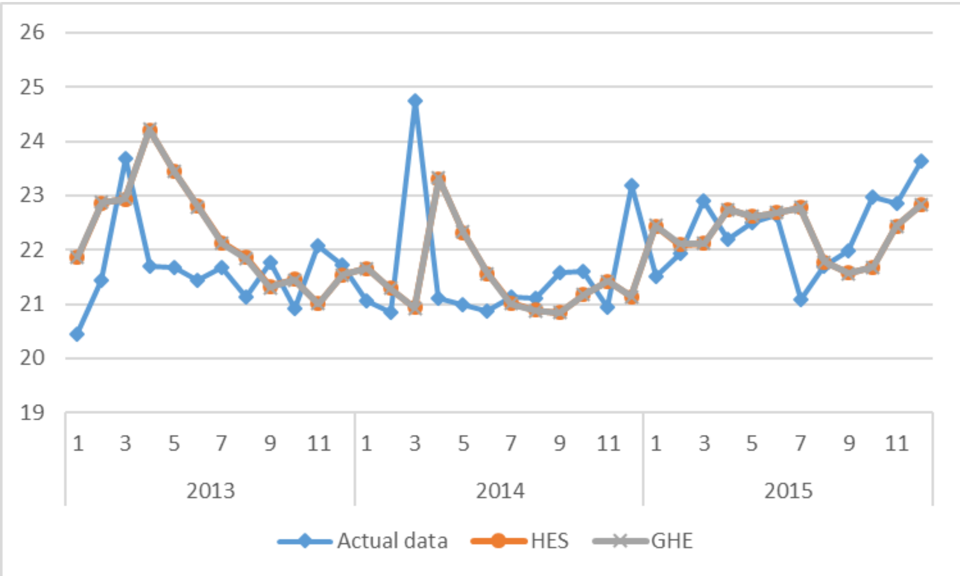


Fig. 9: The fitting graphs of HES and GHE