Mathematical modeling for awareness, knowledge, and perception that influence willpower to pay tax using multiple regression

Achi Rinaldi¹, Yulistia Devi¹

¹Universitas Islam Negeri Raden Intan Lampung, Indonesia

Corresponding author: achi@radenintan.ac.id

Abstract. One effort to increase revenue is to increase the source of funds obtained from within the country, namely taxes. The taxpayer's willingness in their tax payments is essential to optimize tax revenue, so it needs a particular study on this matter. This study discusses the factors that influence the willingness to pay taxes for private taxpayers using multiple regression. The model of willingness to pay taxes through 3 variables, namely: awareness, knowledge, and perception of taxpayers. The study location in Lampung Province was divided into five regions based on the tax office that sheltered it. The samples were taken 100 respondents through a questionnaire with a stratified random sampling technique. The regression model in this study gives significant results for all independent variable coefficients. The variable that has the most significant effect on the willingness to pay taxes is the awareness variable with p-value < 0.05. Stakeholders have to improve awareness, knowledge, and perception to increase state tax revenue.

Keywords: awareness, knowledge, multiple regression, perception, willingness to pay taxes.

1. Introduction

Continuously, the government aims to improve the welfare of the people, both material and spiritual. To realize these objectives requires a large, enough development budget. One effort to achieve increased revenue for development is to explore the source of funds from within the country, namely taxes. Economically, tax collection is state revenue that is used to improve people's lives [1].

The development process carried out smoothly and continuously; a balanced relationship between the state budget and revenue is needed dynamically and proportionally. A tax which is one of the state revenues of the State Budget (APBN), has an essential role in supporting the administration of the state. Based on BPS's data on the realization of state revenues in 2007-2017, the tax sector's total APBN revenue plays a dominant role than non-tax revenue sources. This can be seen from the percentage of tax revenue from 2007 to 2017, which is always above 68%, while non-tax revenue is not still more than 32% of total state revenue. This result shows that revenue from the tax sector has become the mainstay in financing development in Indonesia.

Taxes are people's contributions to the state treasury based on the law (which can be forced) by not getting service fees (contra), which are directly indicated and used to pay public expenses [2]. In some developed countries that impose taxpayers, citizens benefit from the State, such as benefits for the unemployed, free health benefits, free primary education, convenient transportation, etc. These direct or indirect benefits encourage taxpayers to fulfill their obligations to pay taxes with the full awareness that
they will get compensation through facilities that have been designed by the government. Automatically the desire to break the obligation to pay taxes will be eroded. But there are still many taxpayers who do not have the awareness or their willingness to pay taxes.

The willingness of taxpayers to pay their tax obligations is vital to optimizing tax revenue. The willingness to pay tax can be interpreted as a value that is willing to be contributed by someone (as determined by the regulations) used to finance general state expenditure by not getting direct services (contra) [3]. According to the research of Widayati and Nurlis, some factors influence the taxpayer's willingness to pay their tax obligations, namely the awareness factor to pay taxes; good perception of the effectiveness of the taxation system; knowledge and understanding of tax regulations [4].

The Taxpayer, especially personal taxpayers who are freelancers, must-own willingness to pay taxes. Since the issuance of Law Number 6 of 1983, which was later amended by Law No. 28 of 2007 concerning General Provisions and Tax Procedures (KUP), is known as the Self Assessment System that gives taxpayers the confidence to calculate, pay, and report the tax payable themselves. With the Self Assessment System implemented, in addition to awareness and honesty of taxpayers, adequate tax knowledge and technical understanding also play an essential role, so that taxpayers can carry out their tax obligations correctly and adequately.

This research discusses the factors influencing the willingness to pay taxes for personal taxpayers who are freelancers. Some of the studies that underlie this research are [3–10]. Their study uses three variables in their study, namely Tax Paying Awareness, Knowledge, and Understanding of Tax Regulations, and Perceptions for the Effectiveness of the Taxation System. This research conduct in Bandar Lampung and case studies on KPP Pratama Kedaton. There are differences in the results of research conducted by previous studies, such as the results of the study of Widayati and Nurlis namely awareness of paying taxes and a good perception of the effectiveness of the taxation system that has no significant effect on the willingness to pay taxes [4]. While the results of research Hardiningisih and Yulianawati, namely, awareness of paying taxes has a positive effect on the willingness to pay taxes. Variable knowledge, understanding, and perception of the taxation system's effectiveness do not affect the willingness to pay taxes [5].

2. Method

The data obtained in this study were from KPP in Lampung, namely personal taxpayer data and the organizational structure. The sample size was 100 from Slovin formula, while we used stratified random sampling to choose respondents. This research used the R program version 4.0.2 for data quality test, classical assumption test, and multiple linear regression test.

2.1. Data Quality Test

In analyzing the questionnaire, we used the following methods:

a. Instrument Validity test

A validity test used is to measure the validity questionnaire. Testing this validity using the Pearson Correlation approach. If the correlation between the scores of each question item and the total score has a significance level below 0.05, then the question item is said to be valid, and vice versa [11,12].

b. Instrument Reliability Test

This reliability test produces a Cronbach Alpha value. A variable is said to be reliable if it gives a Cronbach's Alpha value > 0.6. If the Cronbach Alpha of a variable > 0.6, then the questions in the research instrument are reliable. Conversely, if the Cronbach Alpha value < 0.6, then the question item is not reliable [13].

2.2. Classical Assumption Test

All the data needed in the study must be tested first, so it does not violate the classical assumptions that exist can achieve the right hypothesis testing results and can be accounted for and produce a significant and representative regression model. The classic assumptions tested are Normality Test and Multicollinearity Test.
The distribution function of $F_n$ for the independent stochastically variable $X_i$ as much as $n$ is defined as: $F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{[-\infty, x]}(X_i)$, where $I_{[-\infty, x]}(X_i)$ is the indicator function which value is 1 if $X_i \leq x$ and 0 for another. Kolmogorov-Smirnov statistic for cumulative distribution of $F(x)$ can be written as: $D_n = \sup_x |F_n(x) - F(x)|$, where $\sup$ is the supremum for the distance set. Based on the theorem, if sample come from the distribution of $F(x)$, the $D_n$ will convergent to 0 if the limit is infinitive number [14]. The hypothesis is $H_0: X \sim \text{Normal}(\mu, \sigma^2)$.

2.3. Multiple Regression

The multiple regression model defined as [15]:

$$E(Y|X) = \beta'X$$
$$\text{Var}(Y|X) = \sigma^2$$

The model using error is:

$$Y = X\beta + \epsilon$$

where $Y$ is the $n \times 1$ vector of response values and $X$ is a $n \times p'$ matrix. If the mean function includes an intercept, then the first column of $X$ is a vector of ones, and $p' = p + 1$. If the mean function does not include an intercept, then the column of one is not included in $X$ and $p' = p$. The $ith$ row of the $n \times p'$ matrix $X$ is $x_i$, $\beta$ is a $p' \times 1$ vector of parameters for the mean function, $\epsilon$ is the $n \times 1$ vector of unobservable errors, and $\sigma^2$ is an unknown positive constant.

We use Ordinary Least Square (OLS) to estimate $\hat{\beta}$ that minimize the residual sum of squares function: $RSS(\beta) = (Y - X\beta)'(Y - X\beta)$. After use some mathematical concept we obtained: $X'X\beta = X'Y$ which called the normal equations. If the inverse of $(X'X)$ is exist and linearly independent, the OLS estimation give: $\hat{\beta} = (X'X)^{-1}X'Y$ [16].

In this study the multiple regression formula used is as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon,$$

where:

$Y$: Willingness to pay personal taxpayers who are freelancer
$\beta_0$: Constants (price of $Y$, if $X = 0$)
$\beta_i$: Regression coefficient, where $i = 1, 2, 3$
$X_1$: Tax Pay Awareness
$X_2$: Knowledge of Taxation Regulations
$X_3$: Perception of the effectiveness of the taxation system
$\epsilon$: error

3. Results and Discussion

The analysis starts by testing the effect of three independent variables on the dependent variable's willingness to pay taxes through the Multiple Regression. But before that, we examine the validity and reliability of each variable instrument.

3.1. Validity and reliability test

| Table 1. Test the Validity of Tax Paying Awareness Instruments |
|---------------------------------------------------------------|
| **Statement** | **Pearson Correlation** | **p-value** | **Explanation** |
|----------------|--------------------------|-------------|-----------------|
| K1             | .8512**                  | .0001       | Valid           |
| K2             | .76502**                 | .0001       | Valid           |
| K3             | .74612**                 | .0002       | Valid           |
| K4             | .7911**                  | .0001       | Valid           |
Table 1 shows that all statement items for variables with Tax Pay Awareness (K) sub-variables are at a significance level below 0.05; it can be concluded that the statements items for this variable are valid.

Table 2. Test the Validity of Knowledge Instruments for Taxation Regulations

| Statement | Pearson Correlation | p-value | Explanation |
|-----------|---------------------|---------|-------------|
| K5        | .7971**             | .0001   | Valid       |
| K6        | .7722**             | .0001   | Valid       |

Table 2 shows that all statement items for variables that have a Tax Regulatory Knowledge (P) sub-variable are at a significance level below 0.05; it means that the statements in this study are valid.

Table 3. Test the Validity of a Good Perception Instrument for System Effectiveness Taxation

| Statement | Pearson Correlation | p-value | Explanation |
|-----------|---------------------|---------|-------------|
| P1        | .6789**             | .001    | Valid       |
| P2        | .6241**             | .0001   | Valid       |
| P3        | .5511**             | .0001   | Valid       |
| P4        | .3541**             | .002    | Valid       |
| P5        | .6542**             | .0002   | Valid       |
| P6        | .5863**             | .0001   | Valid       |
| P7        | .6651**             | .0001   | Valid       |

Table 3 shows that all statements for variables that have a good perception of the Effectiveness of the Taxation System (PE) are at a significance level below 0.05; it can be concluded that the statements for perception variables are valid.

Table 4. Test the validity of the willing instrument to pay taxes which work as a freelancer

| Statement | Pearson Correlation | p-value | Explanation |
|-----------|---------------------|---------|-------------|
| KM1       | .7672**             | .0001   | Valid       |
| KM2       | .7061**             | .0001   | Valid       |
| KM3       | .5831**             | .0002   | Valid       |
| KM4       | .7192**             | .0001   | Valid       |
| KM5       | .6771**             | .0001   | Valid       |
| KM6       | .6292**             | .0001   | Valid       |
| KM7       | .6321**             | .0002   | Valid       |
| KM8       | .6871**             | .0001   | Valid       |

Table 4 shows that all statement items for variables that have a sub-variable Willingness to Pay Taxes for Personal Taxpayers Doing Freelancer (KM) are at a significance level below 0.05, so it can be concluded that the statements in this study are valid. This means that all statement items used in this study were able to reveal something measured on the questionnaire.
After testing validity of variable instruments, then we going through reliability test.

| Variable                                      | Cronbach Alpha | N of Item | Explanation |
|-----------------------------------------------|----------------|-----------|-------------|
| Tax Pay Awareness                             | 0.884          | 6         | Reliable    |
| Knowledge of Taxation Regulations             | 0.701          | 7         | Reliable    |
| Perception of the Effectiveness of the Taxation System | 0.754          | 5         | Reliable    |
| Willingness to pay taxes for personal taxpayers who do freelancer | 0.831          | 8         | Reliable    |

Based on the reliability test results in Table 5, it shows that the statements in this questionnaire are reliable because they have a Cronbach's Alpha value greater than 0.06. This result indicates that each item of statement used will obtain consistent data, which means that if the statement is submitted again, it will get an answer that is relatively the same as the previous answer.

3.2. Normality and multicollinearity test

To know the certainty of the distribution of the data obtained should be tested for normality of the relevant data. The following graph shows the normality test data on the pp-plot graph.

Based on the normal pp-plot graph in Figure 1 shows that the regression model is feasible to be used in this study because, in the normal graph, the plot looks the points spread around the diagonal line and the spread follows the direction of the diagonal line so that it meets the normality assumption.

Multicollinearity testing was conducted to test whether the regression model found a correlation between independent variables. Multicollinearity problems can be detected by looking at the value of the Tolerance and Variance Inflation Factor (VIF) and the magnitude of the correlation between independent variables.
Table 6. Multicollinearity Test

| Model | Unstandardized Coefficients | Standardized Coefficients | Collinearity Statistics |
|-------|-----------------------------|---------------------------|-------------------------|
|       | B | Std. Error | Beta | t | Sig. | Tolerance | VIF |
| (Constant) | 3.413 | 4.345 | .785 | .434 |
| X1 | .485 | .110 | .382 | 4.388 | .000 | .965 | 1.037 |
| X2 | .248 | .115 | .186 | 2.152 | .034 | .980 | 1.021 |
| X3 | .436 | .149 | .258 | 2.929 | .004 | .946 | 1.057 |

Table 6 above shows that each variable has a Tolerance value close to number 1 and the value of VIF around number 1. Awareness of paying taxes has a tolerance value of 0.965, knowledge of tax regulations has a tolerance value of 0.980, and a good perception of the effectiveness of the system taxation has a tolerance value of 0.946. And awareness of paying taxes has a VIF value of 1.037, knowledge of tax regulations has a VIF value of 1.021, and a good perception of the tax system's effectiveness has a VIF value of 1.057. Thus it can be concluded that the regression equation does not have a multicollinearity problem because the tolerance value is above 0.10, and the VIF value is below 10.

3.3. Parameter estimation

Regression in this study is used to calculate the effect of awareness of paying taxes, knowledge about tax regulations, and a good perception of the effectiveness of the taxation system on the willingness to pay taxes for personal taxpayers who do freelance. To determine the regression equation can be seen in the table below.

Table 7. Parameter estimate for multiple linear regression model

| Model | Unstandardized Coefficients | Standardized Coefficients |
|-------|-----------------------------|---------------------------|
|       | B | Std. Error | Beta | T | p-value |
| 1 (Constant) | 3.413 | 4.345 | .785 | .434 |
| X1 | 0.485 | 0.110 | 0.382 | 4.388 | .000 |
| X2 | 0.248 | 0.115 | 0.186 | 2.152 | .034 |
| X3 | 0.436 | 0.149 | 0.258 | 2.929 | .004 |

Based on table 7 above it is known that the coefficient value of the regression equation. From the output we get the regression equation model:

\[ \hat{Y} = 3.413 + 0.485X_1 + 0.248X_2 + 0.436X_3 \]

The results of the regression equation, a constant value of 3,413 means awareness of paying taxes (X1), knowledge of tax regulations (X2), and a good perception of the effectiveness of the taxation system (X3), are considered constant, so the willingness to pay taxes for personal taxpayers who do constant freelance equal to 3,413. The regression coefficient of variable awareness of paying tax (X1) of 0.485 means that awareness of paying taxes has increased by 4%, then the willingness to pay taxes for personal obligations which do freelance (Y) will experience an increase of 0.485 assuming other independent variables have a fixed value.

The regression coefficient of the taxation knowledge variable (X2) of 0.248 means that the knowledge of taxation regulations has an increase of 2%, then the willingness to pay taxes for personal taxpayers
who are freelancer (Y) will experience an increase of 0.248 with the assumption that the other dependent variables are of constant value.

The regression coefficient of a good perception variable on the effectiveness of the taxation system (X₃) of 0.436 means that a good perception of the effectiveness of the taxation system has increased by 4%, then the willingness to pay taxes for personal taxpayers which do freelance (Y) will experience an increase of 0.436 assuming other independent variables the value is fixed.

The coefficient of determination aims to measure how far the ability of the model can explain the variation of the dependent variable. The coefficient of determination is between 0 (zero) and 1 (one). A small R² value means that the ability of the independent variables to explain the variation of the dependent variable is very limited. A value close to one means that the independent variables provide almost all the information needed to predict the variation of dependent variables. This shows how big the percentage of variation in the dependent variable, the coefficient can be seen in the table 8 as follows:

| Mode | R   | R² | Adjusted R² | Std. Error of Estimate | Change Statistics | Durbin-Watson |
|------|-----|----|-------------|------------------------|-------------------|--------------|
| 1    | 0.545| 0.297| 0.275       | 3.9951                 | .297 | 13.513     | 3 | 96 | 0.000 | 1.951 |

Based on Table 8, an adjusted R² value of 0.275 is obtained. This shows that the percentage influence of the independent variable (awareness of paying taxes, knowledge of tax regulations, and a good perception of the effectiveness of the taxation system) on the dependent variable (willingness to pay taxes for personal taxpayers) by 27.5%. Or the variation of the independent variables used in the model (awareness of paying taxes, knowledge of tax regulations, and a good perception of the effectiveness of the taxation system) is able to explain as much as 27.5% of the variance of the dependent variable. At the same time, the rest 72.3% is influenced or explained by other variables not included in this research model, such as variable understanding of tax regulations, sunset policy, and service quality of KPP.

3.4. Analysis of variance
F statistical test in Analysis of variance aims to determine the effect of jointly or simultaneously independent variables on the dependent or dependent variable. The criterion used is if probability > 0.05, then Ho is accepted while vice versa Ho is rejected.

| Model     | Sum of Squares | Df | Mean Square | F     | p-value |
|-----------|----------------|----|-------------|-------|---------|
| Regression| 647.297        | 3  | 215.766     | 13.513| .000    |
| Residual  | 1532.863       | 96 | 15.967      |       |         |
| Total     | 2180.160       | 99 |             |       |         |

Based on Table 9 it is known that the significant value of 0.000 is smaller than 0.05. This result means that the independent variable is awareness of paying taxes, knowledge of tax regulations, and a good perception of the effectiveness of the taxation system, have a significant (simultaneous) effect on the willingness to pay taxes for personal taxpayers who do freelance.

4. Conclusion and suggestion
The regression model in this research is appropriate because there are significant values on each variable as well. Thus the higher the awareness of paying taxes, knowledge of tax regulations, and a good perception of the effectiveness of the taxation system, the higher the level of willingness to pay taxes for freelance personal taxpayers.
References

[1] Agung M 2007 Teori dan Aplikasi Perpajakan Indonesia Jakarta CV. Pustaka Setia
[2] Mardiasmo P 1995 Andi Offset
[3] Rantung T V and Adi P H 2009 Dampak Program Sunset Policy Terhadap Faktor–Faktor yang Mempengaruhi Kemauan Membayar Makal. Simp. Nas. Perpajak. II. Madura
[4] Widayati N and Nurlis A 2010 Faktor-faktor Yang Mempengaruhi Kemauan Membayar Pajak Wajib Pajak Orang Pribadi Yang Melakukan Pekerjaan Bebas (Studi kasus pada KPP Pratama Gambir Tiga) Makal. Simp. Nas. Perpajak. XIII (yang dipresentasikan pada tahun 2010)
[5] Hardiningsih P and Yulianawati N 2011 Faktor-faktor yang mempengaruhi kemauan membayar pajak Din. Keuang. dan Perbank. 3 126–42
[6] Nugroho R A and Zulaikha Z 2012 Faktor–Faktor Yang Mempengaruhi Kemauan Untuk Membayar Pajak Dengan Kesadaran Membayar Pajak Sebagai Variabel Intervening (Studi Kasus Wajib Pajak Orang Pribadi Yang Melakukan Pekerjaan Bebas Yang Terdaftar Di KPP Pratama Semarang Tengah Satu)
[7] Putri A R S and Jati I K 2012 Faktor-faktor yang mempengaruhi kepatuhan wajib pajak dalam membayar pajak kendaraan bermotor di Denpasar E. J. Akunt. 2
[8] Fikrinkingrum W K and Syafruddin M 2012 Analisis Faktor-Faktor Yang Mempengaruhi Wajib Pajak Orang Pribadi Dalam Memenuhi Kewajiban Membayar Pajak (Studi Kasus Pada Kantor Pelayanan Pajak Pratama Semarang Candisari)
[9] Andinata M C 2015 Analisis Faktor-faktor yang mempengaruhi kepatuhan wajib pajak orang pribadi dalam membayar pajak CALYPTRA 4 1–15
[10] Fahluzy S F and Agustina L 2014 Faktor-Faktor yang Mempengaruhi Kepatuhan Membayar Pajak UMKM di Kabupaten Kendal Account. Anal. J. 3
[11] Susanto H, Rinaldi A and Novalia N 2015 Analisis Validitas Reliabilitas Tingkat Kesukaran Dan Daya Beda Pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika Kelas XII Ips Di SMA Negeri 12 Bandar Lampung Tahun Ajaran 2014/2015 Al-Jabar J. Pendidik. Mat. 6 203–18
[12] Syazali M, Putra F G, Rinaldi A, Utami L F, Widayanti W, Umam R and Jermsittiparsert K 2019 Partial Correlation analysis using multiple linear regression: Impact on business environment of digital marketing interest in the era of industrial revolution 4.0 Manag. Sci. Lett. 9 1875–86
[13] Heale R and Twycross A 2015 Validity and reliability in quantitative studies Evid. Based. Nurs. 18 66–7
[14] Rinaldi A 2016 Sebaran Generalized Extreme Value (GEV) dan Generalized Pareto (GP) untuk Pendugaan Curah Hujan Ekstrim di Wilayah DKI Jakarta Al-Jabar J. Pendidik. Mat. 7 75–84
[15] Weisberg S 2005 Applied linear regression vol 528 (John Wiley & Sons)
[16] Myers R H and Myers R H 1990 Classical and modern regression with applications vol 2 (Duxbury press Belmont, CA)