Jurnal Kebijakan Publik, Vol.14, No.1, 2023
p-ISSN 1978-0680, e-ISSN 2655-5204
https://jkp.ejournal.unri.ac.id

POLICY IMPORT TAX INCENTIVE UTILISATION DURING THE COVID-19 PANDEMIC

KEBIJAKAN PEMANFAATAN INSENTIF PAJAK IMPOR SELAMA PANDEMI COVID-19

Rahadian Lungun[1*], Prianto Budi Saptono[2]
[1]Department of Public Policy and Administration, FIA, Universitas Indonesia, Depok, Jawa Barat
[2]Faculty of Administrative Science, Universitas Indonesia, Depok, Jawa Barat
*Koresponden email: rahadian.lungun@ui.ac.id

ABSTRACT

The objective of a policy will be achieved if the policy targets are well-defined. However, it often happens that the policy targets do not utilise the public policies given optimally. Therefore, it is essential to predict the targeted group that will utilise the policy provided so that the policy implementation can be more efficient. This study uses the Data Mining method as a policy decision support tool to analyse the best predictive model of the PPh 22 Import Tax Incentive utilisation during the COVID-19 Pandemic. The predictive model can be practical and valuable for Indonesia's tax authority and policy analysis knowledge. This study uses administrative data of 43.547 taxpayers who are already utilising and not utilising the tax incentive, combined with the data mining method. The results showed that with the Random Forest algorithm, the utilisation of the tax incentive could be predicted with an accuracy above 94%. Furthermore, The Total Sales Value, The Export Value, and The Bonded Zone Category are the variables that mostly predict the utilisation of the tax incentive. Therefore, tax authorities should be able to utilize data mining as a tool to be able to implement tax incentive policies more efficiently and accurately.

Keywords*: Public Policy, Tax Policy, Tax Incentive, Import Activity, Data Mining*

ABSTRAK

Tujuan suatu kebijakan akan tercapai jika sasaran kebijakan ditetapkan dengan baik. Namun seringkali sasaran kebijakan tidak memanfaatkan kebijakan publik yang telah diberikan secara optimal. Oleh karena itu, sangat penting untuk memprediksi kelompok sasaran yang akan memanfaatkan kebijakan yang diberikan agar implementasi kebijakan dapat dilakukan dengan lebih efisien. Penelitian ini menggunakan metode Data Mining sebagai alat pendukung keputusan kebijakan untuk menganalisis model prediksi terbaik dalam pemanfaatan Insentif Pajak PPh Pasal 22 Impor di masa Pandemi COVID-19. Model prediktif ini dapat secara praktis dapat bermanfaat bagi otoritas pajak di Indonesia dan dapat memperkaya pengetahuan terkait analisis kebijakan. Penelitian ini menggunakan data administrasi dari 43.547 wajib pajak yang sudah memanfaatkan dan belum memanfaatkan insentif pajak, dipadukan dengan metode data mining. Hasil penelitian menunjukkan bahwa dengan algoritma Random Forest, pemanfaatan insentif pajak dapat diprediksi dengan akurasi di atas 94%. Selanjutnya, Nilai Total Penjualan, Nilai Ekspor, dan Kategori Kawasan Berikat merupakan variabel yang paling dapat memprediksi pemanfaatan insentif pajak. Oleh karena itu, otoritas pajak seharusnya dapat memanfaatkan data mining sebagai salah satu alat untuk dapat mengimplementasikan kebijakan pemberian insentif pajak dengan lebih efisien dan akurat.

Kata kunci: Kebijakan Publik, Kebijakan Pajak, Insentif Pajak, Kegiatan Impor, Data Mining

INTRODUCTION

The COVID-19 Pandemic has engulfed the world and significantly impacted the global economy. From March 2020 to June 2021, as many as 2.2 million people were infected with the Sars-COV2 virus, which caused 60,000 people to die. After two years of experiencing the Pandemic, many countries take this as a valuable lesson to resolve the economic crisis due to the pandemic situation. At the beginning of the COVID-19 Pandemic, many countries were not ready to handle the COVID-19 Pandemic, as shown by several government policies in every country. This situation happened because the government must quickly handle the impact of the virus on public safety by applying lockdown policies or social scale restrictions. At the same time, the government also has to implement a responsive policy to support economic activity amid the Pandemic. When the Pandemic started in March 2020, the government quickly implemented several policies related to handling the spread of the Sars-COV2 virus, including the Large Scale-Social Restriction (PSBB) policy. At the same time, the government also implemented a policy handling crisis economy through an economic recovery policy.

The government has carried out the fiscal policy as a COVID-19 crisis counter cycle since March 2020. It started with the enactment of Lieu of Law Number 1 of 2020, dated March 31, 2020, followed by Law Number 2 of 2020 concerning State Financial Policy and Financial System

Stability for Handling the Corona Virus Disease Pandemic 2019 (COVID-19). The program which was implemented at that time was called the National Economic Recovery Program (PEN).

Moreover, one PEN program provides a tax incentive for the taxpayers affected by the COVID-19 Pandemic. There are six types of tax incentives that were implemented during the pandemic, namely: Income Tax Article 21 (PPh 21) borne by the government, MSME PPh Final borne by the Government, Income Tax Article 22 Import (PPh 22 Import) exemption, Income Tax Article 25 (PPh 25) Instalment Reduction, Preliminary Return VAT, and Income Tax Final Construction Services Program for the Acceleration of Irrigation Water Utilisation Improvement (P3-TGAI) borne by the government.

According to Indrawati et al. (2022) , from 2020 to 2021, there were as many as 14.2 trillion Rupiah of tax incentives given by the government. The PPh 22 Import incentive is the tax incentive that the taxpayers most utilise. However, based on a survey conducted by the Ministry of Finance (2021) , even though the PPh 22 Import incentive has been considered to be beneficial for taxpayers, it turns out that only 58% of respondents have utilised the tax incentive. On the other hand, as much as 42% of taxpayers choose to refrain from utilising the tax incentives due to other reasons. Furthermore, Indrawati et al. (2022) also state that there is a problem in targeting the taxpayers who are really in need of the tax incentive. Due to this problem, several taxpayers did not utilise the tax incentive during COVID-19. Therefore, there is a problem in the implementation of the tax incentive. Where the tax incentive did not utilise the tax incentive optimally while the government has already allocated the budget for the tax incentive, the other problem is that the government has to face the policy targeting problem due to the lack of data accuracy. Thus, there are also taxpayers who are in need of the tax incentive but cannot utilise it.

In terms of tax policy implementation during the COVID-19 Pandemic, the OECD (2020) has provided guidelines and recommendations. The government should implement a comprehensive yet well-coordinated policy. Especially for developing countries, the government must prioritise supporting healthcare facilities which are very likely insufficient to deal with the COVID-19 Pandemic. Additionally, the OECD also recommends that during the COVID-19 Pandemic, tax incentives should also be applied, especially for business sectors that need support. Moreover, Dye (2017) asserts that a great deal of a policymaking process is the budgetary and appropriateness process. The fiscal range space can describe how strongly the government can implement the policy appropriately. It is argued that the implementation of the tax incentive during the COVID-19 Pandemic requires a considerable budget. Therefore, if the government can predict the amount of the budget that is needed to be provided, then the implementation of the tax incentive policy can be more efficient and effective. In regard to this, the government has realised that the uncertainty caused by the Pandemic requires more budget to allocate. Therefore, the government revised the state budget deficit to allow more than 3% of the deficit. Thus, the government can manage more budget for implementing several policies.

Furthermore, in terms of lack of data accuracy, it is argued that the Directorate General of Taxes (the tax authority of Indonesia) has already administered vast amounts of data related to taxpayers. The data are varied from the taxpayer's identity, the financial report and the tax payment that reflects the detail and the description of the taxpayer's business activity. Therefore, the government should not have to face the data accuracy issue. Rosid (2022) asserts that the tax authority should have utilised the administrative data and combined it with the utilisation of data mining methods to predict taxpayers' behaviour. In this way, the data can be used in a whole new way to support the policy implementation with less of an accuracy issue.

Regarding the data mining method, Atwell et al. (2015) argue that data mining can describe various computationally intensive techniques for finding structures and analysing patterns in data. Data mining is one of the approaches that can process large amounts of data with multiple attributes and use complex algorithms to find patterns in the data. However, Atwell et al., (2015) also asserted that data mining would not replace quantitative analysis with conventional statistical research. The existence of a data mining method can actually complement conventional statistical research through its ability to produce a predictive model with high predictive power. Therefore, the analysis will be more beneficial in practical and future research.

Based on the report by the IMF (2021) concerning the utilisation of tax incentives, the government should pay more attention to the risk of targeting the tax incentive. It is because, in some countries, there are several taxpayers taking advantage of crisis relief/stimulus measures, even

though they are not supposed to use the tax incentive. Furthermore, Satyadini et al. (2019) also state that the utilisation of tax incentive terms may require rigorous scrutiny to prevent or detect potential abuse. She states that the degree of tax avoidance risk varies with the characteristics of taxpayers, whereas higher risks were found in so-called foreign-owned companies. Therefore, the tax authority should pay more attention to the utilisation of tax incentives during the Pandemic. The higher risk of utilisation tends to occur in taxpayers categorised as Medium, Large, and Foreign-Owned Companies taxpayers.

It is argued that tax incentives are significant for helping taxpayers' business activity during the demand and supply shock (Chetty et al., 2020) . Meanwhile, Crystallin (2022) also states that the tax incentive is a must during the Pandemic, and a "do-nothing" policy is a non-viable option. Fiscal stimulus has been proven to bolster trade activities on the demand side to cope with demand and supply shocks. On the other hand, supply-side support by the government is also needed mainly by the physical-contact-intensive and labour-intensive sectors. However, some challenges in implementing fiscal stimulus programs still need to be addressed. Among the many challenges is the accuracy of data and targeting. Better data collection and analysing methods will improve the utilisation of programs to strengthen policy implementation in Indonesia.

To the best of the researcher's knowledge, three studies have been conducted regarding the taxpayers' behaviour toward participation in tax incentives and tax regulation in Indonesia (Rosid, 2022) , (Febriminanto & Wasesa, 2022) , and (Rosid et al., 2022) . Meanwhile, according to Mittelstaedt et al. (2006), there are determinants of the import activity tendency. Their findings showed that large manufacturing firms are more likely to import than small ones. It is also identified that the firm located in the concentrated area or an Industrial area (near other firms in their area) also has more tendency to import. Furthermore, Busom et al. (2017) argued that the participation of taxpayers utilising tax incentives also could be explained by the variable of the number of workers and market share. The salary ratio can measure the number of workers compared to the cost of goods sold or the cost of production. Moreover, the market share can be measured by the proportion of import value to the sales value of a firm.

This study intends to fill the gap in predicting the utilisation of PPh 22 Import using data mining with the identified variables that affect the determinant of import behaviour and the tax incentives utilisation. The prediction model of the PPh 22 Import tax incentive will be considered beneficial to the Directorate General of Taxes as a policy implementer in setting the right target for the tax incentive. Moreover, it is also can enrich the knowledge and fill the gap of how data mining can be a support tool for the policymaker to implement an efficient and effective policy. Therefore, the theoretical framework of this study is as follows:

Source: Processed by Researcher

Figure 1. Theoretical Framework

METHODS

The approach which is used in this research is positivism. According to Neuman (2014) , the positivism paradigm is basically a quantitative approach. Moreover, according to Creswell & Creswell (2017) , quantitative is an approach that tests objective theories by examining relationships between variables. These variables can be measured with instruments so that statistical techniques can be used to analyse the numbered data.

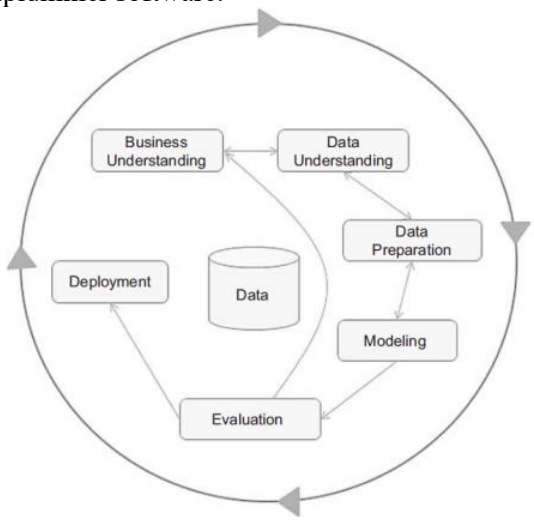
This study uses secondary data of the taxpayers' population which is registered in the Medium Taxpayer's Tax Offices, the Large Taxpayer's Tax Offices, the Foreign Investment Tax Offices and the Stock Listed Company Tax Office. The amount of the population of taxpayers registered in those tax offices will be as follows: (1) Medium Taxpayers Tax Offices: 33.258 taxpayers (2) Large Taxpayers Tax Offices: 4.352 taxpayers (3) Foreign Investment Taxpayers Tax Offices: 4.220 taxpayers (4) Stock Listed Company Tax Office: 1.717 taxpayers.

In terms of research time, this study is carried out to analyse the tax incentive utilisation during the implementation of the PPh 22 Import Tax Incentive. Therefore, this research conducts the analysis from April 2020 to December 2021. This selection considers that the government has provided the tax incentive consistently to the Taxpayers during the COVID-19 Pandemic.

This study uses CRISP-DM (Cross Industry Standard Process for Data Mining) as a data mining process framework. This framework was

developed by a consortium of several companies working in data mining. CRISP-DM is widely used and adopted as a frame of reference in developing data mining solutions. Although there are several other methods of reference framework besides CRISP-DM, all of these methods have similarities in the process and evaluation of the final results.

The data mining process is not linear but has an iterative or continuous process. As depicted in Figure 2, it can be seen that several steps must be taken. According to Kotu & Deshpande (2015) , the data mining process can be divided into five steps: Prior Knowledge, Preparation, Modelling, Application, and Knowledge. The main goal of any process that involves data mining is to answer analytical questions using algorithms available in software or tools such as Rapidminer, R, Weka, SAS, and SPSS. However, in this study, the analysis uses a classifying algorithm using Rapidminer software.



Source: (Kotu & Deshpande, 2015)
Figure 2. CRISP DM Framework

RESULTS AND DISCUSSION

Business and Data Understanding

Regarding the implementation of fiscal policy in the form of tax incentives, whilst it must be confirmed that the effects of policy should be measured, the target and the non-target group should also be identified and separated to ensure the implementation is right on the target. By doing so, the effect of the policy will be impactful had the utilisation of the tax incentive been optimal (Dye, 2017) . Therefore, predicting the utilisation of the PPh 22 Import tax incentive with data mining will be helpful for the tax authority or other policy implementers in setting the policy target effectively.

However, aside from the identified variables which use in this study, it is important to note that there are other factors that may affect the utilisation of tax incentive. For example, differences in the level of outreach carried out by the tax office where the taxpayers are registered, differences in tax literacy levels of taxpayers, and differences in the status of using intermediaries or tax consultants by taxpayers. However, this study uses the variable based on the availability of the administrative data and test the findings in the study by Mittelstaedt et al. (2006) and Busom et al. (2017).

Data mining (DM) describes various computationally intensive techniques for finding structures and analysing patterns in data. Moreover, data mining can be classified as one of the paradigms that can process large amounts of data with multiple attributes and use complex algorithms to find patterns in the data (Atwell et al., 2015).

Kotu & Deshpande (2015) argues that data mining is one of the paradigms that can process large amounts of data with multiple attributes and use complex algorithms to find patterns in the data. There are several ways in data mining to predict the dependent variable. One of them is called a classifier. In some classification algorithms, classification involves developing a statistical predictive model using a data set of independent variables or attributes to predict each data value that will be the target or dependent variable. The results of these predictions can be described at a probability level which is then used to classify the categories of data that will then be predicted.

In this study, the utilisation of the PPh 22 Import tax incentive is classified as the decision of the taxpayers utilising or not utilising the tax incentive. Meanwhile, the independent variables which is used in this study are according to the previous research by Mittelstaedt et al. (2006) and Busom et al. (2017). Those independent variables are (a) the size of the firm measured by the sales value or turnover of the firm, (b) the location of the firm categorised as in bonded zone or non-bonded zone, (c) the number of workers which the salary cost ratio can measure compared to the cost of goods sold or the cost of production, and (d) the market share value measured with the import value compared to the cost of goods sold or cost of production, export value, and business sector category.

The objective of this study is to produce the best prediction model to predict the use of tax incentives based on predetermined variables. The prediction model made in this study uses a

binominal classification algorithm that consists of Logistic Regression, Random Forest, and Decision Tree. Following the business and data understanding steps that have been taken before, the data preparation step will be carried out before going into the modelling prediction.

Data Preparation

Once the dataset has been gathered, it is identified that three variables contain missing values. Therefore, to run the analysis, this study uses the Replace Missing Value command to fill the missing value with each variable's average value. The following step identifies and selects the optimum variable that can be used in the predictive model. This step uses the Weight by Information Gain Ratio and Optimize Selection command to find the weighted data.

Table 1. Variable Optimise Selection

Attribute	Weight
Total Sales Value	0,252
Export Value	0,228
Bonded Zone Category	0,072
Salary Cost Ratio	0,059
Import Intensity	0,040
Business Sector	0,023
Business Age	0,020

Source: Processed by Researcher

Based on the Table 1, it can be seen that the total sales value, export value and bonded-zone category are the most weighted variables to predict the utilisation of the PPh 22 Import tax incentive. Therefore, this variable sequence is in line with the previous research conducted by Mittelstaedt et al. (2006) and Busom et al. (2017). However, in this study, the rest of the variables will also be included as independent variables since this study aims to seek the best predictive model with the highest accuracy. Thus, modelling will be conducted with the three algorithm types mentioned before after knowing the weighted value.

Modelling

The strategy for creating the predictive model is using Split Data. As much as 80% of the data will be set as data training, while 20% will be set as data testing. Afterwards, Cross Validation is used as one of the requirements to produce an effective predictive model so that the resulting model is not overfitting and can be used to deploy data testing or other practical data.

Finally, to measure the accuracy and effectiveness of the predictive model, this study uses the Binominal Classification Performance, which will show the level of accuracy based on the Confusion Matrix and Area Under Curve (AUC).

Table 2. Decision Tree

Accuracy: 93,70% +/- 0.36% (micro average: 93,70%)

	True Not Utilising	True Utilising	Class Precision
Pred. Not Utilising	32.196	1.955	**94,28%**
Pred. Utilising	242	467	**65,87%**
Class Recall	**99,25%**	**19,28%**	

Source: Processed by Researcher

Table 3. Logistic Regression

Accuracy: 93.06% +/- 0.14% (micro average: 93.06%)

	True Not Utilising	True Utilising	Class Precision
Pred. Not Utilising	32.387	2.367	**93,19%**
Pred. Utilising	51	55	**51,89%**
Class Recall	**99,84%**	**2,27%**	

Source: Processed by Researcher

Table 4. Random Forest

Accuracy: 94.39% +/- 0.24% (micro average: 94.39%)

	True Not Utilising	True Utilising	Class Precision
Pred. Not Utilising	32.287	1.803	**94,28%**
Pred. Utilising	151	619	**80,39%**
Class Recall	**99,53%**	**25,56%**	

Source: Processed by Researcher

Evaluation

Based on the Table 2, it can be seen that the overall accuracy of tax incentive utilisation prediction reaches 93,7%. However, the Decision Tree algorithm model only produces 65,87% of accuracy in predicting the taxpayers who will utilise the tax incentive.

Moreover, according to the Table 3, the overall accuracy of Logistic Regression model cannot produce higher accuracy than the Decision Tree model (93,06%). Because, at the same time, the model can only predict 51,89% of taxpayers who will utilise the tax incentive.

On the other hand, the Random Forest algorithm surprisingly can produce highest accuracy. Based on the class precision, the model can

predict to more than 80% of accuracy of the taxpayers who will utilise the tax incentive.

Based on the comparison of the confusion matrix generated from the three algorithms above, it can be evaluated that the Random Forest algorithm produces the best predictive model with an accuracy level of up to 94.39%. The prediction that taxpayers will not utilise the tax incentives reaches 99.53%. Also, the class precision value of taxpayers using tax incentives reaches 80.39%. However, the strength level of the predictive model accuracy cannot rely on a single measure of accuracy. Therefore, Area Under Curve (AUC) and Receiver Operator Characteristics (ROC) should be used to rate a model's performance. According to Figure 3, it can be seen that the predictive model is effective due to its value of AUC reaching 0.949[1].

Deployment

Ultimately, the best predictive model that has been produced in the previous step (Random Forest algorithm model) will be tested using 20% of the data that has been previously separated using split data. Based on the deployment results, it can be seen that the overall accuracy of the model has not changed significantly (94.09%), with the level class recall at 99.65% and class precision value in predicting taxpayers who will use tax incentives reaching 80.82%.

AUC: 0.949 +/- 0.005 (micro average: 0.949) (positive class: Utilising)

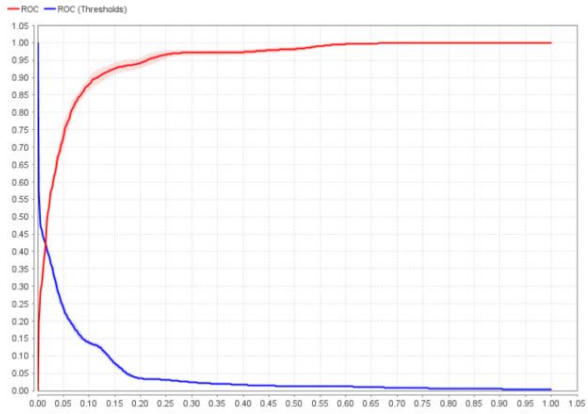


Figure 3. Area Under Curve (AUC) and Receiver Operator Characteristics
Source: Processed by Researcher

[1] Any AUC higher than 0.5 is better than random and the closer it is to 1.0, the better the performance

Table 4. Confusion Matrix Deployment
Accuracy: 94.09%

	True Not Utilising	True Utilising	Class Precision
Pred. Not Utilising	8.08	487	**94,32%**
Pred. Utilising	28	118	**80,82%**
Class Recall	***99,65%***	***19,50%***	

Source: Processed by Researchers

CONCLUSION

This study examines the prediction of the utilisation of the PPh 22 Import tax incentive during the COVID-19 pandemic. Based on the analysis above, using administrative data combined with the data mining method, the utilisation of the tax incentive can be predicted with an accuracy of more than 94%. Based on the weighted value, it can be seen that the variables that can predict the utilisation of the PPh 22 Import tax incentive are Total Sales Value, Export Value, and Bonded Zone Category. The predictive model is consistent with the results of previous studies, which state that the tendency of import activities and the utilisation of tax incentives can be explained by factors such as the size of the firm, the location of the firm, and the market share value. Overall, based on the analysis, the Random Forest Algorithm produce the best predictive model, where the prediction accuracy rate reaches more than 94%. Moreover, after testing, the predictive model is practical and not overfit.

REFERENCES

Atwell, P., Monaghan, D. B., & Kwong, K. (2015). *Data Mining for the Social Sciences: An Introduction* (Vol. 1). University of California Press.

Busom, I., Corchuelo, B., & Martínez-Ros, E. (2017). Participation inertia in R&D tax incentive and subsidy programs. *Small Business Economics*, *48*(1), 153–177. https://doi.org/10.1007/s11187-016-9770-5

Chetty, R., Friedman, J. N., Hendren, N., & Stepner, M. (2020). *The Economic Impacts of COVID-19: Evidence from a New Public Database Built from Private Sector Data*.

Creswell, J. W., & Creswell, D. J. (2017). *Research Design:* Qualitative*, Quantitative, and Mixed Methods Approaches*.

Crystallin, M. (2022). Walking a Thin Line. In *COVID-19 in Indonesia* (pp. 71–100). Routledge. https://doi.org/10.4324/9781003243670-4

Dye, T. R. (2017). *Understanding public policy*. Pearson.

Febriminanto, R. D., & Wasesa, M. (2022). Machine Learning Analytics For Predicting Tax Revenue Potential. *Indonesian Treasury Review*, *7*(3), 193–205. www.pajak.com

IMF. (2021). *Revenue Administration: Reinvigorating Operations to Safeguard Collection and Compliance, IMF COVID-19 Special Series*. https://www.imf.org/en/Publications/SPROLLs/covid19-special-notes.

Indrawati, S. M., Nazara Suahasil, Anas, T., Ananda, C. F., & Verico, K. (2022). *Keeping Indonesia Safe from the COVID-19 Pandemic: Lessons Learnt from the National Economic Recovery Programme*.

Kotu, V., & Deshpande, B. (2015). *Predictive analytics and data mining : concepts and practice with RapidMiner*. Morgan Kauffman.

Melati, M., & Zulkarnaini, H. (2021). Efektivitas Program Bantuan Sosial Tunai Pada Masa Pandemi Covid-19 Di Kelurahan Pasir Pengaraian. Journal Publicuho, 4(1), 19.

Ministry of Finance. (2021). *Pemanfaatan Stimulus 2020 & Outlook Stimulus 2021*.

Mittelstaedt, J. D., Raymond, M. A., & Ward, W. A. (2006). External and Internal Factors Affecting the Tendency to Import. *Journal of Global Marketing*, *19*(2), 27–48. https://doi.org/10.1300/J042v19n02_03

Neuman, W. L. (2014). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Pearson Education Limited.

OECD. (2020). *Tax and fiscal policy in response to the Coronavirus crisis: Strengthening confidence and resilience*.

Rosid, A. (2022). Predicting Firms' Taxpaying Behaviour Using Artificial Neural Networks: The Case of Indonesia. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4185966

Rosid, A., Ardin, G., & Sanjaya, T. B. (2022). Prediksi Keikutsertaan Pelaku Usaha dalam Pemanfaatan Insentif Pajak dengan Artificial Neural Network. *Jurnal Ekonomi Indonesia* •, *11*, 109–142.

Satyadini, A. E., Rea Erlangga, R., & Vdkg, B. S. (2019). Who Avoids Taxes? An Empirical Evidence From The Case of Indonesia. *SCIENTAX*, *1*(1), 1–26.