Aggressiveness as Performance Signaling: Forced Distribution Rating System in Tax Authority

Iswahyudi, Heru

27 July 2021

Aggressiveness as Performance Signaling: Forced Distribution Rating System in Tax Authority

Heru Iswahyudi[*]

July 27, 2021

Abstract

This article is aimed at examining whether introducing the forced distribution rating system (FDRS) in tax authority may improve performance and at exploring alternative explanations for the results utilizing the perspectives of signaling theory. Using Indonesian tax authority as case study and employing the interrupted time-series analysis on tax collection data, it is found that the adoption of FDRS may have worsened the already declining trend in collection. One of the possible explanations for this deterioration is that the introduction of FDRS, combined with the existing system of collection targets, may incentivize tax officials to demonstrate aggressive behaviors toward taxpayers as signals of performance with the desired outcome of being categorized as top performers under the FDRS. These behaviors may erode taxpayers' trust in tax authority, diminish voluntary compliance, and encourage migration into the underground economy; thence spurring a vicious cycle: as the tax base is shrinking because taxpayers are going underground, tax revenue will decline thus prompting tax officers to signal their performance by behaving more aggressively toward taxpayers, hence encouraging more taxpayers to escape to the underground economy. It is important for tax authorities to implement performance evaluation systems which are conducive toward creating and maintaining trustworthy relationships between taxpayers and tax officers.

JEL Classification: H26, M10, M12, M52, M54

Keywords: Signaling Theory, Forced Ranking, Tax Compliance, Underground Economy, Trust in Tax Authority

[*] Ministry of Finance – Indonesia. Email: heru.iswahyudi@gmail.com

This paper was prepared in author's personal capacity. The opinions expressed in this article are the author's own and do not reflect the view of the organization with which the author affiliated.

1. Introduction

Forced distribution rating system (FDRS) is a performance evaluation system that obligates raters to distribute a prearranged percentage of ratees in several categories based on the ratees' performance relative to the performance of other ratees (Moon, Scullen, & Latham, 2016). Business and public organizations are no stranger to this employee rating system, which is also known as 'forced ranking', 'stack ranking', 'rank and yank', or 'the bell curve'. Although typically aimed at improving the performance of organizations, however, adopting FDRS for tax authorities may trigger adverse behavioral response on the part of tax officers. FDRS may incentivize tax officials to act aggressively toward taxpayers as signals for their supervisors regarding their supposed ability or performance in collecting taxes. This signaling effect risks the emergence of negative externalities in the relationships between taxpayers and tax officers, particularly in terms of trust relations between these two parties.

This article examines the Indonesian tax administration (Direktorat Jenderal Pajak – Directorate General of Taxes, henceforth DGT) as a case study because research related to challenges in the country's taxation systems and performance have been mainly focused on identifying defects in tax regulations (see, for example, in Alm (2019); Iswahyudi (2020b)) and limits in the capacity of the tax authority (e.g., Iswahyudi (2020a)). Studies on how performance evaluation systems for tax officers – in this case, the FDRS – may affect the levels of revenue collection have rarely been done; particularly for the case of developing countries with significant size of underground economy such as Indonesia. Moreover, although studies have emphasized the importance of trust for tax collection (e.g., Feld and Frey (2002); Hammar, Jagers, and Nordblom (2009); Kaplanoglou and Rapanos (2013)), however, the system by which tax officers' performance are evaluated and how this evaluation system may affect the dynamics in the relationships between taxpayers and tax officers have rarely been studied.

This article is aimed at filling these gaps in the literature. In particular, it is aimed at examining whether FDRS adoption in a tax authority could improve revenue performance and to seek alternative explanations for the results. Therefore, the first

step is to conduct quantitative examinations on the impact of FDRS adoption on tax collection performance. The second step is to provide possible explanations for the result found in the previous step; these explanations will be explored under the perspectives provided by the signaling theory. To the best of the author's knowledge, this article is the first which examines the effects of FDRS introduction in a tax authority, with specific focus on how this system might affect the behavior of tax officers and how this behavioral response might affect taxpayers' compliance. Further, this article is unique, in the sense that it is the first which tries to apply the signaling theory to explore the possible behavioral responses of tax officials to the introduction of FDRS.

2. Literature Review

2.1. Forced Distribution Rating System

FDRS might closely be associated with Jack Welch who popularized this system (Kwoh, 2012; Lawler, 2002). When he lead General Electric Company its employees were subjected to a 20/70/10 'vitality curve' in which the company's supervisors determined the 'top 20%', 'vital 70%', and 'bottom 10%' performers among their subordinates (Welch & Byrne, 2003). Those in the top 20% typically received bonuses two to three times higher than those in the vital 70%, whereas those in the bottom 10% generally got nothing and their employment were often terminated. Other companies use (or had used) some variations of FDRS with different percentage distributions (see Eichenwald (2012); Osborne and McCann (2004); Ovide and Feintzeig (2013)). Beside affecting employees compensations and terminations, FDRS also plays a key factor in employees promotions with those in the top ranks have relatively higher chances to advance their career (Ovide & Feintzeig, 2013).

This performance evaluation system has been practiced since before World War II to rank the individual performance of U.S. Army officers, and subsequently this system was picked up by industry after the War (Cappelli, 2009). As a performance management tool, FDRS has potential positive effects as well as negative ones. One of the potential positive effects of the FDRS is that this system might reduce the possibility of leniency on the part of the rater and persuade employees to give more

efforts. In an absolute rating system, employee ratings might be inflated thus a disproportionate number of them could be rated as above average (Johannes Berger, Harbring, & Sliwka, 2013; Bretz Jr, Milkovich, & Read, 1992). FDRS could minimize this problem by providing a relative rating system whereby a supervisor is forced to rate their subordinates relative to one another, thus one would know where one stands relative to others (Welch & Welch, 2005). FDRS might also motivate employees to give more effort since on the one side there are higher financial rewards, chances for promotion, as well as the recognition of competence and status in the organization for those in the top ranks while, on the other side, there are threats of termination for those in the low ranks (Boswell & Boudreau, 2002; Jenkins Jr, Mitra, Gupta, & Shaw, 1998; Rynes, Gerhart, & Parks, 2005). The other potential benefit of the FDRS is that the system might support an organization in its effort to attract and retain high performers. Several studies suggested that organizations that reward employees based on job performance tend to attract applicants with higher cognitive ability, self-confidence, and need for achievement (Blume, Rubin, & Baldwin, 2013; Cable & Judge, 1994; Trank, Rynes, & Bretz, 2002; Turban & Keon, 1993).

On the other hand, there are potential negative effects associated with the use of FDRS as a tool for performance management. Adsit, Bobrow, Hegel, and Fitzpatrick (2018) argued that although simulation research might indicate increases in performance, however, in practice the FDRS may result in adverse consequences for organizations. Measuring performance often involves subjective measures and may potentially inaccurate, however, employee rankings compelled by FDRS may have to be based on these flaws (Pfeffer & Sutton, 2006; Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). The problems could be compounded in organizations which require high interdependence and teamwork among units or staffs since the FDRS would increase supervisors' confusions in evaluating the performance of each subordinate (Moon et al., 2016).

Further, FDRS might diminish employees' commitment to organization due to the potential increase in their perceptions of workplace injustices; this, in turn, may lead to negative work behaviors as well as induce counterproductive competitions among workers and reduce collaborations within members of teams or units (Moon et al.,

2016). In similar lines, Whiting and Kline (2007) argued that FDRS might increase employee turnover due to its adverse effects on employee commitment and job satisfaction. Moreover, FDRS might sow workers' distrust in managers, thus diminishes managers' ability to give meaningful coaching and feedbacks (Pfeffer & Sutton, 2006; Schleicher, Bull, & Green, 2009). With the FDRS, rankings from different work groups or teams may be compiled together to arrive at organization wide, overall ranking of employees. As a result, high absolute performers in a work group could find themselves placed in the lowest category and low absolute performers could find themselves placed in the top category (Chattopadhayay & Ghosh, 2012). This may happen due to differences in subjective measures used by different supervisors for different work groups. When differences in quality and performance of employees are insignificant and meaningless supervisors might have no option other than to create fictitious bell-curve distributions of their subordinates since the bell-curve is mandated by FDRS (Schleicher et al., 2009). After a comprehensive review of literature and careful study using an integrative model to assess the effect of FDRS on performance, Moon et al. (2016) concluded that over time the negative effects of FDRS might increasingly outweigh its initial positive effects.

2.2. Signaling Theory

Signaling theory basically concerns with information asymmetry between two parties (Connelly, Certo, Ireland, & Reutzel, 2011; Spence, 2002) and this information asymmetry takes place when "different people know different things" (Stiglitz, 2002, p. 469). Insights from the signaling theory have been widely used to explain business management phenomena (Kharouf, Lund, Krallman, & Pullig, 2020). For example, in the field of human resources management and organizational behavior, the signaling theory has been used to examine recruitment and talent attraction processes (Bangerter, Roulin, & König, 2012; Celani & Singh, 2011) as well as to examine employer branding and corporate reputation strategies (Martin & Groen-in't Woud, 2011; Taj, 2016).

There are four main elements in the signaling theory: the signaler, the signal, the receiver, and feedback (Bergh, Connelly, Ketchen Jr, & Shannon, 2014). Signaler is an

insider who has specific information that is not available to outsider. This signaler's private information may involve knowledge on the quality of some aspects of an individual, product, or organization (see, as examples, in Joël Berger (2019); Li, Tang, Jiang, Yen, and Liu (2019); Thomas, Darby, Dobrzykowski, and van Hoek (2020)). Signal is information sent by the insider to outsider with the intention to achieve the insider's desired outcome, receiver is the outsider who intent to receive information privately owned by the insider or the signaler, and feedback refers to the perceptions and action of the receiver after receiving and interpreting the signal sent by the signaler (Connelly et al., 2011; Taj, 2016).

In organizations, supervisors' organizational hierarchy and their ability to decide rewards (or punishments) for subordinates may make them powerful signalers and important receivers at the same time, thus they may have credible capacity to demand certain desirable behaviors from subordinates (Briker, Hohmann, Walter, Lam, & Zhang, 2021; Magee & Galinsky, 2008; Shin & Zhou, 2003). On the other hand, the personal characters of subordinates may impact how they interpret the signals from their supervisors (Costa & McCrae, 2008). Hence, signals sent by supervisors may act as stimuli to induce the desired behavior of subordinates; whereas subordinates would construe these stimuli according to each subordinate's personality-derived perceptions and cognitive boundaries (Ehrhart & Klein, 2001; Felfe & Schyns, 2006).

3. Data

As a government agency responsible for administering tax collection at national level, DGT collects income taxes, value-added taxes, sales taxes (for goods deemed as luxurious), land and building taxes (for agricultural, forestry, and mining industries), and stamp duty. DGT started using FDRS in May 2018 and this adoption was part of bureaucratic reforms program in the DGT, with the implicit aim is to improve the revenue performance (Ministry of Finance, 2021; PER-12/PJ/2018, 2018). In this system, DGT employees are categorized into five performance criteria: S (for top 15% performers), A (for the upper middle 20%), B (for the middle 30%), C (for the lower middle 20%), and D (for the bottom 15%). Those in the top category will receive higher salaries and chance to be promoted.

Because one of the main tasks of DGT is to collect tax revenue, the effect of the adoption of FDRS would be examined on the basis of its impacts on collection performance. In this article the tax ratio – i.e., the ratio of tax revenue to gross domestic product (GDP) – was employed to measure collection performance. Data were collected from Badan Pusat Statistik (2019, 2021); CEIC (2021). Data on GDP only available in quarterly frequency thus the data frequency for tax ratio would have to follow accordingly and would cover the period from Quarter 1 – 2015 to Quarter 1 – 2021. This period was selected under time proportion consideration: from the start of FDRS adoption (Quarter 2 – 2018) to the last available data (Quarter 1 – 2021) consisted a total of 12 quarters, hence the period from Quarter 1 – 2015 to Quarter 1 – 2018 would provide a more or less balanced collection data for the period before the FDRS adoption.

This article focused its study on two types of tax: income tax and value-added tax. This was because these two tax instruments accounted for the majority of revenue: 97.8% of total revenue collected by DGT is from income taxes and value-added taxes (Direktorat Jenderal Pajak, 2019). Therefore, if there was any impact from the adoption of FDRS in DGT then it should be reflected on the collection performance of these two tax instruments. The reason for using tax ratio is that this ratio may show how much of a country's economic output that are able to be collected by the tax authority as tax payments under the applicable laws, hence one of the crucial factors that affects this ratio is the administrative capacity of the tax authority (Bird, 2004). It is worth noting here that the majority of fiscal incentives to support Indonesian taxpayers affected by the Covid-19 pandemic have been provided in the form of 'taxes borne by the government'; i.e., it is the government – not taxpayers – who pays the taxes, hence these incentives are recorded in the government's balance sheet as tax revenue. Table 1 presents statistical description for the data on tax ratio for the period under study.

4. Methodology

Interrupted time-series analysis (ITSA) was employed to examine the impact of FDRS on the performance of DGT. ITSA is considered to be the suitable research design in this study because the intervention (i.e., the introduction of FDRS) is expected to

'interrupt' the level or trend of the time series under examination (i.e., the tax ratio). Moreover, ITSA might offer a quasi-experimental research design with a potentially high degree of internal validity because multiple observation in the pre-intervention period (i.e., tax ratio before FDRS adoption) and the post-intervention period (i.e., tax ratio after FDRS adoption) can be obtained (Linden, 2015, 2017).

Table 1. Descriptive Statistics

Mean	8.34
Standard Error	0.36
Median	8.15
Mode	N/A
Standard Deviation	1.80
Sample Variance	3.23
Kurtosis	0.02
Skewness	0.60
Range	7.20
Minimum	5.59
Maximum	12.79
Sum	208.56
Count	25

Sources: Badan Pusat Statistik (2019, 2021); CEIC (2021)

The ITSA would provide several empirical results. First, it would show the baseline conditions before the introduction of the policy – that is, the baseline level and trend (or slope) of the tax ratio before FDRS adoption. Second, it would show the condition in the period immediately after the introduction of the policy relative to the baseline – i.e., relative change in the level of tax ratio immediately after FDRS adoption. Third, it would show the long-term condition in period after the introduction of the policy relative to the baseline – i.e., relative change in the slope of tax ratio. Fourth, the ITSA would provide assessment on the long-term impact of the policy (i.e., the post-intervention slope of tax ratio after FDRS).

The model was structured as follows:

$$Y_t = \beta_0 + \beta_1 T_t + \beta_2 X_t + \beta_3 X_t T_t + \epsilon_t \quad (1)$$

here, Y_t was the dependent variable of tax ratio at time t and T_t was the time elapsed since the start of the study. X_t was a dummy variable representing the FDRS introduction; with 0 represented periods before the use of FDRS (before Quarter 2 – 2018) and 1 represented the periods after FDRS adoption (Quarter 2 – 2018 and beyond). $X_t T_t$ was an interaction term. β_0 denoted the intercept or starting level of the independent variable of tax ratio. β_1 denoted the trend (or slope) of tax ratio in the period before the introduction of FDRS. β_2 denoted change in the level of tax ratio that occurs in the period immediately after FDRS adoption. β_3 denoted the difference between trends in tax ratio in periods before and after the adoption of FDRS. Newey-West standard errors with a maximum lag of 1 was used for the autocorrelation structure and a p-value of less than 0.05 was regarded as statistically significant. The data were processed using ITSA statistical package provided by STATA.

5. Results

Table 2 summarized the regression results from ITSA and Figure 1 presents the visual display of these results.

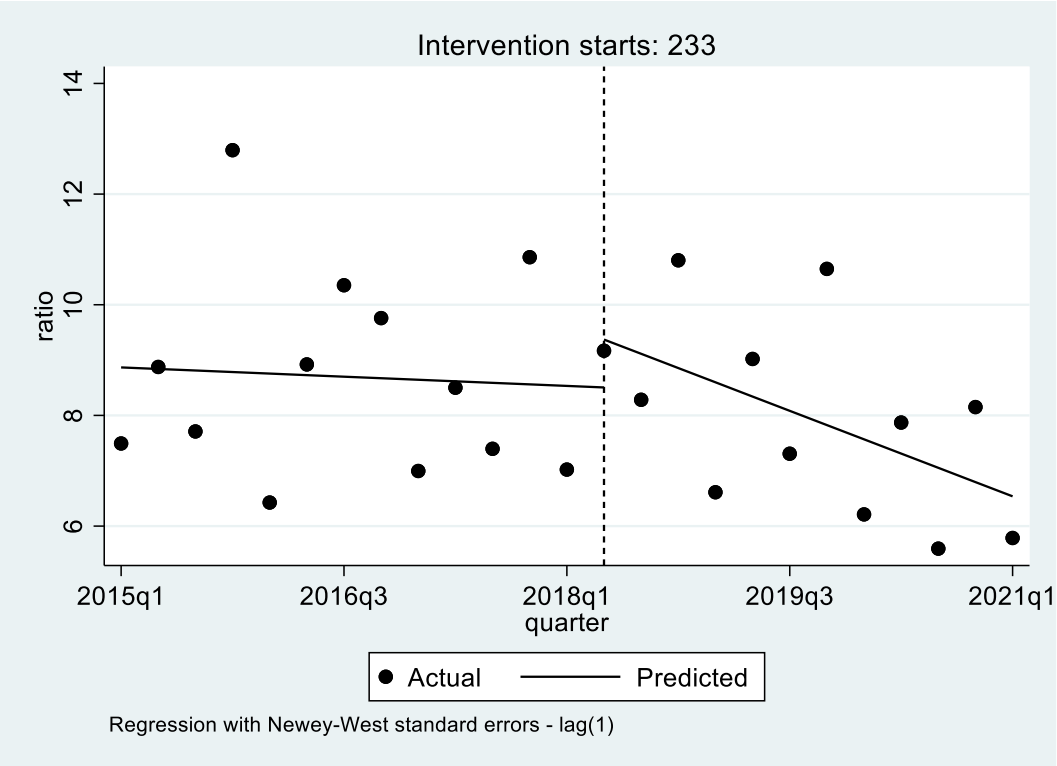
Table 2. Results from ITSA

Study period	Baseline level (β_0)	Baseline slope (β_1)	Change in level	p-value	Change in slope	p-value	Post-intervention slope	p-value
2015q1 – 2018q1	8.89	-0.03	–	–	–	–	–	–
2018q2 – 2021q1	–	–	0.87	0.258	-0.23	0.044	-0.26	0.000

Source: Author

As shown in Table 2, the starting (or the baseline) level of the ratio of tax revenue to GDP was at 8.89%, and this ratio appears to decrease steadily prior to FDRS adoption at an estimated declining rate of 0.03 percentage points every quarter. In the first quarter immediately after FDRS adoption (Quarter 2 – 2018), there seems to be a significant increase in the level of tax ratio of 0.87 percentage points relative to the baseline. However, its p-value (0.258) is not statistically significant, implying that this

increase may not be able to be attributed to the introduction of FDRS in the tax authority.



Source: Author

Figure 1. Tax Ratio Before and After FDRS

Results from Table 2 also show that, relative to the pre-FDRS period, the trend of tax collection in post-FDRS period seems getting worse: in every quarter, the post-intervention tax ratio is 0.23 percentage points lower than the pre-intervention ratio. The p-value (0.044) is statistically significant, implying that this worsening performance after FDRS adoption may not occur by chance.

It could also be seen in Table 2 that after the introduction of FDRS, tax ratio persistently decreases at an estimated rate of 0.26 percentage points every quarter. This can be observed in Figure 1 where the negative slope post-FDRS is much steeper than the negative slope pre-FDRS. With a p-value of 0.000 it seems that deteriorations in the trend of tax collection after FDRS adoption may not happen by coincidence.

6. Discussion

Empirical results in this article show that DGT's collection performance after FDRS adoption is worse off than it was before the adoption. In this section, possible explanations for these results will be discussed under the framework of signaling theory. It may be important to firstly point out one of the key performance indicators (KPI) used to assess the achievements of DGT personnel. At DGT, regional and district managers are responsible for collecting certain amount of tax revenue. This revenue target or quota is included in the managers' KPI and thus how much of the allotted target can be achieved may determine these managers' ranks according to the FDRS. Under this condition, a manager has the incentives to persistently send signals expecting subordinates to focus on accomplishing the manager's collection target. Since manager may have limited information on the true capacity of each of his subordinate, therefore he may rely on the observable characteristics of how each subordinate puts in efforts to achieve the collection target and rank these subordinates accordingly.

Aggressiveness toward taxpayers as performance signaling may emerge when there is widespread expectation among tax officials that the feedback from these signals are rewards (or punishments) associated with achieving (or failing to achieve) the collection targets. The adoption of FDRS for performance evaluation may lend credibility to this expectation. Once this expectation is in place, tax officials would be incentivized to signal their performance to their managers and the credibility of these signals would be directly related to the observable 'costs' incurred by these tax officials when they produce communicative signals of aggressiveness in collecting taxes.

These costs do not need to be purely material (Marquez, 2020). Engaging in aggressive behaviors toward taxpayers may cost the tax officer peer disapprovals as well as loss of trust from taxpayers. These costs could exacerbate when the revenue target is grossly incongruent with the real financial capacity and the actual legal liability of the taxpayers hence treating taxpayers in aggressive manners may fail to result in desired payments. Although aggressive tactics may not always result in actual tax payments, however, the aggressiveness may be received by manager as

credible signal that the tax officer possesses commitment toward achieving the collection target. To signal aggressiveness, tax officers can threaten taxpayers with higher taxes than the true unless taxpayers revise their tax returns and included in the revisions increases in tax payments. Taxpayers may perceive this threat to be credible because tax officers typically have various devices to over-state the payable taxes: e.g., tax officers may challenge or disallow legitimate deductions, they may charge tax on non-taxable transactions, and tax regulations may be interpreted in ways that favor the tax officers. Taxpayer rights may be suppressed to 'save' current year's collection target: e.g., claims for tax refunds may be delayed until the next fiscal year and taxpayer's request for transfer to other district may be denied because it may erode the revenue of the tax office in the original district.

Tax officials may not receive direct financial gains from these aggressive tactics (for example, through bribery). However, under the FDRS employees who sit at the top ranks will have the opportunities to be promoted; and with these promotions follow increases in monthly salaries which can reach twice the initial levels as well as facilities that are only available for officials at higher echelons. These may serve as strong incentives for tax officials to send signals to their supervisors that their performance (i.e., aggressiveness) are worth rewarded with the top ranks. As tax officials compete to signal their performances there may be a process of 'aggressiveness inflation' in which these officials try to outcompete each other by sending signals which indicate that they are more aggressive than others; this would make supervisors to demand their subordinates to employ more aggressive tactics toward taxpayers as the price for the top tier ranks in FDRS. In this inflation process, signals of moderate level of aggressiveness may not be enough since other tax officers may have sent the same signals at the same moderate level, therefore an inflated level of aggressiveness may be needed to stand out in the competition.

The main cost of aggressiveness as performance signaling is the diminishing trust taxpayers put in tax authority. Large body of research has emphasized the importance of trust relations between citizens (or taxpayers) and the government (including the tax authority) – see, for example, in Feld and Frey (2002). The level of tax compliance is the result of the dynamic interactions between the power held by tax authority and

taxpayers' trust in tax authority (Kirchler, Hoelzl, & Wahl, 2008). With its legal power, tax authority can enforce compliance mainly through the threat of punishment; however, this approach is not always effective in creating and maintaining compliant behavior from taxpayers. On the other hand, a high level of trust put by taxpayers in tax authority may increase voluntary compliance and minimize the need of costly audits. High level of trust may motivate taxpayers to pay their share of taxes as fair contributions to public goods even when the tax authority is unable to exercise complete supervision to each individual taxpayer. If tax officers treat taxpayers as equal partners – not as would be 'robbers' who would always try to avoid paying taxes – and these tax officers play their role as advisors for taxpayers in navigating the complexity of tax regulations then the taxpayers may perceive these treatments and relations as fair and trustworthy; they may therefore voluntarily respond with reciprocity and compliance. On the contrary, when taxpayers distrust the tax authority, they may continuously engage in rational decision making: weighing the benefits gained from evading taxes vis-à-vis the probabilities for detection, and evade whenever the former outweighs the latter.

In the perspective of taxpayers who have to encounter aggressive treatments from tax officers, there are at least two possible responses available to them: fight or flight. With the fight response taxpayers may choose the legal approaches by submitting formal objections to tax authority and appeal to tax court. This fight option, however, is costly in terms of money and time. Small and medium taxpayers may not have adequate financial resources to hire tax lawyers and consultants to represent them in the complex and lengthy objection processes and hearing sessions in tax court.

When taxpayers perceive that the (legal) fight response is too costly then they may choose the flight response: that is, to move out of the formal economy and moving into the underground economy – which is typically hard to tax (Alm, Martinez-Vazquez, & Schneider, 2004). Contrary to popular believe, in terms of taxation taxpayers are not necessarily a powerless subject. The always available option of moving into the underground economy – although its propensity may vary across nations – gives taxpayers the power to counter, albeit illegally, against burdensome taxation as well as against aggressive, unfair treatments and harassments from tax

officials. In the perspective of these taxpayers, engaging in the underground economy may serve two purposes: firstly, they would be able to continue their operations and maintain their sources of income and, secondly, to avoid the costly efforts of legally challenging the aggressive tactics from tax officials. Even for taxpayers who cannot completely escape to the underground economy – because of their social status as prominent people, for example – they still have the option of playing on both sides: one business may stay in the formal economy so as not to attract too much attention from the tax authority if it disappears completely and one business operates in the underground economy so as to minimize the compliance costs emerging from the aggressive behavior of tax officials.

Hence, tax officials' aggressiveness as demonstration of performance may induce a vicious cycle: as the tax base is shrinking because taxpayers are moving into the underground economy, tax revenue will decline thus widening the budget deficits; these conditions will put pressures on tax authority to increase collections and to reward those who are able to achieve the increasing collection targets; this will further incentivize tax officers to signal their performance by behaving more aggressively, hence encouraging more taxpayers to escape to the underground economy and worsening the problems of shrinking tax base, declining tax collections, and the widening budget deficits.

Moreover, the pandemic of Covid-19 may exacerbate this vicious cycle. The lockdown imposed by the government to curb the pandemic may cause businesses to experience financial hardships hence adversely affect their ability to pay the taxes due. Although there are various measures of tax facilities intended for taxpayers affected by the pandemic, however, the FDRS still applies hence this performance evaluation system may still incentivize tax officials to act aggressively to signal their supervisors that they can maintain their performance even when taxpayers experience financial constraints due to the pandemic. This condition may encourage more taxpayers to migrate to the underground economy thus shrinking the tax base and depressing the collection even more.

All this does not mean that non-compliance with tax laws should be left unchecked. Instead, detection of and punishment for non-compliance are important

elements in building taxpayers' trust in the taxation system. Taxpayers would put their trust in the taxation system if they believe that the tax authority has the capacity to detect and punish non-compliance effectively thus ensuring that all citizens are paying their fair share of taxes. This paper, however, suggests that tax laws should be enforced in an environment of 'service and client' attitude, in which the tax authority is intent on delivering fair, transparent, respectful, and supportive treatment for taxpayers (Kirchler et al., 2008). In this context, it may be important for tax authorities to consider exploring and implementing more suitable employee performance evaluation systems which, unlike the FDRS, are conducive toward creating and maintaining trustworthy relationships between taxpayers and tax officers.

7. Conclusion

Using Indonesian tax authority as case study, empirical results under the interrupted time-series analysis found that the adoption of forced distribution rating system (FDRS) for evaluating the performance of tax officials may have worsened the already declining trend in tax collection. Before the introduction of FDRS, the ratio of tax collection to gross domestic product has a moderately declining trend; the slope of the decline, however, is getting steeper after the adoption of FDRS.

This article provides alternative explanation for the mechanism by which this deterioration occurs using the perspectives provided by the signaling theory. It suggests that one of the possible sources for this worsening collection performance after FDRS adoption may be the migration of taxpayers to the underground economy. This migration may be encouraged by the aggressive behaviors of tax officials and this aggressive behavior could be the result of the introduction of the FDRS. The combination of FDRS for performance evaluation and the system of collection targets may incentivize tax officials to employ aggressive tactics toward taxpayers as signals for their supervisors that they are committed to achieving the collection targets and thus are worthy to be rewarded with the top ranks available under the FDRS.

One of the costs of these aggressive behaviors are erosions in the trust taxpayers put in tax authority. When taxpayers cannot trust that the tax authority would be willing to solve tax disputes in fair manners and it is too costly for taxpayers to solve

tax disputes through the legal systems, the taxpayers may choose to move out of the formal economy and moving into the underground economy. Hence, there may be a vicious cycle: as the tax base is shrinking because taxpayers are going underground tax revenue will decline, thus prompting tax officers to signal their performance by behaving more aggressively toward taxpayers hence encouraging more taxpayers to escape to the underground economy. This article suggests that it may be important for tax authorities to explore and implement more suitable performance evaluation systems which, unlike the FDRS, are conducive toward creating and maintaining trustworthy relationships between taxpayers and tax officers.

References

Adsit, D. J., Bobrow, W. S., Hegel, P. S., & Fitzpatrick, B. G. (2018). The return on investment of rank and yank in a simulated call-center environment. *Consulting Psychology Journal: Practice Research, 70*(2), 113.

Alm, J. (2019). Can Indonesia Reform Its Tax System? Problems And Options. *Journal of Poverty Alleviation and International Development, 10*, 71-99.

Alm, J., Martinez-Vazquez, J., & Schneider, F. (2004). 'Sizing'the Problem of the Hard-to-Tax. *Contributions to Economic Analysis, 268*, 11-75.

Badan Pusat Statistik. (2019). *Indikator Ekonomi Desember 2019*.

Badan Pusat Statistik. (2021). *Indikator Ekonomi April 2021*.

Bangerter, A., Roulin, N., & König, C. J. (2012). Personnel selection as a signaling game. *Journal of Applied Psychology, 97*(4), 719.

Berger, J. (2019). Signaling can increase consumers' willingness to pay for green products. Theoretical model and experimental evidence. *Journal of Consumer Behaviour, 18*(3), 233-246.

Berger, J., Harbring, C., & Sliwka, D. (2013). Performance appraisals and the impact of forced distribution—an experimental investigation. *Management Science, 59*(1), 54-68.

Bergh, D. D., Connelly, B. L., Ketchen Jr, D. J., & Shannon, L. M. (2014). Signalling theory and equilibrium in strategic management research: An assessment and a research agenda. *Journal of Management Studies, 51*(8), 1334-1360.

Bird, R. M. (2004). Administrative Dimensions of Tax Reform. *Asia-Pacific Tax Bulletin, 10*(3), 134-150.

Blume, B. D., Rubin, R. S., & Baldwin, T. T. (2013). Who is attracted to an organisation using a forced distribution performance management system? *Human Resource Management Journal, 23*(4), 360-378.

Boswell, W. R., & Boudreau, J. W. (2002). Separating the developmental and evaluative performance appraisal uses. *Journal of Business Psychology, 16*(3), 391-412.

Bretz Jr, R. D., Milkovich, G. T., & Read, W. (1992). The current state of performance appraisal research and practice: Concerns, directions, and implications. *Journal of management, 18*(2), 321-352.

Briker, R., Hohmann, S., Walter, F., Lam, C. K., & Zhang, Y. (2021). Formal supervisors' role in stimulating team members' informal leader emergence: Supervisor and member status as critical moderators. *Journal of Organizational Behavior*.

Cable, D. M., & Judge, T. A. (1994). Pay preferences and job search decisions: A person-organization fit perspective. *Personnel Psychology, 47*(2), 317-348.

Cappelli, P. (2009). What's old is new again: Managerial "talent" in an historical context. In *Research in personnel and human resources management*: Emerald Group Publishing Limited.

CEIC. (2021).

Celani, A., & Singh, P. (2011). Signaling theory and applicant attraction outcomes. *Personnel Review, 40*(2), 222-238. doi:10.1108/00483481111106093

Chattopadhayay, R., & Ghosh, A. K. (2012). Performance appraisal based on a forced distribution system: its drawbacks and remedies. *International Journal of Productivity Performance Management*.

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of management, 37*(1), 39-67.

Costa, P. T. J., & McCrae, R. R. (2008). The Revised Neo Personality Inventory (NEO-PI-R). In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment, Vol. 2. Personality measurement and testing* (pp. 179 - 198): Sage Publications, Inc.

Direktorat Jenderal Pajak. (2019). *Laporan Tahunan DJP 2019*. Retrieved from Jakarta: https://www.pajak.go.id/sites/default/files/2020-12/Laporan%20Tahunan%20DJP%202019%20-%20INDONESIA.pdf

Ehrhart, M. G., & Klein, K. J. (2001). Predicting followers' preferences for charismatic leadership: The influence of follower values and personality. *The Leadership Quarterly, 12*(2), 153-179.

Eichenwald, K. (2012). Microsoft's Lost Decade. *Vanity Fair*.

Feld, L. P., & Frey, B. (2002). Trust breeds trust: How taxpayers are treated. *Centre for Tax System Integrity Working Paper No. 32 Australian National University and Australian Taxation Office*.

Felfe, J., & Schyns, B. (2006). Personality and the Perception of Transformational Leadership: The Impact of Extraversion, Neuroticism, Personal Need for Structure, and Occupational Self-Efficacy 1. *Journal of Applied Social Psychology, 36*(3), 708-739.

Hammar, H., Jagers, S. C., & Nordblom, K. (2009). Perceived tax evasion and the importance of trust. *The Journal of Socio-Economics, 38*(2), 238-245.

Iswahyudi, H. (2020a). Have Diminishing Returns Set in? The Efficiency of Indonesian Tax Administration. *Bisnis & Birokrasi Journal, 26*(3).

Iswahyudi, H. (2020b). *The problem of gross receipts taxes in Indonesia: Economic distortions and policy options*. [Working Paper]. Munich Personal RePEc Archive.

Jenkins Jr, G. D., Mitra, A., Gupta, N., & Shaw, J. D. (1998). Are financial incentives related to performance? A meta-analytic review of empirical research. *Journal of Applied Psychology, 83*(5), 777.

Kaplanoglou, G., & Rapanos, V. T. (2013). Tax and trust: The fiscal crisis in Greece. *South European Society Politics, 18*(3), 283-304.

Kharouf, H., Lund, D. J., Krallman, A., & Pullig, C. (2020). A signaling theory approach to relationship recovery. *European Journal of Marketing, 54*(9), 2139-2170. doi:10.1108/EJM-10-2019-0751

Kirchler, E., Hoelzl, E., & Wahl, I. (2008). Enforced versus voluntary tax compliance: The "slippery slope" framework. *Journal of Economic Psychology, 29*(2), 210-225.

Kwoh, L. (2012). Rank and yank" retains vocal fans. *The Wall Street Journal, 31*, 2012.

Lawler, E. E., III. (2002). The Folly of Forced Ranking. *Strategy+Business, Third Quarter 2002*.

Li, J., Tang, J., Jiang, L., Yen, D. C., & Liu, X. (2019). Economic Success of Physicians in the Online Consultation Market: A Signaling Theory Perspective. *International Journal of Electronic Commerce, 23*(2), 244-271. doi:10.1080/10864415.2018.1564552

Linden, A. (2015). Conducting interrupted time-series analysis for single-and multiple-group comparisons. *The Stata Journal, 15*(2), 480-500.

Linden, A. (2017). A comprehensive set of postestimation measures to enrich interrupted time-series analysis. *The Stata Journal, 17*(1), 73-88.

Magee, J. C., & Galinsky, A. D. (2008). 8 social hierarchy: The self-reinforcing nature of power and status. *Academy of Management Annals, 2*(1), 351-398.

Marquez, X. (2020). The Mechanisms of Cult Production: An Overview. In K. Postoutenko & D. Stephanov (Eds.), *Ruler Personality Cults from Empires to Nation-States and Beyond: Symbolic Patterns and Interactional Dynamics*: Forthcoming.

Martin, G., & Groen-in't Woud, S. (2011). Employer branding and corporate reputation management in global companies. In H. Scullion & D. G. Collings (Eds.), *Global Talent Management* (pp. 87-110). London: Routledge.

Ministry of Finance. (2021). Reformasi Birokrasi Untuk Kesejahteraan Masyarakat. https://www.kemenkeu.go.id/sites/default/files/artikel_pajak_2509.pdf

Moon, S. H., Scullen, S. E., & Latham, G. P. (2016). Precarious curve ahead: The effects of forced distribution rating systems on job performance. *Human Resource Management Review, 26*(2), 166-179.

Osborne, T., & McCann, L. (2004). Forced Ranking and Age-Related Employment Discrimination. *Human Rights Magazines*.

Ovide, S., & Feintzeig, R. (2013). Microsoft Abandons 'Stack Ranking' of Employees. *The Wall Street Journal*.

Peraturan Direktur Jenderal Pajak Nomor Per - 12/PJ/2018 tentang Pengelolaan Kinerja di Lingkungan Direktorat Jenderal Pajak, (2018).

Pfeffer, J., & Sutton, R. I. (2006). Evidence-based management. *Harvard business review, 84*(1), 62.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879.

Rynes, S. L., Gerhart, B., & Parks, L. (2005). Personnel psychology: Performance evaluation and pay for performance. *Annual Review Psychology, 56*, 571-600.

Schleicher, D. J., Bull, R. A., & Green, S. G. (2009). Rater reactions to forced distribution rating systems. *Journal of management, 35*(4), 899-927.

Shin, S. J., & Zhou, J. (2003). Transformational leadership, conservation, and creativity: Evidence from Korea. *Academy of Management Journal, 46*(6), 703-714.

Spence, M. (2002). Signaling in retrospect and the informational structure of markets. *American Economic Review, 92*(3), 434-459.

Stiglitz, J. E. (2002). Information and the Change in the Paradigm in Economics. *American Economic Review, 92*(3), 460-501.

Taj, S. A. (2016). Application of signaling theory in management research: Addressing major gaps in theory. *European Management Journal, 34*(4), 338-348.

Thomas, R., Darby, J. L., Dobrzykowski, D., & van Hoek, R. (2020). Decomposing social sustainability: Signaling theory insights into supplier selection decisions. *Journal of Supply Chain Management*.

Trank, C. Q., Rynes, S. L., & Bretz, R. D. (2002). Attracting applicants in the war for talent: Differences in work preferences among high achievers. *Journal of Business Psychology, 16*(3), 331-345.

Turban, D. B., & Keon, T. L. (1993). Organizational attractiveness: An interactionist perspective. *Journal of Applied Psychology, 78*(2), 184.
Welch, J., & Byrne, J. A. (2003). *Jack: Straight from the Gut*. New York: Grand Central Publishing.
Welch, J., & Welch, S. (2005). *Winning*. New York, NY: Harper Business.
Whiting, H. J., & Kline, T. J. B. (2007). Testing a model of performance appraisal fit on attitudinal outcomes. *The Psychologist-Manager Journal, 10*(2), 127-148.