Tax administration resources and Income inequality

Nguyen, Trang T.T.

Department of Business, Universitat Autònoma de Barcelona, Spain

30 October 2016

TAX ADMINISTRATION RESOURCES AND INCOME INEQUALITY[☆]

Trang T.T Nguyen[a,*]

[a] *Department of Business, Universitat Autònoma de Barcelona, Spain*

Abstract

The paper contributes to the research field as the first quantitative study deploying OECD Tax administration database to provide empirical evidence with respect to the impact of Tax administration resources on income inequality. Methodologically speaking, we apply, for the unique data set of 46 countries extracted from OECD Tax administration 2015 database, the cross-sectional multiple regression employing the OLS estimator before justifying with other estimators i.e. truncated regression, quantile regression and weighted least squared (WLS). We finally obtain the robustly negative relationship between Tax administration resources (employees and non-salary expenditure) and income inequality as hypothesize. The paper ends with some policy recommendations following the limitations and directions for further research.

JEL Classification: D63, H20.

Keywords: income inequality, Tax administration, OECD Tax administration database, Tax administration resources, Tax administration employees, non-salary expenditure.

[☆] Preliminary Draft, please do not cite or circulate without authors' permission.

[*] Corresponding author.

Email addresses: trang-ntt@hotmail.com (Trang T.T Nguyen), trangntt.uab@gmail.com (Trang T.T Nguyen)

ABBREVIATIONS (Abbr.)

OECD	Organization for Economic Co-operation and Development
EU	European Union
G20	Group of Twenty - an international forum for the governments and central bank governors from 20 major economies.
WDI	World Development Indicators
WIID	World Income Inequality database
WGI	Worldwide Governance Indicators
GINI	Gini coefficients extracted from World Development Indicators (WDI)
GINI_WIID	Gini coefficients extracted from World Income Inequality Database
TE	Tax employee
NSE	Non-salary expenditure
RL	Rule of law
GE	Government effectiveness
CC	Corruption control
RQ	Regulatory quality
UE	Unemployment
NC	Natural capital
ILO	International Labor Organization
GPDpc	GPD per capita
OLS	Ordinary least squared

1. Introduction

Inequality in general and inequality in income distribution together with the way how inequality is generated and operated in particular has been considered a major concern and drawn the remarkable attention of researchers, economists and policy makers for decades. It can be easily seen a considerable number of papers working on this research matter and contributing to the research field with different inequality determinants i.e. institutional quality (Chong & Gradstein, 2007; Fuentes et al., 2014; Perera & Lee, 2013...), unemployment rate (Galbraith et al., 1999; Aaberge et al., 2000; Helpman et al., 2010...), economic growth (Kuznets, 1955; Aghion et al., 1999; Barro, 2000...), educational attainment level (Gregorio & Lee, 2002; Yang & Qiu, 2016; Breen & Chung, 2015...), natural resource endowment (Auty, 1997; Gylfason & Zoega, 2002; Fum & Hodler, 2010; Carmignani, 2013...) and so on. However, the relationship between income inequality and Tax administration resources seems far from being well understood. And it is likely that there have not been found any available empirical paper mentioning about the role of Tax administration resources to income distribution inequality.

In order to close this gap in the research field, the paper makes a combination of literature on GINI determinants and Taxation administration for elaborating the contribution and effect of Tax administration resources on income equality (or "social welfare"). To some certain extent, it seems the fact that Tax administration as a part of "service paradigm" accompanied with traditional "enforcement paradigm" and "trust paradigm" might help serve as a large set of alternative strategies to reduce Tax evasion (Alm, 2012). This outcome in case of being well accompanied with efficiency in Tax transfer in the areas of Tax governance, enforcement and compliance would be expected to diminish or even avoid the "leaky bucket" on the process of income redistribution for getting more equality and better social welfare (Browning, 2008) by deteriorating the deadweight loss. In addition, as stated by Chong and Calderon (2000), institutional framework is linked with how a society's gains are distributed and redistributed among its members, it seems the institutional quality plays an indispensable role in income distribution.

The purpose of this paper is to review the literature dealing with income inequality to investigate the other potential determinants besides the existing ones that may explain this phenomenon. In details, it is to explore the effect of Tax administration with regards to resources employment i.e. Tax administration-related employees and non-salary expenditure on providing the necessary condition on Income inequality reduction, and to examine the effect of institutional factors on explaining income inequality across countries. These are actualized by finding the answers to the following questions saying "Do number of Tax administration employees relate negatively to the likelihood of income inequality?", "How is income inequality different among countries, in terms of non-salary expenditure?", and "Does institutional quality matter for income distribution inequality? If so, how?" And after that, we would make some policy recommendations for the improvement of income

inequality to achieve a better social welfare.

The study contributes to social welfare in relation to income redistribution, Tax transfer efficiency, and Taxation with the stress on Tax administration literature in some ways. This research paper aims to identify the new and potential factors that might help explain the inequality of income distribution along with the existing determinants. To add on, by using OECD Tax administartion database and presenting cross-country evidence, it is the first quantitative empirical study, to the best of our knowledge, to bring to the field of research with the role of Tax administration on account of resources allocation, especially the number of Tax employees and non-salary expenditure of Tax administration-related functions in reduction of likelihood of income inequality throughout the World, including both OECD and non-OECD countries. Additionally, the contribution of institutional factors to predicting the phenomenon is also mentioned.

The rest of the paper is organized as follows. The next section is about theoretical backgrounds in which we review the relevant papers and literature underlying reasons for proposed hypotheses. The methodology is discussed in the third section before the results shown in section 4. Then, section 5 deals with discussion and conclusion, followed by some limitations and future research in section 6. The last section ends the paper with some implications for policy makers.

2. Theoretical background and hypotheses

The paper generally discusses about the determinants of income inequality with the special focus on the impact of Tax administration resources on income inequality as a proxy of social welfare based on income redistribution, besides the review of existing literature on relevant determinants of income inequality underlying reasons for proposing the hypotheses.

2.1. Inequality and Tax administration

It was argued by Leach (2004) that competitive markets have considerably devoted to material well-being, but they are always with their flaws. A better level of material welfare might thus be obtained if the government operates in consort with the market system in order to mend these flaws and moderate its harsher tendencies. On the one hand, the government tends to repair the market for improving the resource allocation (first theorem). On the other hand, it is the redistribution of income (second theorem). However, it was also stated that if a Pareto optimal (or "first best") allocation cannot be reached, policy decisions revolve around making appropriate compromises. And, the best available compromise is believed to be the "second best" allocation.

The second theorem argues that if the government wish to achieve a more equitable distribution of income, it need only transfer income from the economically advantaged to the economically disadvantaged i.e. the Government tax those who are initially earning the high wage and subsidize

those who are initially earning the low wage. Taxes, for example, are levied on the purchases of goods and the receipt of income. And welfare payments are made to those who have no other source of income. It is the redistribution of income in favor of social justice and efficiency (Leach, 2004). In other words, redistribution can be accomplished without a loss of economic efficiency.[1] However, in reality, redistributive policies often entail a loss of economic efficiency. This optimal redistribution holds if and only if the redistribution is lump-sum ("first best") which is impossible due to incentive incompatibility. A deadweight loss will doubtlessly result from any attempt to redistribute income. The government can then reduce efficiency losses but cannot entirely eliminate them. As a result, it was suggested that a government, which is intent on redistribution, should design its policies including a tax system with the correct degree of progressivity as "second best" through tagging and targeting policies (Leach, 2004). Because one of the effects of a progressive income tax is to reduce inequality in the distribution of income (Creedy, 1994). Taxes are, according to Wilkinson (1992), also used to redistribute income; and, for Keynesians, they take a part in demand management in the society, added to the role of allocation resources. Over and above, it is suggested that the design of redistributive policies is extremely important to the extent of the negative link between the deadweight loss associated with income transfer between the rich and the poor. In this case, tax policy could play a main role in making the post-tax income distribution less unequal. And, tax policy is crucial for raising revenues to finance public expenditure on transfers, health and education that tend to favor low-income households, to the same extent as on growth-enabling infrastructure that can also increase social equity (Carter & Matthews, 2012). All of these afore-discussed help indicate the close relationship between taxation and equality in income distribution as argued that "the income tax is more efficient than the legal system in redistributing income" (Kaplan & Shavell, 1994). More importantly, it was declared that tax and transfer systems can reduce overall income inequality in all countries. On average across the OECD, three quarters of the reduction in inequality is due to transfers while the rest to direct household taxation (OECD, 2012).

To the extent of Taxation and Tax administration, it is said that tax is one of two only things certain in life and at least far from inevitable, resulting in a variety of actions to reduce the tax liabilities (Alm, 2012). Tax avoidance is acknowledged to be the change of the behavior in order to reduce the tax liability; there is thus nothing illegal and it was called "the only intellectual pursuit that carries any reward" by John Maynard Keynes (Rosen & Gayer, 2009). On the contrary, tax evasion is failing to pay legally due tax. It is regarded as illegal and intentional actions taken by individuals to decrease their legally due tax obligations. And, it has been found widespread everywhere all over the World (Alm, 2012), because the lump-sum tax is believed to be the only tax that cannot be evaded, resulting from being not linked to anything within the control of the taxpayer (Leach, 2004).

[1] See Bakija, J. (2014). The deadweight loss (net reduction in aggregate economic surplus in society) from taxation means that redistribution from rich to poor makes the rich worse off by more than it makes the poor better off when measured in currencies.

Tax administration deals with collecting and processing tax revenue. As being said by Leach (2004) that "everyone would be better off if no one attempted to evade the tax", tax transfer efficiency (or "income redistribution") will therefore spend the concentration on tax evasion as an evidence, the tax governance, compliance and enforcement, and the efficiency in managing the transfer bucket as optimal by limiting the leakage on the way as well, so that redistributive bucket is truly beyond being only a sieve before embarking on a further expansion in welfare policies (Browning, 2008). The things that people do to avoid paying taxes alter the allocation of resources in ways that reduce economic welfare (Leach, 2004). In this sense, it is implied that the Tax administration is of course highly correlated with the effectiveness and efficiency of Taxation through the role of income redistribution. Put differently, the effectiveness of Tax administration will perhaps lead to the income transfer efficiency with the aim of retaining a greater social welfare (or "income equality") by a minimal deadweight loss. Taken together, Tax administration might be expected to have the influence on social welfare or income distribution of society.

With a view to Tax administration operations, it always needs costs of running the tax system i.e. the costs of administering the income tax because no tax system is costless to administer (Rosen & Gayer, 2009). Tax administrative costs are the costs of collecting and processing tax revenue (Salanie, 2003). Enhancing tax administration effectiveness entails improvement of tax compliance and enforcement, but lowering collection costs (OECD, 2012). And, Tax administration as "service paradigm" together with traditional "enforcement paradigm" and "trust paradigm" will form a large set of alternative strategies to reduce tax evasion (Alm, 2008).

For achieving this target, it seems clear for the demand of resources facilitating the effectiveness. It is said that the overall level of resources devoted to tax system administration is a critical and thematic issue for governments, revenue organizations and external observers (OECD, 2015).

Moreover, being recorded as the most popular inequality index that is mostly associated with the descriptive approach to inequality measurement (Bellù & Liberati, 2006), the Gini coefficient is well established as a conventional, ad hoc measure of income inequality (Dorfman, 1979). These things might cause Gini coefficient to be used as a proxy for social welfare.

In brief, Tax administration in view of resources allocation might be used with Tax employees (proxy for tax inspectors) and non-salary expenditure i.e. the share of Tax administration's aggregate expenditure with the exception of salary costs. And, it is our hypothesis that:

Hypothesis 1. The tax administration resources are negatively associated with the likelihood of income distribution inequality across countries.

It appears reasonable and logical to expect that the appropriate and remarkable number of Tax administration staffs may contribute to higher income equality. We would thus propose that:

Hypothesis 1.a. The higher number of Tax administration-related employees, the lower rate of income inequality (lower Gini coefficient) across countries

Aside from the salary costs, it is necessary for governments to make the appropriate allocation of non-salary expenditure. It is well recognized that IT, HRM, infrastructure ... have attracted the increasingly large share among resources for the effectiveness and efficiency targets. It was found in Carter and Matthews (2012) that the use of new IT systems in revenue administrations increasingly embrace instruments such as sophisticated risk engines to explore potential missing revenues. Also, efforts to curb offshore non-compliance by information exchange among tax authorities more effectively have been given a new impetus. With the same token, tax evaders, who are often wealthy, have fewer places to hide their money.

It is our expectation to find the negative relationship between Tax administration's non-salary expenditure and the propensity of income inequality. Hence, we predict that:

Hypothesis 1.b. The higher proportion of non-salary expenditure in Tax administration, the lower rate of income inequality (lower Gini coefficient) across countries

2.2. Inequality and other relevant variables

Besides Tax administration resources as the determinant of income inequality, as regards the existing literature on inequality, there have been accumulated a wide range of literature dealing with various inequality determinants and sources.

It can be seen that Browning and Johnson (1984) discussed about the tradeoff between equity and efficiency, and that "the marginal cost of reducing income inequality with a policy that has distributional effects similar to the present tax-transfer system is quite high even for modest labor supply elasticities". Whereas, it was stated "Income inequality leads to increased mortality via disinvestment in social capital" (Kawachi et al., 1997). Sarel (1997) developed a cross-section empirical framework to examine the relationship between macroeconomic environment as well as the trends in income distribution. And, it was found higher growth rate, higher income level, higher investment rate, real depreciation, and improvement in terms of trade would be associated with improvement in income distribution.

Jong-Sung and Khagram (2005) considered the increase of level of corruption due to income inequality through material and normative mechanisms.

Recently, Lee et al. (2013) reviewed the trends and identified the determinants of income inequality in Korea from 1980 to 2012 in which they discovered the impact of macroeconomic index such as the government spending, investment share, population are the important factors in widening the gap of income inequality. However, both Kuznets' hypothesis of inverted U-shaped relationship between income inequality and economic growth and Barro's hypothesis of U-shaped relationship are not empirically supported.

Yunker (2016) dealt with this issue on a different way. He mentioned the overall relationship between economic inequality and optimal redistribution, as well as specific insights into the effect of various economic parameters on this relationship.

In addition to these generally income inequality- related issues, there has been found with a great amount of literature dealing with particular aspects of income inequality as determinants.

2.2.1. Inequality and institutions

It is said that the institutional framework is linked with not only economic performance, but also how a society's gains are distributed and redistributed among its members, that is, with the distribution of income (Chong & Calderon, 2000).

First of all, talking about the relationship between income inequality and institutional quality, Chong and Calderon (2000) presented the first systematic cross-section empirical research between institutions and the distribution of income. It was generally found a negative correlation, say, the lower the quality of the institutions, the higher the degree of income inequality. However, it was also stated that institutional quality displayed a significantly and robustly quadratic relationship with income inequality. To be specific, for poor countries institutional quality was found positively linked with income inequality, but the opposite effect for rich countries.

Meanwhile, the paper "Inequality and institutions" by Chong and Gradstein (2007) presented "theory and evidence on the relationship between inequality and institutional quality". The authors exhibited a model illustrating the mutual dynamical reinforcement for testing this relationship by sing a broad array of institutional measures. And it was empirically established with the double causality between a more equal distribution of income and institutional strength with employment of dynamic panel and linear feedback analysis.

Sharing the theme, Fuentes et al. (2014) argued that the deterioration of institutional quality influences skilled labor mobility across sectors (in developing economies) thereby decreasing (increasing) long term income inequality. And, there found the negative effect of institutional quality on long term income inequality.

Plus, Perera and Lee (2013) contributed to the research field with finding that "improvements in corruption, democratic accountability, and bureaucratic quality are associated with a worsening of the income distribution" in Asia. This led to recommendations of taking measures for improving the level of institutional quality in developing countries of East and South Asia on addressing the problems of poverty and income distribution.

Taking all of above-mentioned literature into account, we expect to see a negative relationship between income inequality and institutional quality. We then hypothesize that:

Hypothesis 2. The income distribution inequality (Gini coefficient) is greater in countries with lower institutional quality i.e. weaker rule and law

2.2.2. Inequality and unemployment

Next, the previous literature has demonstrated the association of inequality and unemployment. In "Inequality and Unemployment in Europe: The American Cure", it was concluded with the positive relationship between inequality and unemployment across the European continent, within countries, between countries and through time (Galbraith et al., 1999). On the one hand, inequality was considered to be the cause of unemployment as the expression of frustration with low-wage work. On the other hand, unemployment itself also brought inequality showing that in the periodic slumps and recoveries of the United States, this appeared to be the principal chain of cause and effect. To be in line with this research interest, Aaberge et al. (2000) paid attention to the link of unemployment shocks and income distribution with respect to "how the Nordic countries i.e. Denmark, Finland, Norway and Sweden fared during their crises? of late 1980s and early 1990s when unemployment rose dramatically. Anyhow, it was suggested that the Gini coefficient, a standard measure of inequality, was surprisingly stable in all these 4 countries during the referred time. In addition, it was Helpman et al. (2010) in "Inequality and Unemployment in a global economy" who developed a new framework for "examining the determinants of wage distributions emphasizing within-industry reallocation, labor market frictions, and differences in workforce composition across firms". In this paper, it was said that the opening of trade fostered wage inequality and could either raise or reduce unemployment. At the same time, wage inequality was found higher in a trade equilibrium than in autarky, gradual trade liberalization exaggerated at the beginning and later diminished the inequality.

With the literature on unemployment – inequality association, it is likely to be noted that unemployment ought to be treated as a determinant of Gini coefficient. This can thus facilitate the appropriateness for us to treat "unemployment" as a control variable in this paper.

2.2.3. Inequality and natural resources

As regards the literature on Gini's determinants, it should be noted with the link between inequality and natural resources that Academics have so far put into research. In fact, resource curse has drawn the considerable attention of academics, economists and policymakers in the sense that natural resource might be an economic curse instead of a blessing (Auty, 1997). It was pointed out by Leamer et al. (1999) that natural resource abundance might increase Latin American income inequality and income inequality was higher in this region than in East Asia. This was possible to explain by the fact that natural-resource-intensive sectors often absorb capital that might otherwise flow to manufacturing, resulting in depression of workers' incentive to accumulate skill and industrialization delays.

Being on the same platform, Gylfason and Zoega (2002) demonstrated in theory as well as empirically that the "increased dependence on natural resources tends to go along with less rapid economic growth

and greater inequality in the distribution of income across countries". And, natural resources were declared to "raise income inequality in ethnically polarized societies, but reduce income inequality in ethnically homogenous societies" (Fum & Hodler, 2010). Later, it was explained by Goderis and Malone (2011) about "the time path of income inequality following natural resource booms in resource-rich countries". In other words, they proposed and justified that inequality fell immediately after a boom, and then rose steadily over time until the disappearance of boom's initial impact under "the condition of a relatively unskilled labor intensive nontraded sector", especially for oil and mineral booms. Besides that, uncertainty about future commodity prices would increase long-run inequality. Also, Carmignani (2013) by the paper entitled "Development outcomes, resource abundance, and the transmission through inequality" conducted studies and found evidence supporting the positive effect of resource abundance on the inequality of income distribution within a country.

Taking together, the afore-discussed papers reveals the truth that, it may be worth using natural capital as control variable in our paper in the field that "The more increasing dependence on natural resources, the greater inequality in the distribution of income across countries".

2.2.4. Inequality and Economic growth

Among the relevant literature on determinants of income inequality, economic growth has been regarded as the classical factor. This is illustrated by Kuznets (1955) when showing character and causes of long- term changes in the personal distribution of income, particularly the potentially non-linear relationship between inequality and development (hereafter called Kuznets curve). Aghion et al. (1999) analyzed the perspective of the new growth theories on the basis of inequality - economic growth correlation. On examining the effect of inequality on growth, it was stated that it is not a necessity of a trade-off between equity and efficiency in case of imperfect capital markets. On the other hand, it was drawn attention with indicating several mechanisms whereby growth may raise wage inequality, both across and within education cohorts. At the same time, Barro (2000) contributed to the research field with evidence from a broad panel of countries showing little overall relation between income inequality and rates of growth i.e. higher inequality tends to hinder growth in poor countries and advance growth in richer places.

It seems undeniable that there occurs the impact of economic growth on income inequality, to some extent. That is hence the reason why we decide to take it into account in the form of an additional control variable.

2.2.5. Inequality and Education

As other potential determinants of income inequality, education has been acknowledged to be a piece of the issue. Gregorio and Lee (2002) raised the idea that education is related to income distribution. In more details, educational factors i.e. higher educational attainment and more equal distribution

of education were thought to play a significant role in advance the equality of income distribution. Furthermore, it was mentioned by Yang and Qiu (2016) about the impact of education on income inequality and intergenerational mobility. They argued that innate ability and family investment in early education take a crucial role in estimating income inequality and intergenerational income mobility. Nonetheless, Breen and Chung (2015) argued that the between-education contribution to inequality is small, even when it is considered with only adjusted inequality omitting the within-person component. As well, feasible educational policy was asserted to have only a limited impact on income inequality.

In brief, in light of the relevant literature entailing the relationship between education and inequality, it should not be doubtful reason to use education as a candidate for additional control variable in our paper.

3. Data and methodology

3.1. Sample and data collection

The paper, based on OECD Tax administration database 2015, initially employs a chosen set of 56 advanced and emerging economies including all 34 OECD and 22 non-OECD countries i.e. EU and G20 members. Due to missing values of Gini coefficients and non-salary expenditure, we ultimately work with 46 countries. The full list of these countries will be found in Appendix 1.

The unique data set is expected to collect with variables corresponding to regression and regressors from published sources, i.e. World Bank, WDI, WIID, WGI, and OECD.

3.2. Data description

The entire list of variables that have been applied with their sources in details can be found in Appendix 2. We consider, besides Tax administration resources variables - employees and non-salary expenditure, some explanatory variables with the role of "control variables" to observe the influences on social welfare in particular the income inequality. Here in this subsection, we would explain briefly the key variables that have been used.

3.2.1. Gini coefficients

Income inequality is estimated by Gini index. It is a uni-dimensional index measuring Inequality for socio-economic impact of policies through the extent to which the distribution of income (or, in some cases, consumption expenditure) among individuals or households within an economy deviates from a perfectly equal distribution (World Bank).

We deploy the Gini coefficients extracted from World Development Indicators (WDI) - World Bank estimate (hereafter called GINI) averaged over 5 years (2008-2012) to compensate for the scattering and insufficiency of Gini coefficients in some cases.

Furthermore, for the purpose of results' consistency check, we also use Gini coefficients basically extracted from World income inequality database - WIID (hereafter called GINI_WIID). Owing to the varying levels and sources of reports, and considerable missing values, we scan the available comparable data and then average the Gini coefficients over the afore-referred period from 2008 to 2012.

The Gini index, used as dependent variable, is of the interval between 0 and 1 in which 0 represents perfect equality, while an index of 1 implies perfect inequality.

3.2.2. Tax employee

The first independent variable, denoted Tax employee (TE), measures the number of Tax administration staffs over every 10,000 active citizens (aged 15-64) in which, the former is extracted from OECD Tax administration database and, and the latter is adopted from WDI (World Bank).

The paper utilizes the log Tax employee to see the elasticity of GINI with respect to the number of Tax employees.

3.2.3. Tax administration's non-salary expenditure

Non-salary expenditure (NSE), the second independent variable, represents the share of non-salary expenditure for Tax administration-related functions, covering infrastructures, information technology (IT), human resource management (HRM)... The variable is the share, calculated by the rest of aggregate expenditure for all Tax administration – related functions after total salary costs for all Tax administration – related functions i.e. 1 - (salary expenditure/aggregate expenditure). It is adopted from OECD Tax administration database.

3.2.4. Rule of law

Rule of law (RL), an institutional indicator, is used as a proxy for the quality of institutions that is extracted from Worldwide Governance Indicators (WGI). This index ranges from -2.5 to +2.5 where positive number indicates a higher quality.

In fact, all of four institutional indicators (rule of law - RL, government effectiveness - GE, corruption control - CC, and regulatory quality - RQ) have the same effect on GINI (in this model). And, they are highly correlated.[2] Therefore, we decide to use rule of law (RL) showing the strength of law and order as the best fit to be the representative of institutional quality index.[3]

[2]See Appendix 3. Correlation matrix among institutional variables

[3]See Appendix 4. Baseline regression model with different institutional indicators

3.2.5. Control variables

The control variables are unemployment (UE) and natural capital (NC).

UE refers to the rate of unemployment of total labor force, and is extracted from International Labor Organization (ILO), Key indicators of the labor market database WDI.

And, NC per capita used as a proxy for natural resources abundance demonstrates the log natural capital per capita (US$) in the most recent statistics of 2005. This variable is extracted from World Bank (2011).

3.2.6. Additional control variables

Besides, we also introduced some variables for controlling the omitted variables.

To evaluate the economic production and growth, we use the logarithm of GDPpc extracted from WDI. Also, in order to control the educational attainment, we acquire the average years of schooling corresponding to levels (total, primary, secondary, tertiary).

Figure 1 illustrates the distribution of main variables used in the paper.

3.3. Econometric methods and Model

3.3.1. Method

The paper is basically applied with Cross-sectional multiple regression employing the OLS estimator in Stata and then, for robustness check, it is justified with other estimators i.e. truncated regression and weighted least squared.

It is acknowledged that, in order to estimate a model with multiple outcomes, it can be utilized with multiple regressions for cross-sectional data to learn more about the relationship between several predictor variables and a criterion variable (Wooldridge, 2009).

With a unique data set of 46 countries, the explained variable – GINI and all explanatory and control variables (TE, NSE, RL, UE, NC, etc...) are associated with one period or point in time – 2000s, the choice of a cross-sectional regression appears to be appropriate. Due to the specific characteristics of the variables, especially GINI, this paper fails to address a time-series regression or longitudinal regression in which the variables are considered to be associated with a sequence of points in time to witness any change trend during time or causal relationship then (Wooldridge, 2009).

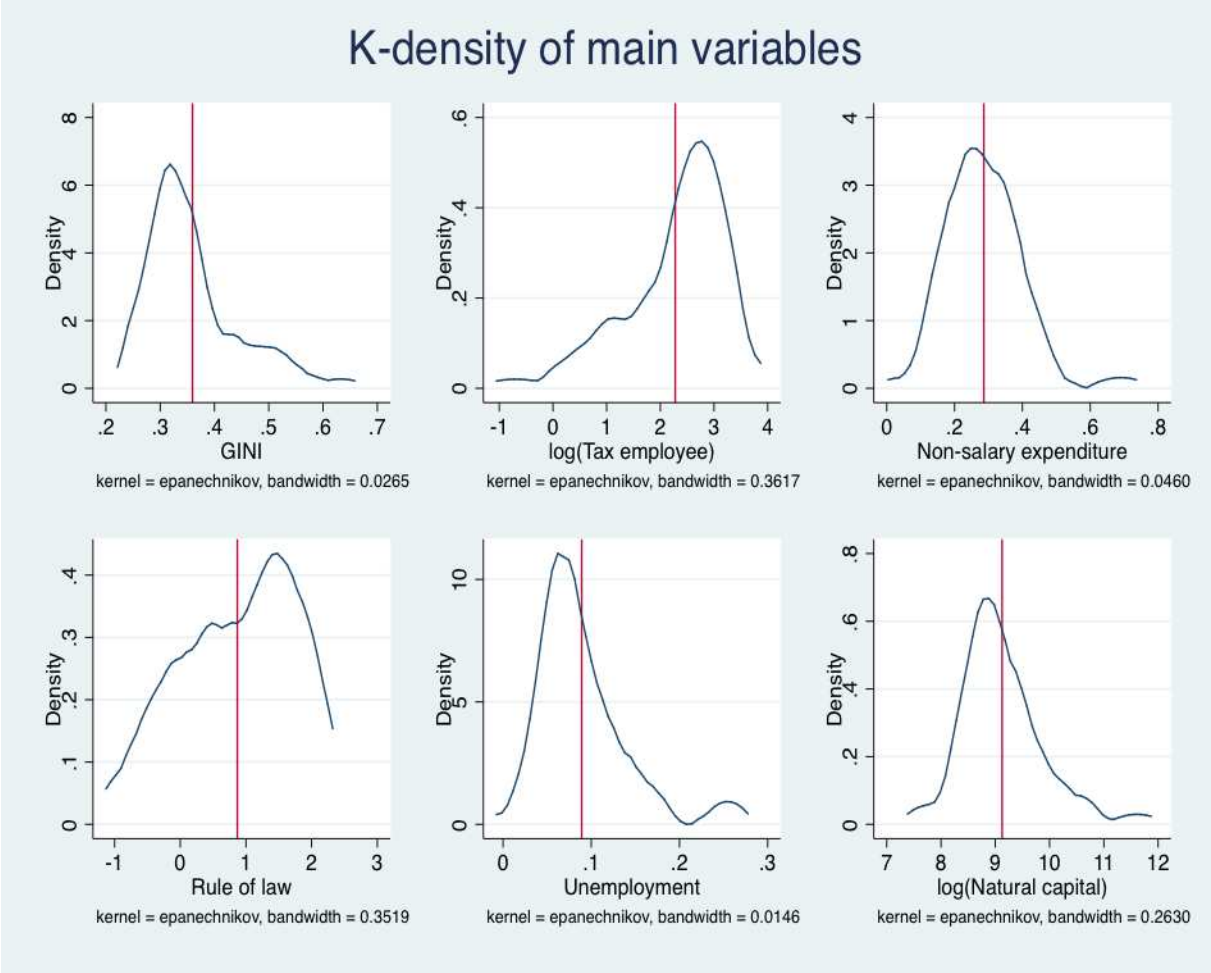

Figure 1: The distribution of main variables (vertical line represents the mean values).

3.3.2. Model

$$GINI = \beta_0 + \beta_1 \times [Tax\,employee] + \beta_2 \times [Non - salary\,expenditure] + \beta_3 \times [Rule\,of\,law] + \\ + \sum \delta_i \times [Control_i] + \epsilon$$

in which Control is a vector of control variables entailing unemployment rate (UE) and log natural capital per capita (NC), and ε is error term.

Table 1 shows correlations among major variables employed in the paper.

	Abbr.	GINI	log(TE)	NSE	UE	RL	log(NC)
Gini coefficients	GINI	1					
log(Tax employee)	TE	-0.5821	1				
Non-salary expenditure	NSE	-0.1036	-0.1305	1			
Unemployment	UE	0.1142	0.1917	0.0129	1		
Rule of law	RL	-0.5201	0.4099	-0.1619	-0.0997	1	
log(Natural capital)	NC	-0.0534	0.1494	-0.0347	-0.2552	0.3216	1

Table 1: Correlation matrix

Table 2 demonstrates the descriptive statistics for all main variables.

Statistics	GINI	log(TE)	NSE	UE	RL	log(NC)
Mean	0.3595	2.2781	0.2863	0.0892	0.8687	9.1258
Median	0.3379	2.6221	0.285	0.0745	0.9850	8.9632
Min	0.2479	-0.6931	0.0500	0.0070	-0.7800	7.6468
Max	0.632	3.5068	0.6900	0.2630	1.9700	11.6097
Standard Deviation	0.0849	0.9152	0.1100	0.0508	0.8408	0.7330
Variance	0.0072	0.8377	0.0121	0.0026	0.7070	0.5373
Standard Error (mean)	0.0125	0.1349	0.0162	0.0075	0.1240	0.1118
Observations	46	46	46	46	46	43

Table 2: Summary statistics

4. Empirical Results

4.1. Main results

Dependent variable: GINI				
	(1)	**(2)**	**(3)**	**(4)**
log(Tax employee)	-0.0562***	-0.0427**	-0.0475**	-0.0490***
	(0.0182)	(0.0176)	(0.0182)	(0.0164)
Non-salary expenditure	-0.141*	-0.172**	-0.174**	-0.185**
	(0.0777)	(0.082)	(0.081)	(0.0803)
Rule of law		-0.0371***	-0.0332***	-0.0376***
		(0.0104)	(0.0102)	(0.0106)
Unemployment			0.305	0.445
			(0.304)	(0.331)
log(Natural capital)				0.0245*
				(0.0136)
Constant	0.528***	0.538***	0.519***	0.298**
	(0.0532)	(0.0463)	(0.0472)	(0.135)
Observations	46	46	46	43
R-squared	0.372	0.482	0.513	0.565
P-value	0.0037	0.000	0.000	0.000

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Main estimation results

Table 3 reports the main results of the paper from the cross-sectional multiple regression with robust option to control the heteroscedasticity, modeling the relationship between GINI and explanatory variables as well as how these independent variables help predict GINI. Particularly, it is shown with the value of the estimated coefficients, the robust standard errors and an indication of the significance level for each and every model. Observing the R-squared accompanied with p-values, all of the models are recorded to reach significance levels below 0.01 (1%) and the equivalent variation in statistics GINI might be explained by independent variables. This may confirm the good fit to the data. Also, the null hypothesis that all estimated coefficients are equal to zero is thus rejected in all these models.

As can be seen from Table 3, starting by model 1 for only two main explanatory variables concerning Tax administration resources (TE & NSE), it is found that this model is fit and both variables are statistically significant in explaining GINI. Then, model 2 is added with the third independent variable - RL. Again, the model is fit and all these three variables are found to be significant.

We then add the two control variables – UE and NC to the models 3 and 4 respectively. It can be observed that, the two models are with increasing R-squared, especially the model 4 – our baseline regression - that could explain 56.5% of the GINI's variation. And all explanatory variables are found statistically significant. It seems remarkable that, signs of all coefficients appear to be suitable as expected. That is to say, the coefficients on Tax employee, non-salary expenditure and rule of law are witnessed to be negative and highly significant at level of 1% and 5%. And these results therefore generally support all hypotheses. In an average country, Tax administration with the focus on resources usage represented by Tax administration staffs and non-salary expenditure as well as the stronger law of public and order may contribute to reducing the inequality of income distribution. More specific, according to our first hypothesis (H1a), TE is negative and statistically significant; therefore, the increasing number of Tax administration staffs helps decrease the rate of income inequality (GINI coefficient). Similarly, NSE is negatively and statistically significant, revealing the negative relationship between TE and GINI. And, the higher proportion of non-salary expenditure in Tax administration reduces the degree of income distribution inequality as proposed in hypothesis H1b.

Regarding the hypothesis H2, the quality of institution factors – RL – tends to be capable of controlling the GINI. With the negative and statistically significant RL coefficient, it is suggested that, high GINI countries will be generally witnessed with low RL index, as proposed in hypothesis 3.

In addition, it is worthwhile to pay attention to the robustness of all three independent variables namely TE, NSE, and RL when they appear and remain strongly significant in all models with considerable coefficients indicating their robust effect on predicting GINI's variability, no matter what the existence of control variables i.e. UE and NC.

Figure 2 shows the partial effects of these explanatory variables on GINI.

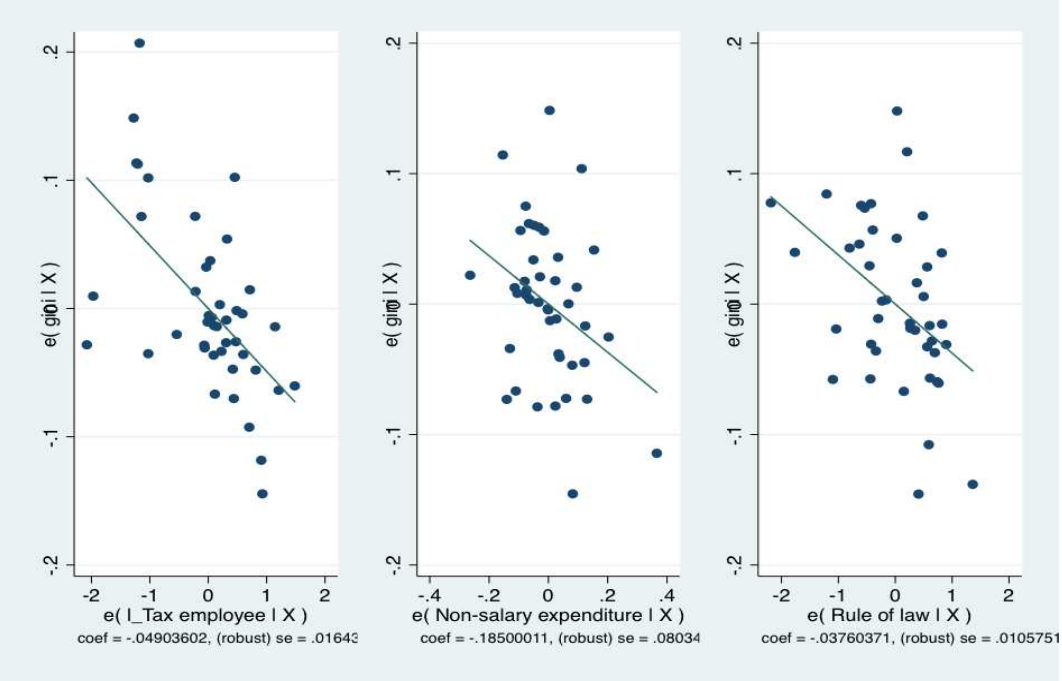

Figure 2: The partial effects of explanatory variables on GINI

4.2. Additional control variables

Then, for the purpose of omitted variables, we add more control variables to the specification. As shown in table 4, the major results appear to be unchanged when controlling for these aforementioned potential covariates of GINI. The two main hypotheses concerning Tax administration and institutional quality remain the significance in explaining GINI's variability in all three models, say, adding economic production and growth (GDP), total schooling, and other joint variables of educational attainment respectively. And the noticed thing is that there is not any change in explanatory power of these three models i.e. (5) (6) and (7) demonstrated by observed R-squared (around 56%) compared to baseline regression in model (4). This helps confirm the suitability and validity of the baseline model.

4.3. Further robustness checks

In this sub-section we present some robustness checks as an exercise for the core regression coefficient estimation. Consequently, it is said that one of the most essential issues is to interpret the impact of explanatory variables on the variation of GINI at different points of time. We hence generate the model with the utilization of data 2011. Table 5 presents the outcomes. And then we also make use of other estimators for checking the findings (see table 6 for details).

Dependent variable: GINI			
	(5)	**(6)**	**(7)**
log(Tax employee)	-0.0512***	-0.0488***	-0.0482***
	(0.0158)	(0.0167)	(0.0174)
Non-salary expenditure	-0.176*	-0.185**	-0.184**
	(0.0887)	(0.0815)	(0.0842)
Rule of law	-0.0455*	-0.0373***	-0.0380***
	(0.0231)	(0.0111)	(0.0108)
Unemployment	0.451	0.445	0.458
	(0.341)	(0.336)	(0.348)
log(Natural capital)	0.0223	0.0246*	0.0239
	(0.0147)	(0.0142)	(0.0156)
log(GDPpc)	0.0104		
	(0.0297)		
Total schooling		-0.0408	
		(0.00717)	
Primary schooling			-0.00259
			(0.0123)
Secondary schooling			-0.00124
			(0.0105)
Tertiary schooling			0.0089
			(0.0347)
Constant	0.222	0.300**	0.314*
	(0.268)	(0.139)	(0.168)
Observations	43	43	43
R-squared	0.567	0.565	0.566
P-value	0.000	0.000	0.000
Robust standard errors in parentheses			
*** p<0.01, ** p<0.05, * p<0.1			

Table 4: Additional control variable

4.3.1. The baseline regression with data of 2011[4]

Dependent variable: GINI		
	2013 (4)	**2011 (8)**
log(Tax employee)	-0.0490***	-0.0330*
	(0.0164)	(0.0192)
Non-salary expenditure	-0.185**	-0.136*
	(0.0803)	(0.0793)
Rule of law	-0.0376***	-0.0416***
	(0.0106)	(0.0095)
Unemployment	0.445	0.131
	(0.331)	(0.170)
log(Natural capital)	0.0245*	0.0210
	(0.0136)	(0.0143)
Constant	0.298**	0.305**
	(0.135)	(0.136)
Observations	43	35
R-squared	0.565	0.538
Robust standard errors in parentheses	0.000	
*** p<0.01, ** p<0.05, * p<0.1		

Table 5: Baseline regression with data of 2011

[4]See Appendix 2 for description of data used in this model

As seen in table 5, it is likely that the model is still fit with data of 2011. And all of the coefficients of independent variables i.e. TE, NSE and RL are recorded significantly negative as expected. The hypotheses remain supported although the expanatory powers of the whole model as well as two main independent variables i.e. TE and NSE are found rather lowered compared to the baseline regression i.e. model (4). This means, the relationship between Tax administration resources and rule of law and income inequality remains.

4.3.2. Application of other techniques for baseline regression

Bearing the nature of GINI i.e. belongs to 0-1 interval, we make the application of truncated regression for estimation. Moreover, we utilize the weighted least squared estimator for the heteroscedasticity matter. And, it is then acknowledged that the models are fit, and especially the main results survive these tests as expected, even the significance of certain variable (NSE) are improved (Table 6). Particularly, the model (10) with WLS is obtained with the highest explanatory power (R-squared = 57.48%), and the coefficients of variables NSE and NC are acknowledged to be more strongly significant at levels of 1% and 5% respectively instead of 5% and 10% in the baseline model.

Dependent variable: GINI		
	Truncated Reg. (9)	**WLS (10)**
log(Tax employee)	-0.0490***	-0.0444***
	(0.0154)	(0.0162)
Non-salary expenditure	-0.185**	-0.238***
	(0.0754)	(0.0726)
Rule of law	-0.0376***	-0.0473***
	(0.00993)	(0.0130)
Unemployment	0.445	0.612
	(0.311)	(0.390)
log(Natural capital)	0.0245*	0.0392**
	(0.0128)	(0.0174)
Constant	0.298**	0.162
	(0.127)	(0.169)
Sigma	0.0559***	
	(0.0067)	
R-squared		0.5748
Observations	43	43

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Baseline regression with truncated regression and weighted least squared (WLS) estimator

4.4. Alternative measure of GINI

In this sub-section, we make a replacement of GINI with GINI_WIID for observing any potential difference. It is our hope the main results may survive the test.

Table 7 shows the empirical results with the usage of GINI_WIID[5] instead of GINI. And it is the fact that, all explanatory variables are estimated with significantly negative coefficients, demonstrating the support for all three above-mentioned hypotheses.

Dependent variable: GINI / GINI_WIID		
	GINI (4)	**GINI_WIID (11)**
log(Tax employee)	-0.0490***	-0.0463**
	(0.0164)	(0.0174)
Non-salary expenditure	-0.185**	-0.176**
	(0.0803)	(0.0723)
Rule of law	-0.0376***	-0.0259**
	(0.0106)	(0.0111)
Unemployment	0.445	0.456
	(0.331)	(0.358)
log(Natural capital)	0.0245*	0.0217*
	(0.0136)	(0.0129)
Constant	0.298**	0.319**
	(0.135)	(0.132)
Observations	43	43
R-squared	0.565	0.496
P-value	0.000	0.0002

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Baseline regression with GINI_WIID

Clearly, the choice of GINI might satisfy the research criteria of reliability and validity for "reducing the possibility of getting the answer wrong means" (Saunders et al., 2009, p.156). In other words, it is our effort of well selecting and collecting data derived from reliable sources that are relevant to the research objectives. In this sense, the findings seem really about what they appear to be about as mentioned in Saunders et al. (2009, p.157).

To conclude this section, taken together these results confirms the robustness of the models, results, and the significance and explanatory power of independent variables in explaining GINI's variation as well. It is also noticeable with the reliability and validity of the data applied in the paper, particularly the GINI from WDI.

5. Discussion and Conclusions

As referred earlier, the target of this paper is to advance our empirical understanding about the determinants of income inequality. Specifically, through the proposed hypotheses, we would stress the

[5]GINI coefficients extracted from World Income Inequality Database - See 3.2. Data description and Appendix 2 for details

role of Tax administration resources on disciplining the income inequality besides the institutional quality as GINI's determinant. The remarkable thing is that, this paper tends to contribute to the research field by the first time providing empirical evidence with regards to the negative effect of Tax administration resources on income distribution inequality. The study places the emphasis on Tax administration practices and performance that, the devotion of resources to Tax system administration is an important and topical issue for governments (OECD, 2015). Good tax administration always matters (Carter & Matthews, 2012). In details, in view of amelioration in tax governance, compliance and enforcement, the suitable number of Tax administration-related function staffs and the considerable proportion of non-salary expenditure i.e. infrastructure, information technology (IT), human resource management (HRM)... might dedicate to fostering the equality of income distribution (or "social welfare") by partly eliminating the deadweight loss on the income transfer process.

The paper shows the empirical results strongly supporting the hypotheses, especially the first two hypotheses on Tax administration resources' explanatory power in GINI's variation. The findings, via various models, are robust to data 2013, 2011 and even with different techniques (OLS regression, truncated regression, weighted least squared), to controlling for unemployment rate and natural capital as well as educational attainment and economic production and growth (GDPpc). The alternative measure of GINI is also utilized for validation and robustness test, and it is demonstrated with result survival.

Regarding the first hypothesis on the relationship between Tax administration resources and income distribution inequality, it appears to be noticeable with this robustly negative association and that effectiveness of Tax administration to the extent of resources allocation is crucial for moderating the likelihood of income inequality across countries. More concretely, the results indicate the significantly negative relationship between the number of Tax administration staffs and the GINI index. That is to say, on conditional of other relevant factors, in countries possessing higher number of Tax administration-related functions corresponding to the size of population and economy, it would be found with lower rate of income inequality. This can be seen with Latin American countries such as Colombia, Brazil, Costa Rica, Chile, Mexico, Argentina which have been acknowledged among the most unequal distribution of income (GINI: 0.44 – 0.55) with the comparably low number of employees.[6]

Also, the share of non-salary expenditure i.e. expenditure of Tax administration-related functions besides total salary expenditure is recorded with negative link to GINI index. This suggests that, countries with higher proportion of afore-mentioned investment might predict a lower inequality of income distribution. It seems undeniable that, on the base of supporting tax collection as well as inspecting and preventing tax fraud and evasion potentials, IT operations giving support for

[6]See Appendix 6. Data file for details

administrative operations together with HRM functions providing administrative operations in terms of personnel, payroll, recruitment, learning and developing... have been believed to increasingly greater, for efficiency and effectiveness, as Governments' decisions on reducing the public wage costs for other kind of Tax administration – related investment (OECD, 2015).

As institutional framework is linked with how a society's gains are distributed and redistributed among its members (Chong & Calderon, 2000), the institutional quality plays an important role in income distribution. As far as the second hypothesis on institutional quality represented by rule of law in this research paper is concerned, our findings support the hypothesis that the strong law and order diminish the income inequality. This paper, therefore, helps consolidate the previous study by Chong and Calderon (2000), Chong and Gradstein (2007), Perera and Lee (2013), and Fuentes et al. (2014) to the extent of causality between institutional strength and a more equal distribution of income. The Scandinavian countries Denmark, Finland, Sweden, and Norway whose RL index (among the highest) are 1.8733, 1.9257, 1.9541 and 1.9681 respectively accompanied with 0.2907, 0.2757, 0.2702 and 0.2617 for GINI (among the lowest) are considered clear examples.[7] This may help confirm results from other studies.

It is likely the truth that, organizational and institutional reforms play essential parts in Tax administration's effectiveness and efficiency, resulting in the worthy decrease of income distribution inequality. Taken together, Tax administration with specific respect to resources allocation accompanied with institutional quality are certified effective and helpful for income equality, to some certain extent. This also helps commit to the fact that whether the boom of industrialization and modernization, human always take a certain central position in all fields. Therefore, besides the institutional quality and non-salary expenditure (investment in technology, automation...), it appears to be undeniable with the sustainably key role of number of Tax administration-related officials.

In a nutshell, Tax administration as a part of "service paradigm" relates to the likelihood of social welfare demonstrated with income distribution measured by GINI index. The findings are novel in the literature and the paper contributes to the research field as being the first to reveal the effect of Tax administration in terms of resources deployment on income inequality. And it can be undeniably recognized with the need of effective Tax administration to gain better efficiency in income redistribution (or "tax transfer") for the purpose of higher equality by avoiding the "leaky bucket" (or minimizing the deadweight loss) on the process of income redistribution as referred in Browning (2008).

[7] See Appendix 6. Data file for details

6. Limitations and future research

It is often witnessed with limitations of any research, and this paper is thus not without its limitations, opening the fruitful spaces and directions for future research.

In relation to the data set, this paper is limited to 46 countries (fewer than 25% of total Globe) that mainly belong to high and upper-middle income classification. It might fail to get the generality of the findings. A further research conducting with a broader and more various collection of data set would be better in understanding the relationship between Tax administration resources and income distribution inequality.

Methodologically speaking, the paper uses cross-sectional method, in spite of robustness checks, this paper fails to address time dimension in which the variables are considered to be associated with a sequence of points in time to witness any change trend during time or causal relationship then. This study is hence likely to miss the opportunity to catch the causal effect, suggesting the improvement with panel approach.

In other respects, owing to the objective limitation caused by Tax administration database that has been consolidated with only three available versions saying Tax Administration 2006, 2013 and 2015 for 2004, 2011 and 2013 data respectively, resulting in the limited choice of corresponding data i.e. 2013 and 2011 in this paper. As a consequence, further research overcoming this drawback and applying bigger and more availably sufficient Tax administration databases in the future would definitely facilitate the relevant knowledge in terms of other potential aspects of Tax administration linking to income inequality.

Also, it might be remarkable that more study on this new relationship and effect of Tax administration on social welfare should be conducted as a hint of this paper. In the meantime, there arises study with different data to survive the findings over times and examine if the relationship between Tax administration resources and income inequality remains over the time. This could definitely advance the better acknowledgement of GINI's underlying and implicit determinants, navigating the better policy with reference to income equality and equity.

Furthermore, it is noted with slight limit when this paper basically deals with the effect of Tax administration to the extent of resources usage and exploitation. In this sense, it would be of interest to conduct further research discovering this relationship in respect of other aspects of Tax administration i.e. institutional arrangement of Tax administration, organization of revenue bodies, strategic management, human resource management, operational performance management..., or even more resource-related variables namely cost of collection ratios, Tax administration expenditure / GDP... By doing so, that could consolidate the impact as well as explanatory power of Tax administration on income inequality.

In this paper, tax rates are implicitly assumed equal throughout the countries. That would be hence

valuable for future research, in view of tax reform, to make effort on trying to correct for tax rates i.e. tax pressure as stated in the law. Conjointly, interacting rule of law with Tax administration resources would show light in the respect of policy issue in using current resources more efficiently.

Finally, the paper, with the first time, discovers the mutual link and effect of Tax administration resources on income inequality when other factors held as equal. It will then be interesting if future research may further explore and extend this finding and evidence in a deeper way i.e. examine whether and how this effect will be differently moderated and mediated under the differed environmental factors such as organizational, institutional, historical, economic, political and social elements that can dedicate to intervening, advancing or hindering this effect. Additionally, the research that focuses and analyses this finding under the lens of Inequality and Public economics instead of Institutional economics would be an another direction.

7. Implications for policy-makers

As mentioned earlier, income inequality has been considered among important issues facing all nations throughout the World. This could be thanks to the fact that income inequality leads to increased mortality via disinvestment in social capital (Kawachi et al., 1997) and this phenomenon is thought to increase the level of corruption through material and normative mechanisms (Jong-Sung & Khagram, 2005).

From the findings of this paper, it is undeniable that, income inequality may be affected by a variety of elements besides existing GINI's determinants i.e. population, economic growth, trade openness, natural resources abundance, educational attainment, institutional quality... Consequently, it should be paid more attention to legislation of the suitable Tax administration policies. Obviously, the optimal employment of resources namely considerable number of relevant staffs as well as the notable Tax administration's non-salary expenditure could absolutely improve and facilitate more efficiency in tax collection, leading the better distribution of income through avoidance of redistributive pressures (Jong-Sung & Khagram, 2005).

In addition, the governments should, to the extent of policy execution, be aware of the combination of alternative strategies i.e. traditional enforcement paradigm, service paradigm, and trust paradigm (Alm, 2012) to reduce tax evasion and thus increase the tax transfer efficiency, resulting in better social welfare with lower income inequality.

And lastly, it finds likely to suggest that increase of Tax administration resources might increase social welfare. Nevertheless, there still exists a policy issue in using current resources more efficiently. That hence results in the crucial policy that notices and moderates the interaction of Tax administration resources and institutional quality indicators as referred previously for future research direction.

[1] Aaberge, R., Wennemo, T., Bjorklund, A., Jantti, M., Pedersen, P. J., & Smith, N. (2000). Unemployment shocks and income distribution: how did the Nordic countries fare during their crises?. The Scandinavian Journal of Economics, 102(1), 77-99.

[2] Aghion, P., Caroli, E., & Garcia-Penalosa, C. (1999). Inequality and economic growth: The perspective of the new growth theories. Journal of Economic literature, 37(4), 1615-1660.

[3] Alm, J. (2011). Designing alternative strategies to reduce tax evasion. In Tax evasion and the shadow economy. p.13-32. Edward Elgar Pub

[4] Auty, R. M. (1997). Natural resource endowment, the state and development strategy. Journal of International Development, 9(4), 651-663.

[5] Bakija, J. (2014). Social Welfare, Redistribution, and the Tradeoff between Efficiency and Equity, with Developing Country Applications.

[6] Barro, R. J. (2000). Inequality and Growth in a Panel of Countries. Journal of economic growth, 5(1), 5-32.

[7] Barro, R. J., & Lee, J. W. (2000). International Data on Educational Attainment: Updates and Implications," CID Working Paper No. 042, April.

[8] Bellù, L. G., & Liberati, P. (2006). Inequality analysis: The gini index. FAO, EASYPol Module, 40.

[9] Breen, R., & Chung, I. (2015). Income inequality and education. Sociological Science, 2, 454-477.

[10] Browning, E. K. (September 5, 2008). The Leaky Bucket. Retrieved from http://www.independent.org/newsroom/article.asp?id=2309

[11] Browning, E. K., & Johnson, W. R. (1984). The trade-off between equality and efficiency. The Journal of Political Economy, 175-203.

[12] Carmignani, F. (2013). Development outcomes, resource abundance, and the transmission through inequality. Resource and Energy Economics, 35(3), 412-428.

[13] Carter, A. & Matthews, S. (2012). How tax can reduce inequality. Retrieved August 26, 2016, from OECD Observer: http://www.oecdobserver.org/news/fullstory.php/aid/3782/How_tax_can_reduce_inequality.html

[14] Chong, A., & Calderon, C. (2000). Institutional quality and income distribution. Economic Development and Cultural Change, 48(4), 761-786.

[15] Chong, A., & Gradstein, M. (2007). Inequality and institutions. The Review of Economics and Statistics, 89(3), 454-465.

[16] Creedy, J. (1994). Taxation, Poverty and Income Distribution. Edward Elgar Publishing.

[17] Dorfman, R. (1979). A formula for the Gini coefficient. The Review of Economics and Statistics, 146-149.

[18] Fuentes, R., Mishra, T., Scavia, J., & Parhi, M. (2014). On optimal long-term relationship between TFP, institutions, and income inequality under embodied technical progress. Structural Change and Economic Dynamics, 31, 89-100.

[19] Fum, R. M., & Hodler, R. (2010). Natural resources and income inequality: The role of ethnic divisions. Economics Letters, 107(3), 360-363.

[20] Galbraith, J. K., Conceicao, P., & Ferreira, P. (1999). Inequality and unemployment in Europe: the American cure.

[21] Goderis, B., & Malone, S. W. (2011). Natural resource booms and inequality: theory and evidence. The Scandinavian Journal of Economics, 113(2), 388-417.

[22] Gregorio, J. D., & Lee, J. W. (2002). Education and income inequality: new evidence from cross-country data. Review of income and wealth, 48(3), 395-416.

[23] Gylfason, T., & Zoega, G. (2002). Inequality and economic growth: Do natural resources matter?.

[24] Helpman, E., Itskhoki, O., & Redding, S. (2010). Inequality and unemployment in a global economy. Econometrica, 78(4), 1239-1283.

[25] Jong-Sung, Y., & Khagram, S. (2005). A comparative study of inequality and corruption. American Sociological Review, 70(1), 136-157.

[26] Kaplow, L., & Shavell, S. (1994). Why the legal system is less efficient than the income tax in redistributing income. The Journal of Legal Studies, 23(2), 667-681.

[27] Kawachi, I., Kennedy, B. P., Lochner, K., & Prothrow-Stith, D. (1997). Social capital, income inequality, and mortality. American journal of public health, 87(9), 1491-1498.

[28] Kuznets, S. (1955). Economic growth and income inequality. The American economic review, 45(1), 1-28.

[29] Leach, J. (2004). A course in public economics. Cambridge University Press.

[30] Leamer, E. E., Maul, H., Rodriguez, S., & Schott, P. K. (1999). Does natural resource abundance increase Latin American income inequality?. Journal of development Economics, 59(1), 3-42.

[31] Lee, H. Y., Kim, J., & Cin, B. C. (2013). Empirical analysis on the determinants of income inequality in Korea. International Journal of Advanced Science and Technology, 53, 95-110.

[32] Nathans, L. L., Oswald, F. L., & Nimon, K. (2012). Interpreting multiple linear regression: A guidebook of variable importance. Practical Assessment, Research & Evaluation, 17(9).

[33] OECD 2012, "Income inequality and growth: The role of taxes and transfers", OECD Economics Department Policy Notes, No. 9. January 2012.

[34] OECD (2015), Tax Administration 2015: Comparative Information on OECD and Other Advanced and Emerging Economies, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/tax_admin-2015-en

[35] Perera, L. D. H., & Lee, G. H. (2013). Have economic growth and institutional quality contributed to poverty and inequality reduction in Asia?. Journal of Asian Economics, 27, 71-86.

[36] Rosen, H.S. & Gayer, T. (2009). Public Finance, 9th Edition. McGraw-Hill/Irwin

[37] Ross, M. L. (2007). How mineral-rich states can reduce inequality. In: Sachs, J.D., Stiglitz, J.E., Humphreys, M. (Eds), Escaping the resource curse. Columbia University Press, New York, p.237-255.

[38] Salanie, B. (2003). The economics of taxation. MIT press.

[39] Sarel, M. M. (1997). How macroeconomic factors affect income distribution: the cross-country evidence (No. 97-152). International Monetary Fund.

[40] Saunders, M., Lewis, P. & Thornhill, A. (2009). Research Methods for Business Students (5th edition). Essex, England: Pearson.

[41] Solt, F. 2014. "The Standardized World Income Inequality Database." Working paper. SWIID Version 5.0, October 2014

[42] The Barro & Lee Educational attainment dataset version 2.1 (02/2016). Retrieved from http://www.barrolee.com/

[43] Wilkinson, M. (1992). What are Taxes For? In Taxation (pp. 1–14). London: Macmillan Education UK. http://doi.org/10.1007/978-1-349-21989-6_1

[44] Wooldridge, J. M. (2009). "Part 1: Regression Analysis with Cross Sectional Data". Introductory econometrics: a modern approach (4th ed.). Cengage Learning. ISBN 978-0-324-66054-8.

[45] World Bank, 2011. The changing wealth of nations: Measuring sustainable development in the new millennium. Retrieved from https://siteresources.worldbank.org/ENVIRONMENT/Resources/ChangingWealthNations.pdf

[46] World Development Indicators. World Bank. Retrieved from http://data.worldbanl.org/wdi

[47] World Wide Governance Indicators. World Bank. Retrieved from http://info.worldbank.org/governance/wgi/index.aspx#home

[48] Yang, J., & Qiu, M. (2016). The impact of education on income inequality and intergenerational mobility. China Economic Review, 37, 110-125.

[49] Yunker, J. A. (2016). Economic inequality and optimal redistribution: A theoretical and empirical analysis. Journal of Policy Modeling.

Appendices

Appendix 1. *List of 46 countries in the sample*

No.	Country	OECD (Yes / No)	Income Group Classification[1]	GINI – Average 5 years (2008-2012)
1	Argentina	No	Not classified	0.440
2	Australia	Yes	High income	0.350
3	Austria	Yes	High income	0.310
4	Belgium	Yes	High income	0.280
5	Brazil	No	Upper middle income	0.540
6	Bulgaria	No	Upper middle income	0.350
7	Canada	Yes	High income	0.340
8	Chile	Yes	High income	0.510
9	Colombia	No	Upper middle income	0.550
10	Costa Rica	No	Upper middle income	0.490
11	Croatia	No	High income	0.320
12	Cyprus	No	High income	0.320
13	Denmark	Yes	High income	0.290
14	Estonia	Yes	High income	0.320
15	Finland	Yes	High income	0.280
16	France	Yes	High income	0.330
17	Germany	Yes	High income	0.310
18	Hungary	Yes	High income	0.290
19	Iceland	Yes	High income	0.280
20	India	No	Lower middle income	0.340
21	Indonesia	No	Lower middle income	0.350
22	Ireland	Yes	High income	0.320
23	Israel	Yes	High income	0.430
24	Italy	Yes	High income	0.340
25	Japan	Yes	High income	0.320
26	Latvia	No	High income	0.350
27	Lithuania	No	High income	0.350
28	Luxembourg	Yes	High income	0.320
29	Malaysia	No	Upper middle income	0.460
30	Mexico	Yes	Upper middle income	0.480
31	Netherlands	Yes	High income	0.290
32	Norway	Yes	High income	0.260
33	Poland	Yes	High income	0.330
34	Portugal	Yes	High income	0.360
35	Romania	No	Upper middle income	0.280
36	Russia	No	Upper middle income	0.410
37	Slovak Republic	Yes	High income	0.270
38	Slovenia	Yes	High income	0.250
39	South Africa	No	Upper middle income	0.630

[1] calculated using the World Bank Atlas method: low-income economies are defined as those with a GNI per capita of $1,045 or less in 2014; middle-income economies are those with a GNI per capita of more than $1,045 but less than $12,736; high-income economies are those with a GNI per capita of $12,736 or more. Lower-middle-income and upper-middle-income economies are separated at a GNI per capita of $4,125

No.	Country	OECD (Yes / No)	Income Group Classification[1]	GINI – Average 5 years (2008-2012)
40	Spain	Yes	High income	0.360
41	Sweden	Yes	High income	0.270
42	Switzerland	Yes	High income	0.330
43	Thailand	No	Upper middle income	0.390
44	Turkey	Yes	Upper middle income	0.390
45	United Kingdom	Yes	High income	0.340
46	United States	Yes	High income	0.400

Appendix 2. *Full list of data description, measure and sources*

No.	Variable	Description	Measure & Source
1.	Gini coefficients (GINI)	The proxy for social welfare. The uni-dimensional index measuring Inequality for socio-economic impact of policies. Gini index measures the extent to which the distribution of income (or, in some cases, consumption expenditure) among individuals or households within an economy deviates from a perfectly equal distribution. A Lorenz curve plots the cumulative percentages of total income received against the cumulative number of recipients, starting with the poorest individual or household. The Gini index measures the area between the Lorenz curve and a hypothetical line of absolute equality, expressed as a percentage of the maximum area under the line. Thus a Gini index of 0 represents perfect equality, while an index of 1 implies perfect inequality.	Average GINI over 5 years (2008-2012), source: WDI - World Development Indicators (World bank estimate). GINI applied for both 2013 and 2011. GINI_WIID: extract the data gaps and then, average available comparative Gini coefficients over the period from 2008 to 2012. Source: World Income Inequality database (WIID). GINI_WIID applied for 2013.
2.	Tax employee (TE)	Number of employees of tax administration per 10,000 active citizens (total citizens aged 15-64)	Calculated by number of 10,000 active citizens (adopted from World Development Indicators – World Bank) divided by number of tax administration staffs (OECD Tax administration 2015 – Table 2.4 for 2013 & OECD Tax administration 2013 – Table 2.4 for 2011) Usage of log(TE)
3.	Non-salary expenditure (NSE)	This represents the share of non-salary expenditure in aggregate expenditure	Calculated by 1 – (minus) proportion of salary expenditure (percentage of salary costs for tax functions and related overheads / aggregate expenditure for tax functions and related overheads), adopted from OECD Tax administration 2015 - Table A.6 & A.5 for 2013 & OECD Tax administration 2013 – Table 5.1 for 2011.
4.	Unemployment (UE)	Total unemployment (% of total labor force) (modeled ILO estimate) Unemployment refers to the share of the labor force that is without work but available for and seeking employment.	WDI – Adopted from International Labor Organization, Key Indicators of the Labor Market database. Data for year 2013 & 2011

No.	Variable	Description	Measure & Source
5.	Rule of law (RL)	An indicator, used as proxy for institutional quality, measures the extent to which individuals / agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, the police, and the courts, as well as the likelihood of crime and violence. Ranges from -2.5 to +2.5 Higher values mean stronger law and order	Worldwide Governance Indicators (WGI) – World Bank Data of 2013 & of 2011
6.	Natural capital (NC)	A proxy for natural resources abundance An important share of total wealth, is the sum of nonrenewable resources (including oil, natural gas, coal, and mineral resources), cropland, pastureland, forested areas (including areas used for timber extraction and non-timber forest products), and protected areas. Most natural resources are valued by taking the present value of resource rents—the economic profit on exploitation—over an assumed lifetime	The paper uses log of NC per capita (US$) in 2005 – the most up-to-date statistics. Source: World Bank 2011 (Appendix C, Table C.1 – Wealth Estimates for 2005). Data used for both 2013 and 2011
7.	GDP per capita (GDP)	GDP per capita is gross domestic product divided by midyear population. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in current U.S. dollars.	Data of year 2013 & 2011, extracted from WDI.
8.	Total schooling (TS)	Average years of total schooling i.e. average years of education completed among people over age 15	Data of year 2010, used for both 2013 and 2011 Source: The Barro & Lee Educational attainment dataset version 2.1 (02/2016) - http://www.barrolee.com/
9.	Primary schooling (PR)	Average years of primary schooling among people over age 15	Data of year 2010, used for both 2013 and 2011 Source: The Barron & Lee Educational attainment dataset version 2.1 (02/2016) - http://www.barrolee.com/
10.	Secondary schooling (SE)	Average years of secondary schooling among people over age 15	Data of year 2010, used for both 2013 and 2011 Source: The Barron & Lee Educational attainment dataset version 2.1 (02/2016) - http://www.barrolee.com/
11.	Tertiary schooling (TR)	Average years of tertiary schooling among people over age 15	Data of year 2010, used for both 2013 and 2011 Source: The Barron & Lee Educational attainment dataset version 2.1 (02/2016) - http://www.barrolee.com/

Appendix 3. *Correlation matrix among institutional variables*

	RL	RQ	GE	CC
Rule of law (RL)	1			
Regulatory quality (RQ)	0.926	1		
Government effectiveness (GE)	0.9635	0.9117	1	
Corruption control (CC)	0.9716	0.8969	0.9603	1

Appendix 4. *Baseline regression model with different institutional indicators*

Dependent variable: GINI				
	(4)	(12)	(13)	(14)
log(Tax employee)	0.0490***	0.0516***	0.0508***	0.0514***
	(0.0164)	(0.0179)	(0.0170)	(0.0168)
Non-salary expenditure	-0.185**	-0.165**	-0.183**	-0.192**
	(0.0803)	(0.0782)	(0.0801)	(0.0772)
Unemployment	0.445	0.481	0.462	0.414
	(0.331)	(0.350)	(0.337)	(0.340)
log(Natural capital)	0.0245*	0.0210	0.0225	0.0253*
	(0.0136)	(0.0146)	(0.0138)	(0.0136)
Rule of law	0.0376***			
	(0.0106)			
Regulatory quality		-0.0312**		
		(0.0139)		
Government effectiveness			-0.0379**	
			(0.0144)	
Corruption control				0.0302***
				(0.00944)
Constant	0.298**	0.324**	0.322**	0.294**
	(0.135)	(0.150)	(0.140)	(0.138)
Observations	43	43	43	43
R-squared	0.565	0.502	0.534	0.549

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix 5. *Do-file (commands for Stata)*

gen l_taxemp = ln(taxemp)
gen l_nc = ln(natcap)
Figure 1. The K-density of main variables
kdensity gini

graph save den_gini

kdensity l_taxemp

graph save den_taxemp

kdensity non_sal_exp

graph save den_nse

kdensity ruleoflaw

graph save den_ruleoflaw

kdensity unemployment

graph save den_unemployment

kdensity l_natcap

graph save den_natcap

graph combine den_gini.gph den_taxemp.gph den_nse.gph den_ruleoflaw.gph den_unemployment.gph den_natcap.gph

Table 1. Correlation matrix

pwcorr gini l_taxemp non_sal_exp ruleoflaw unemployment l_natcap if year==2013

Table 2. Summary statistics

tabstat gini l_taxemp non_sal_exp ruleoflaw unemployment l_natcap if year==2013, statistics (mean median min max sd var semean count)

Table 3. Main estimation results

reg gini l_taxemp non_sal_exp if year==2013, robust

reg gini l_taxemp non_sal_exp ruleoflaw if year==2013, robust

reg gini l_taxemp non_sal_exp ruleoflaw unemployment if year==2013, robust

reg gini l_taxemp non_sal_exp ruleoflaw unemployment l_natcap if year==2013, robust

Figure 2. The partial effects of explanatory variables on GINI

avplot l_taxemp if year==2013

avplot non_sal_exp if year==2013

avplot ruleoflaw if year==2013

Table 4. Additional control variable

gen l_gdp = ln(gdp)

reg gini l_taxemp non_sal_exp ruleoflaw unemployment l_natcap l_gdp if year==2013, robust

reg gini l_taxemp non_sal_exp ruleoflaw unemployment l_natcap total_sch if year==2013, robust

reg gini l_taxemp non_sal_exp ruleoflaw unemployment l_natcap prim_sch sec_sch ter_sch if year==2013, robust

Table 5. Baseline regression with data of 2011

reg gini l_taxemp non_sal_exp ruleoflaw unemployment l_natcap if year==2011, robust

Table 6. Baseline regression with truncated regression and weighted least squared (WLS) estimator

truncreg gini l_taxemp non_sal_exp ruleoflaw unemployment l_natcap if year==2013, robust ll(0) ul(1)

regwls gini l_taxemp non_sal_exp ruleoflaw unemployment l_natcap if year==2013, wvars(l_taxemp non_sal_exp ruleoflaw unemployment l_natcap) type(loge2) robust

Table 7. Baseline regression with GINI_WIID

reg gini_wiid l_taxemp non_sal_exp ruleoflaw unemployment l_natcap if year==2013, robust

Appendix 3 - Correlation matrix among institutional variables

pwcorr ruleoflaw reg_quali gov_eff corr_con

Appendix 4 - Baseline regression model with different institutional indicators

reg gini l_taxemp non_sal_exp reg_quali unemployment l_natcap if year==2013, robust

reg gini l_taxemp non_sal_exp gov_eff unemployment l_natcap if year==2013, robust

reg gini l_taxemp non_sal_exp corr_con unemployment l_natcap if year==2013, robust