DRAFT

WIDER Development Conference

Public economics for development

5-6 July 2017 | Maputo, Mozambique

Income tax policy, non-compliance and its redistributive effects in Ghana

Edward Asiedu (University of Ghana), Dan Pavelesku (World Bank),
Ryoko Sato (World Bank), Tomomi Tanaka (World Bank)[1]

Abstract

Ghana's tax collection is very low compared with other lower middle-income countries. The revenue from income tax is particularly low, lower than the average of low-income countries. Non-compliance of tax payments is an urgent issue in Ghana, as the government has been suffering from a widening fiscal deficit and a rising debt burden. This paper combines data from household surveys, business census, and administrative income tax data, and examines the scale of non-compliance, and potential revenue gains from the enforcement of tax collection in Ghana.

Business census data suggests the actual number of formal sector workers is higher than the number of formal sector workers reported both in the household survey and administrative income tax data in Ghana. The income tax files do not include the individuals who work for formal sector firms which failed to file PAYE (Pay As You Earn), whereas the household data suffers from under-sampling and under-reporting of high wage income earners. By combining the household survey data and the administrative income tax data, and adjusting the number of formal sector workers using the business census, we reconstruct the distribution of wage earners in the formal sector who are subject to income tax, and estimate potential income tax revenue gain from the enforcement of tax collection. We find the income tax revenue could have been higher by 22 percent (equivalent to 0.5 percent of GDP) if everyone who files income tax in 2014 had paid full amounts of income tax due. The income tax revenue could have been higher by 53 percent (equivalent to 1.2 percent of GDP) if all formal sector firms and organizations, regardless of whether they actually filed income tax or not in 2014, had paid full amounts of PAYE for all their employees in 2014. In 2016, we observe a further reduction of income tax revenue. The total income tax revenue could have been higher by 86 percent (1.7 percent of GDP) if all formal sector firms and organizations, regardless of whether they actually filed income tax or not in 2016, paid the full amounts of PAYE for all their employees.

1.   Introduction

This paper combines data from household surveys, business census, and administrative income tax data, and examines the scale of income tax non-compliance in 2014 and 2016, and potential revenue gains from income tax compliance in Ghana.

Developing countries often collect too little tax. Ghana is not an exception. Table 1 shows that Ghana's tax collection, as a percentage of GDP, is very low compared with other lower middle-income countries. Income tax revenue is particularly low, lower than the average of low-income countries. Tax compliance is an urgent issue in Ghana, as the government has been suffering from a widening fiscal deficit and a rising debt burden.

Besley and Persson (2014) attribute low tax collection in developing countries to inadequate administrative capacity, presence of a large informal sector, weak checks and balances, and lack of social norms for tax compliance. They discuss how collecting income taxes is harder than collecting other taxes, such as trade taxes, because income tax collection requires a much more elaborate system of monitoring, enforcement, and compliance. Ahmed and Rider (2013) demonstrate that the evasion of income tax is estimated to be around 60 percent of potential revenues in Pakistan, which is much higher than the estimated evasion of indirect taxes.

Like many other countries, the vast majority of income taxes are paid as PAYE (Pay As You Earn) in Ghana. In order to encourage firms to file income taxes for their employees, the government of Ghana raised the minimum taxable income from 1,584 cedi to 2,592 cedi, and shifted the income tax bands upward, as shown in Table 2. Under the new tax act, wage earners who make below 2,592 cedi per year do not have to pay income tax. The maximum tax rate of 25 percent is only applied to people who earn more than 38,881 cedi a year, while the maximum tax rate of 25 percent was applied to people who earn at least 31,681 cedi until 2015. The tax rate change was applied from 2016. Experiences from other countries suggest the reduction in tax rates may reduce tax evasion. Fisman and Wei (2004) show that a one percentage point decrease in the tax rate is associated with a three percent reduction in corporate tax evasion in China.

Alvaredo and Gasparini (2013) demonstrate that only a few developing countries have made income tax data available, and incomes reported to the tax authorities are considerably higher than those reported in household surveys. Alverado and Piketty (2010, 2014) point out that top income individuals are usually not covered in household surveys, but are included in the administrative income tax data. Alvaredo and Londoño (2013) compare the Colombian household survey with administrative tax micro-data between 2007 and 2010 and find that household surveys poorly capture income at the top of the distribution.

For Ghana, we have access to both household data and administrative income tax data aggregated by income brackets in addition to business census. The business census (IBES) data suggests the actual number of formal sector workers is higher than the number of formal sector workers reported in both the household survey and income tax files. The

income tax files do not include individuals who work for formal sector firms which failed to file PAYE, whereas the household data suffers from under-sampling and under-reporting of high wage income earners. By using the algorithm developed by Fournier (2015) and Blanchet, Founier and Piketty (2017) and combining the household survey data and the income tax files, and adjusting the number of formal sector workers using the business census, we reconstruct a nationally representative distribution of wage earners in the formal sector who are subject to income tax. We then estimate potential income tax revenue and the scale of tax evasion in 2014 and 2016. Our results suggest the income tax revenue could have been higher by 22 percent (equivalent to 0.5 percent of GDP) if everyone who files income tax in 2014 had paid full amounts of income tax due. In addition, the total income tax revenue could increase by 53 percent if all formal sector firms and organizations, regardless of whether they actually filed income tax or not in 2014, pay full amounts of PAYE for all their employees in 2014. The tax evasion is equivalent to 1.2 percent of total GDP. In 2016, we observe a further reduction of income tax revenue. In 2016, the total income tax revenue could have been 86 percent higher if all formal sector firms and organizations, regardless of whether they actually filed income tax or not in 2016, paid the full amounts of PAYE for all their employees. The tax evasion is equivalent to 1.7 percent of total GDP.

The paper is organized as follows: Section 2 describes data and the empirical strategy. Section 3 shows results, and section 4 provides policy recommendations and conclusions.

2.   Data and methodology

2.1 Data

We use four types of data in this paper, namely, administrative income tax data, household survey data (2012/13 GLSS 6), business census data (2014 IBES), and the 2015 Labour Force Survey (LFS) data.

We have aggregated administrative income tax by the public and private formal sectors for 2011, 2012, 2014, and 2016. Figure 1 summarizes the income tax revenue by the public and private formal sectors for these years. Administrative income tax data contains aggregated Pay As You Earn (PAYE) and personal income tax for given income brackets by the public and private formal sector. The PAYE contributions are withholdings from salaries of employees. Self-employed persons are required to pay personal income tax at graduated rates in four equal installments. Figure 1 shows that income tax revenue increased dramatically between 2011 and 2014, but slightly declined in 2016. The increase in income tax collection from 2011 to 2014 was observed predominantly in the private sector. The income tax revenue rose by 267 percent in the private sector during that period.

The Ghana Living Standards Survey Round Six (GLSS 6) was conducted from 18th October 2012 to 17th October 2013. The data covers a nationally representative sample of 16,772 households. Detailed information was collected on demographic characteristics of households, education, health, employment, agricultural production, household

enterprises, household expenditure, income, and other factors. We combine GLSS 6 (2012/2013) and 2014 administrative income tax data to estimate the scale of tax evasion for 2014. Note the number of taxpayers was not updated by GRA between 2013 and 2014. Thus, the number of taxpayers we use for the 2014 income tax data was counted in 2013.

The 2015 Labour Force Survey (LFS) is a nationally representative sample of a household-based survey. Unlike GLSS 6, the 2015 LFS is focused on work-related questions. We combine wage data from the 2015 LFS and the administrative income tax data from 2016 to estimate the distribution of wage earners and income tax evasion in 2016.

In developed countries, administrative income tax data is suitable for studying distributions of income, as only a few individuals fail to file income tax. However, in developing countries, administrative income tax data suffers from a considerable number of missing wage workers and under-reporting of wages. Household survey data is also affected by under-sampling of wage earners, especially the top income individuals. This creates an extra challenge to capture the distributions of wage earners and potential taxpayers, and to estimate tax evasion and potential increases in government revenue from the enforcement of tax compliance.

We complement household survey data and administrative income tax information with business census, which covers all firms and organizations in Ghana. The 2014 Integrated Business Establishment Survey (IBES) is a business census that covers all establishments (firms and organization) undertaking economic activities in Ghana. IBES also covers firms and organizations without employees (i.e., self-employed individuals).

From IBES, we know the exact number of firms and the number of employees for each firm in each district and each industry. We also have information on whether each firm is registered with the Ghana Revenue Authority (GRA). Table A.1 in the Appendix summarizes the number of firms and organizations in each region, the percentage of firms and organizations registered with GRA, the total number of wage workers, and the number of wage workers who are working for the firms and organizations registered with GRA. There is a large difference in the percentage of firms and organizations registered with GRA across regions. Both the numbers of employers and firms are the largest in Greater Accra. Moreover, the percentage of firms which are registered with GRA is also highest in Greater Accra than other areas. Poor regions with smaller numbers of employers, namely Volta, Northern, Upper East and Upper West, have the lowest percentages of firms registered with GRA. Similarly, the percentage of workers who are working for the firms registered with GRA is highest in Greater Accra (76.5 percent). In Volta, Northern, Upper East and Upper West, the proportion of workers who are working for the firms registered with GRA are below 40 percent.

Table 3 compares the number of workers across datasets. The number of private sector workers who filed income tax was 1.145 million in 2014, while IBES data suggests there are 1.417 million people who were working for the firms registered with GRA in November 2013. It implies there are about 272 thousand formal private sector workers who are registered with GRA through their employers but their income tax was not filed.

For GLSS 6 and LFS, the number of workers was estimated using household weights. The estimated number of wage earners is 1.068 million in GLSS 6, which is smaller than both the number of formal private sector workers in the income tax files and IBES. There are two possible explanations. First, GLSS 6 under-sampled the workers in the formal private sector. It is a well-known fact that household surveys under-sample top income households, which are often wage earners. Second, there are 2.766 million wage earners who did not indicate whether their employers are formal or informal firms. These wage earners are categorized as "unknown" in Table 3.

It is likely that the public sector employees, especially top-income, are also under-sampled in GLSS 6. However, IBES did not conduct interviews with government offices. Therefore, we do not know the exact number of public sector employees from IBES.

IBES reports about 1 million private-sector employees work for firms which are not registered with GRA. It is smaller than the estimated number of informal private sector workers in GLSS 6 (1.247 million). This may be because IBES does not cover all the informal private sector firms, such as self-employed individuals.

The administrative income tax data contains only around 310 thousand private formal sector workers in 2016, which is much smaller than the number of private sector workers who filed income tax in 2014 (1.145 million). GRA staff explains that the number of firms that filed income tax declined significantly in 2016. When the headcounts of employees were conducted from October to November 2016, the government failed to supply sufficient electricity. It forced many firms to temporarily shut down their businesses. It is possible that the number of employees was higher in other months in 2016.

The number of public sector employees in the administrative income are updated only every few years. In the administrative income data set, the number of employees in the public sector were the same in 2013 and 2014. This suggests the headcounts of public sector employees were conducted sometime in 2013.

The estimated number of private formal sector workers is only around 398 thousand in the 2015 LFS data. This suggests LFS under-sampled private formal sector wage earners, or there are some problems with the weights.[2]

Figure 2 compares the number of income taxpayers reported in the administrative income tax data in 2014, and the estimated number of wage earners in the formal sector in GLSS 6 by income brackets. We use household weights and wage information for each individual worker to estimate the number of formal sector workers for each income bracket in GLSS 6. The estimated number of wage earners in GLSS 6 is much larger than the income tax payers in lower gross income brackets (annual gross income below 7,800 cedi) in both the public and private sectors. This is probably because people earning less than 7,800 cedi of

---

[2] LFS data is still not open to public. It is possible that the weights will be revised before GSS makes the data available to the public.

gross income are not paying income taxes, thus, GRA has no information about these individuals.

In contrast, the estimated number of wage earners in GLSS 6 is smaller than the income tax payers in higher gross income brackets (annual gross income of above 7,800 cedi) in both the public and private sectors. It implies GLSS 6 under-sampled wealthy individuals in the formal sector. This also indicates the Gini coefficient calculated using GLSS 6 may be underestimated, as GLSS 6 misses a large portion of the top income individuals.

Figure 3 compares the number of income tax payers in the administrative income tax data in 2016 and the estimated wage earners in the formal sector in the 2015 LFS data. We use household weights and wage information from each individual worker to estimate the number of formal sector workers for each income bracket for the LFS. In the public sector, the estimated number of wage earners in the LFS data is larger than the income tax payers in lower brackets of annual gross income below 12,000 cedi. In the private formal sector, the estimated number of wage earners in the LFS data is larger than the income tax payers for annual gross income below 24,000 cedi. This suggests that a considerable percentage of low income wage earners are not paying income tax, thus they are missing in the administrative income tax data. In contrast, the estimated number of wage earners in the LFS data is smaller than the actual income tax payers for higher gross income brackets in both the public and private sectors. It suggests LFS also under-sampled wealthy wage earners in the formal sector, however, the gap of the number of wage earners is much smaller for the LFS and 2016 administrative income tax data. It is partly due to a large portion of wage earners missing in the 2016 administrative income tax data.

2.2 Methodology

By combining the household survey data and the income tax files, and adjusting the number of formal sector workers using the business census, we reconstruct a nationally representative distribution of wage earners in the formal sector who are subject to income tax. Below, we explain step-by-step how we reconstruct the distribution. In the first step, we use administrative income tax data, and generate continuous generalized Pareto curves from aggregated income tax data, which is discrete. In the second step, we estimate the total number of potential income tax payers for each income bracket by applying the number of formal sector workers from business census to the distribution of formal sector employees, which we derive in the first step. The resulting distribution of formal sector workers include those who did not file income tax and are outside of the administrative income tax data.

First step

Using the algorithm developed by Fournier (2015) and Blanchet, Founier and Piketty (2017), we first attain continuous distributions of formal sector wage earners from the administrative income tax data by approximating generalized Pareto curves for the public and private formal sectors separately. We use the non-parametric estimation method

6

developed by Fournier (2015) and derive continuous generalized Pareto curves from aggregated income tax data which is discrete.

There are three critical factors in conducting this simulation, 1) mean income of each income bracket, 2) minimum and maximum income levels of each bracket, and 3) (inverted) Pareto coefficient. The continuous distributions of wage income are defined as the curve of inverted Pareto coefficients b(p), where p is the percentile rank and b(p) is the ratio between the average income above the percentile p and the income threshold at percentile p. We use Matlab for the simulation. Tables 4 and 5 summarize the mean income of each income bracket estimated from GLSS 6 and LFS, and the simulated b(p) for each income bracket for 2014 and 2016 administrative data, respectively. We estimate the mean income for each bracket using GLSS 6 and LFS since the mean income for each income bracket is not available from the administrative data. Substituting mean income from GLSS 6 and LFS is likely to underestimate the distribution of wage income.[3]

Figures 4 and 5 show the estimated continuous distributions of formal sector workers in the public and private sectors, as well as the distribution of total workers from the simulation. They are continuous versions of the distribution of taxpayers.

To compare the simulated distributions of workers from the administrative income tax data with GLSS 6 and LFS, we plotted them together in Figures 4 and 5. The blue bars represent the simulated distributions of workers from the administrative income tax data (generalized Pareto curves), and the red bards represent the distributions of formal sector workers in GLSS 6 and LFS, respectively. Figure 4 clearly shows that the simulated distribution of income tax payers in both the public and private formal sectors contains more high-income individuals, compare with the distribution of formal sector workers from GLSS 6. Figure 5 also shows that the simulated distribution of income tax payers contains more high-income individuals, especially in the private formal sector, compared with the distribution of formal sector workers from LFS. However, the difference is smaller between LFS and the administrative income tax data than GLSS 6 and the administrative income tax data. It is largely due to a significant number of missing taxpayers in the 2016 administrative income tax data.

Second step

In the second step, we estimate the total number of potential income tax payers in the private sector for each income bracket by applying the number of private formal sector workers from business census to the distribution of formal sector employees, which we derive in the first step. The resulting distribution of private formal sector workers include those who did not file income tax, thus are outside of the administrative income tax data.

Czajka (2016) also applies Fournier's estimation method to tax administrative data, combines it with household survey data, and re-estimates the Gini coefficient for Cote d'Ivoire. Czajka assume that the tax administrative data in Cote d'Ivoire completely covers

---

[3] GRA is going to provide us with the mean income for each income bracket. Once we receive the data, we will replace the mean income from the administrative data with the ones we are currently using.

all formal sector workers in the economy. From the analysis in Section 2, we are aware that it is not the case in Ghana. The tax administrative data suffers from missing individuals who failed to file income tax. For that reason, we adjust the number of the private formal sector workers using the business census. Since we do not know the number of public sector workers, we assume GLSS 6 data is closest to the true distribution of the public sector workers.

We first replace the wage income of individuals in GLSS 6 and LFS with the wage distributions from the continuous generalized Pareto curves for the public and private sectors separately. Note that it is not the national representative of wage workers, since the continuous generalized Pareto curves better captures top income, but underestimate low wage earners. However, it is closer to the distribution of potential taxpayers who are above the minimum threshold of income, subject to income tax payments. We follow Czajka (2016) and use household weights to match the workers in GLSS 6 and LFS to the distribution of wages from the generalized Pareto curves. The reason why we use the household weights from household surveys and replace the public sector workers in the household survey with the continuous generalized Pareto curves is because we believe household surveys are still closer to nationally representative distributions of workers.

Next, we use IBES data to adjust the number of private sector workers. The purpose is to add missing tax payers in the private sector. We assume that the proportion of missing private sector workers in each income bracket are identical across income bracket. Since there are 2.766 million wage earners who do not know whether they are working in the formal or informal sector in GLSS 6, we randomly select them and add them to the pool of formal private sector workers without changing the distributions of workers of the continuous generalized Pareto curves for each income bracket. We understand this is a strong assumption, and the simulation results heavily depend on the assumption of the distribution of missing tax payers. We will discuss alternative assumptions we can make in the conclusion section.

Table 6 contains the number of wage earners, their distribution across income brackets and mean income in each bracket before and after the expansion. There is no significant difference between mean wage income in each bracket before and after the expansion. The number of wage earners increases from 1.069 million to 1.423 million after the expansion.

For 2016, we do not have enough workers in the "unknown" category. So, we increasing the number of people proportionally. Table 7 contains the number of wage earners, their distribution across income brackets and mean income in each bracket before and after the expansion for 2016. As you see, the distribution was increased by the same proportion in each income bracket.

In the following result section, we conduct two types of simulations which correspond to First and Second Steps, respectively. Simulation I uses the distribution of formal sector workers derived in First Step, while Simulation II uses the distribution of formal sector workers derived in Second Step.

Simulation I
The first simulation (Simulation I) uses the distribution of wage earners we derived in Step One. The gap between the expected income tax revenue under Simulation I and the actual income tax revenue corresponds to the loss of tax revenue due to under or no payment of income tax among the formal sector wage earners who filed income tax. For the public sector, GRA keep the record of the number of registered workers, and uses that number to calculate income tax due. So it is not the individuals who actually filed income tax.

Simulation II
The second simulation (Simulation II) uses the distribution of wage earners we derived in Second Step. It the approximates the loss of income tax revenue due to nonpayment of income tax by individuals who did not file income tax even though their organizations and firms are registered with GRA. One shortcoming of Simulation II is it does not take into account the individuals who are exempted from income tax. The Public Expenditure Review of the World Bank (2017) point out that the income tax exemption is equivalent to 0.1 percent of GDP.

3. Results

Let us first compare the actual tax revenues with potential tax revenues under the assumption that all wage earners within each income bracket earn the minimum (or maximum) income of that bracket. Theoretically, we expect the actual tax revenue to be between the two estimated revenues.

Figure 6 shows the actual tax revenue and the estimated tax revenues at minimum and maximum levels of income for each income bracket for 2014. In the public sector, there are four income brackets in which actual tax revenues are smaller than the estimated tax revenues under the assumption that every wage earner in the bracket earned the minimum wage of the corresponding bracket. The income brackets cover gross income from 4201 to 30,000 cedi. This implies income tax was not fully paid by taxpayers in these income brackets. Consequently, the total tax revenue in the public sector was only 461 million cedi, which is much lower than the estimated total tax revenue under the assumption that people made the minimum income in their respective income brackets (743 million cedi).

In the private sector, we do not find significant underpayments of income tax in 2014. On the contrary, people in the income bracket of 30,0001 to 42,000 slightly overpaid income tax, as the actual tax revenue for that bracket is slightly higher than the estimated tax revenue under the assumption that people made maximum income within the brackets. In other income brackets, the actual tax revenues are between the two estimated tax revenues.

The last graph in Figure 6 summarizes the total tax revenue and its comparison with the estimated tax revenues at minimum and maximum levels of income for each income bracket. Between the income brackets of 4201 to 18,000 cedi, we observe underpayments of income tax, while we observe an overpayment of income tax for the income bracket of 30,001 to 42,000 cedi. This reflects the patterns of underpayment of income tax in the public sector and overpayment of the private sector for these income brackets.

Figure 7 shows the actual tax revenue and the estimated tax revenue at minimum and maximum levels of income for each income bracket for 2016. It is clear from the first graph on the public sector that the underpayment of income tax in the public sector disappeared in 2016. Moreover, we observe overpayments of income tax in the private sector. Thus, it seems that the large tax evasion we observe in 2014 disappeared in 2016. However, as shown in Table 3, the number of private formal sector workers reduced from over 1 million to about 398 thousand between 2014 and 2016. We will conduct simulations below and show that the types of tax evasions are different between 2014 and 2016.

Simulation I

<u>2014</u>
With the estimated continuous distribution of wage income earners, we estimate expected tax revenue for each income bracket, under the assumption that there was no underpayment of income tax among the formal sector workers who filed income tax. Figure 8 shows the estimation results. It indicates the total income tax revenue will be 22 percent higher than the actual tax revenue if there was no underpayment among people who filed income tax. It is equivalent to 0.5 percent of GDP.

The source of non-compliance is observed predominantly in the public sector. The total income tax revenue from the public sector will be 168 percent higher than the actual tax revenue if there was no underpayment among people who filed income tax in the public sector. As shown in Figure 8, the gap between the expected tax revenue in Simulation I and the actual tax revenue is high for the income brackets between the income of 4,201 to 18,000 and above 42,000 Cedi. The expected tax revenue is higher than the actual income tax revenue by the factor of 3.59 for the income bracket of 4,201 to 7,800. For the income bracket of 7,801 to 12,000, the expected tax revenue is higher than the actual income tax revenue by 9.14. For the income bracket of 12,001 to 18,000, the expected tax revenue is 2.95 times higher than the actual income tax. For the income bracket of above 42,000, the expected tax revenue is 2.26 times higher than the actual income tax.

The research staff at GRA explain that the payments of income tax come as aggregated amounts from government offices, and no information of individuals who contributed to the income tax payments is provided. Therefore, GRA cannot tell who contributed to the income tax and how much each individual contributed. In addition, it is reported that the government has recently removed 50,000 'ghost names' from its payroll.[4] It seems that the tax evasion in the public sector results from the lack of transparency, and the existence of ghost workers in the public sector.

<u>2016</u>
The results of Simulation I for 2016 are shown in Figure 9. Simulation results demonstrate that the private sector workers overpaid tax in 2016. The actual tax revenue (1.925 billion

---

[4] Source: https://www.ghanabusinessnews.com/2017/04/14/ghana-government-removes-50000-ghost-names-from-payroll-imf/

cedi) in the private formal sector is larger than the theoretical tax revenue under the assumption that all taxpayers who filed income tax pay full amounts of the tax due (1.146 billion cedi). It may be because many private sector firms and organizations were not aware of the changes in the income tax rate, or they did not update the wage levels of individuals and paid income tax for updated wages.

In contrast to 2014, the tax evasion in the public sector largely disappeared in 2016. We are not sure why tax evasion vanished in the public sector in 2016. GRA research staff explain that the accounting system was simplified in 2016, and the income tax was sent to a single account, instead of multiple accounts. It may have reduced no-payments or under-payments of PAYE to GRA.

Simulation II

2014
The second simulation examines the other type of tax evasion, that is, nonpayment of income tax among employees who work for the firms and organizations which are registered with GRA, but did not file income tax. This type of tax evasion is expected more in the private sector than the public sector, since GRA already has the number of public sector workers in their record. The number of public sector workers is likely to be overestimated, instead of underestimated, as there exist a number of ghost workers in the public sector. So, we make the adjustment of formal sector workers only in the private sector, as shown in Table 6. Having made the adjustment of the number of wage earners in each income bracket in the private sector, we estimate potential tax revenue for each income bracket and in each sector. The results of Simulation II for 2014 are shown in Figure 8. The results suggest the total income tax revenue could increase by 53 percent if all formal sector workers pay income tax. It is equivalent to 1.2 percent of GDP.

The potential increase in tax revenue is higher in the public sector than the formal private sector, because the estimated evasion under Simulation II includes the evasion estimated in Simulation I. The potential growth in income tax revenue is 221 percent and 28 percent in the public sector and the formal private sector, respectively.

2016
The results of Simulation II for 2016 are shown in Figure 9. We use the expanded number of wage earners presented in Table 7. The simulation results suggest the total income tax revenue could increase by 86 percent if all formal sector workers registered with GRA paid income tax in 2016. It is equivalent to 1.7 percent of GDP.

The potential increase in tax revenue is higher in the private sector than the public sector in 2016, since the number of taxpayers dramatically declined in the private sector in 2016. The potential growth in income tax revenue is 22 percent and 172 percent in the public sector and the formal private sector, respectively.

Simulation III

11

We conduct one more simulation (Simulation III) for 2016. The third simulation (Simulation III) assumes that the number of taxpayers remains the same between 2014 and 2016. The simulation yields a modest increase in the potential income tax revenue (See Figure 9). The total income tax revenue could increase by 54 percent if all formal sector workers who paid income tax in 2014 paid income tax in full amounts in 2016. The potential increase in tax revenue is higher in the private sector than the public sector again, since the number of taxpayers dramatically declined in the private sector between 2014 and 2016. The potential increase in income tax revenue is 6 percent and 118 percent in the public sector and the formal private sector, respectively.

4. Conclusion and policy recommendations

This paper examines the scale of tax evasion and potential revenue gains from the enforcement of tax compliance in Ghana.

By combining the household survey data and the administrative income tax data, and adjusting the number of formal sector workers using the business census, we reconstruct a nationally representative distribution of wage earners in the formal sector, and estimate potential income tax revenue gain from the enforcement of tax compliance for 2014 and 2016. We find the income tax revenue could have been higher by 53 percent (which is equivalent to 1.2 percent of GDP) if all formal sector firms and organizations had paid full amounts of PAYE for all their employees in 2014. The source of non-compliance is observed predominantly in the public sector. The total income tax revenue from the public sector will be 168 percent higher than the actual tax revenue if there was no underpayment among people who filed income tax in the public sector. It seems that the tax evasion in the public sector results from the lack of transparency, and the existence of ghost workers in the public sector.

In 2016, there was a further reduction of income tax revenue. Our simulation results suggest the total income tax revenue could have been higher by 86 percent (which is equivalent to 1.7 percent of GDP) if all formal sector firms and organizations paid the full amounts of PAYE for all their employees. In 2016, tax evasion took in one form, the presence of a larger proportion of private sector wage earners who did not file income tax. On the other hand, private formal sector workers who filed income tax overpaid the tax. The underpayment of income tax we observed in 2014 largely disappeared in 2016. We speculate that the tax compliance improved in the public sector because the income tax collection system was simplified, making it harder for governmental organizations to evade. In contrast, the number of private formal sector workers who filed income tax significantly reduced between 2014 and 2016. Our simulation results indicate that if the same number of people filed income tax in 2016 as 2014, and if they have fully paid the tax, then the total income tax revenue could have been higher by 54 percent, which is equivalent to 1.0 percent of GDP.

Our simulation results indicate improving tax compliances among the registered formal sector firms and organizations can improve tax revenue dramatically. The government needs to improve its incentive mechanisms among taxpayers as well as tax collectors. The concern on tax compliances is as important or even more important in developing countries than developed countries as they have been struggling to increase the tax revenues.

Hallsworth (2014) reviews recent impact evaluations in tax compliances in developed countries. He separates interventions into two: deterrence and non-deterrence approach. Deterrence approach includes the use of audits and the introduction of fines. Non-deterrence approach includes various interventions such as the use of norms and fairness.

Overall, most deterrence interventions that increased the perceived enforcement or perceived severity of the enforcement increased tax compliance (for example, Kleven et al. (2011) and Harju et al. (2013)). The typical treatments under the deterrence approach include sending a letter to randomly selected taxpayers along with their bills, that impose a threat of an audit by informing them about the possibility of them being audited. Kahn, et al. (2001) demonstrate an increase in fines improves tax enforcement in Brazil.

On the other hand, interventions using non-deterrence have mixed results in tax compliance (for example, Del Carpio (2013); Wenzel and Taylor (2004); Ariel (2012)). Del Carpio (2013) studies the effect of norms on tax compliance by giving tax payers the information on average compliance rates, and find positive effects of the interventions. Blumenthal, Christian and Slemrod (2001) evaluate the effect of a letter that uses normative appeals to taxpayers. They find no behavioral change in tax compliance. Ariel (2012) finds sending letters that use moral persuasion increases the tax deductions among corporations, which reduces the tax revenues of the government.

Although there are a number of studies on tax compliances in developed countries as discussed above, empirical evidence is limited in developing countries. In addition, developing countries face additional challenges such as weak institutions, corruption, and informal economy. This study shows weak institutions cause tax evasion not only in the private sector, but also in the public sector.

Thus it is important to evaluate whether a particular intervention that works in developed counties also work in developing countries, and also in the public sector. Even if prior studies suggest that deterrence-approach increases the tax compliance in developed countries, it should be noted that conducting such interventions is costly and providing the right incentives to tax officials to put their efforts to monitor can be challenging in developing countries (Ahmed et al. (2012)). The deterrence approach seems to work well where there is rigid law enforcement. In countries with weak institutions, such approach might not be effective. More empirical evidences are needed from developing countries before we can make solid policy recommendations for developing countries.

As discussed in introduction, the availability of administrative income tax data is limited in developing countries. It makes it extremely difficult to understand the nature of tax evasion in developing countries. The availability of administrative income tax data together

with rich household and business census data makes Ghana a great example to study tax compliances. This paper shows income tax compliances have significant impacts of tax revenue, however, the simulation results heavily depends on the assumption and the quality of data. Further improvement of assumptions and data quality is desirable to estimate the scale of non-compliance more accurately. We used the mean income from household surveys for simulation. Ideally, the mean income should be calculated from the administrative income tax data. Also, we assumed that the number of missing private sector workers are distributed proportionally across income brackets. It is possible that missing tax payers exists more in lower income brackets. If this is true, then Simulation II overestimate tax evasion. We need to conduct further simulations and test other assumptions.

Income tax compliance has a great redistributive effect as income tax is disproportionally paid by top the top segment of the society. Investigating redistributive effects and income inequality is left for future research.

References

Ahmed, N., R. Chetty, M. Mobarak, A. Rahman and M. Singhal (2012). "Improving Tax Compliance in Developing Economies." International Growth Center Working Paper.

Ahmed, R. A. and M. Rider (2013). "Using microdata to estimate pakistan's tax gap by type of tax." Public Finance Review **41**(3): 334-359.

Alvaredo, F. and L. Gasparini (2013). Recent trends in inequality and poverty in developing countries, Documento de Trabajo.

Alvaredo, F. and J. Londoño (2013). High incomes and personal taxation in a developing economy: Colombia 1993-2010, CEQ Working Paper.

Alvaredo, F. and T. Piketty (2010). "The dynamics of income concentration in developed and developing countries: a view from the top." Declining inequality in Latin America: A decade of progress: 72-99.

Alvaredo, F. and T. Piketty (2014). Measuring top incomes and inequality in the Middle East: Data limitations and illustration with the case of Egypt, CEPR Discussion Paper

Ariel, B. (2012). "Deterrence and moral persuasion effects on corporate tax compliance: findings from a randomized controlled trial." Criminology **50**(1): 27-69.

Besley, T. and T. Persson (2014). "Why do developing countries tax so little?" The Journal of Economic Perspectives **28**(4): 99-120.

Blanchet, T., J. Fournier and T. Piketty (2017). Generalized Pareto Curves : Theory and Applications, Paris School of Economics.

Blumenthal, M., C. Christian, J. Slemrod and M. G. Smith (2001). "Do normative appeals affect tax compliance? Evidence from a controlled experiment in Minnesota." National Tax Journal **54**(1): 125-138.

Czajka, L. (2016). Combining Fiscal and Survey Data to Measure Inequality : Ivory Coast 2014, Paris School of Economics.

Del Carpio, L. (2013). Are the neighbors cheating? Evidence from a social norm experiment on property taxes in Peru, Princeton University Working Paper.

Fisman, R. and S.-J. Wei (2004). "Tax rates and tax evasion: evidence from "missing imports" in China." Journal of political Economy **112**(2): 471-496.

Fournier, J. (2015). Generalized Pareto curves: Theory and application using income and inheritance tabulations for France 1901-2012, Master's thesis, Paris School of Economics.

Hallsworth, M. (2014). "The use of field experiments to increase tax compliance." Oxford Review of Economic Policy **30**(4): 658-679.

Harju, J., T. Kosonen and O. Ropponen (2013). "Do honest hairdressers get a haircut? On tax rate and tax evasion." Oxford Review of Economic Policy **30**(4): 658–679.

Kahn, C. M., E. C. Silva and J. P. Ziliak (2001). "Performance-based wages in tax collection: The Brazilian tax collection reform and its effects." The Economic Journal **111**(468): 188-205.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen and E. Saez (2011). "Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark." Econometrica **79**(3): 651-692.

Wenzel, M. and N. Taylor (2004). "An experimental evaluation of tax-reporting schedules: a case of evidence-based tax administration." Journal of Public Economics **88**(12): 2785-2799.

Table 1: Comparison of tax revenue in 2012

	Revenue as % GDP		
	VAT	Corporate Tax	Income Tax
Low income	5.04	2.24	3.30
Lower-middle	5.99	3.51	4.38
Upper-middle	6.20	3.40	4.86
High incoe	6.60	3.65	8.35
Total	6.09	3.33	5.53
Ghana	4.65	2.87	2.52

Source: USAID tax database

Table 2: Annual income tax rates

Until 2015

Income	Rate
0- 1,584	0
1,585-2,376	5
2,377-3,480	10
3,481-31,680	17.5
31,681 and above	25

From 2016

Income	Rate
0- 2,592	0
2,593-3,888	5
3,889-5,700	10
5,701-38,880	17.5
38,881 and above	25

Table 3: Comparison of number of wage earners

	Public	Private-Formal	Private-Inform	Unknown
2013/14 Income tax data	503,913	1,145,183	-	-
2013 IBES	-	1,417,368	1,009,632	-
2012/13 GLSS 6	617,972	1,068,925	1,247,116	2,766,436
	Public	Private-Formal	Private-Inform	Unknown
2016 Income tax data	512,228	309,763	-	-
2013 IBES	-	1,417,368	1,009,632	-
2015 Labour Force Survey	585,838	397,987	689,166	203,391

Table 4: Estimated mean income from GLSS 6 and Pareto coefficients

Public Sector

Income bracket	Estimated mean income	Inverted Pareto coefficient
0-1,980	980	7.87
1,981-4,200	3,063	3.85
4,201-7,800	5,942	2.35
7,801-12,000	9,578	2.18
12,001-18,000	14,586	2.73
18,001-30,000	23,257	2.23
30,001-42,000	36,008	2.38
above 42,000	99,931	-

Private Sector

Income bracket	Estimated mean income	Inverted Pareto coefficient
0-1,980	1,121	7.57
1,981-4,200	2,912	4.07
4,201-7,800	5,809	2.41
7,801-12,000	9,643	2.07
12,001-18,000	14,698	1.93
18,001-30,000	22,485	2.02
30,001-42,000	36,096	2.36
above 42,000	99,017	-

Total

Income bracket	Estimated mean income	Inverted Pareto coefficient
0-1,980	1,077	7.67
1,981-4,200	2,949	4.00
4,201-7,800	5,866	2.39
7,801-12,000	9,604	2.11
12,001-18,000	14,643	2.12
18,001-30,000	22,786	2.10
30,001-42,000	36,073	2.37
above 42,000	99,452	-

Table 5: Estimated mean income from the LFS and Pareto coefficients

Public Sector

Income bracket	Estimated mean income	Inverted Pareto coefficient
0-2,592	1,746	6.47
2,593-6,000	4,269	2.88
6,001-12,000	9,209	1.62
12,001-24,000	16,392	1.54
24,001-60,000	32,536	1.17
60,001-120,000	60,537	1.42
above 120,000	170,937	-

Private Sector

Income bracket	Estimated mean income	Inverted Pareto coefficient
0-2,592	1,756	8.68
2,593-6,000	4,318	5.11
6,001-12,000	8,338	3.54
12,001-24,000	16,181	2.31
24,001-60,000	36,176	1.97
60,001-120,000	69,921	1.90
above 120,000	227,996	-

Total

Income bracket	Estimated mean income	Inverted Pareto coefficient
0-2,592	1,753	7.21
2,593-6,000	4,306	3.50
6,001-12,000	8,845	2.09
12,001-24,000	16,356	1.96
24,001-60,000	33,837	1.74
60,001-120,000	66,451	1.81
above 120,000	217,612	-

Table 6: Estimated number of private formal sector workers, its distribution and mean income of each income bracket before and after the expansion for 2014

	Before expansion			After expansion		
	Number	Distribution	Mean income	Number	Distribution	Mean income
0-1,980	248,645	23.3%	1,121	335,253	23.6%	1,081
1,981-4,200	332,444	31.1%	2,912	439,308	30.9%	2,906
4,201-7,800	212,472	19.9%	5,809	285,867	20.1%	5,822
7,801-12,000	98,219	9.2%	9,643	122,462	8.6%	9,784
12,001-18,000	83,873	7.8%	14,698	111,261	7.8%	14,724
18,001-30,000	51,945	4.9%	22,485	68,705	4.8%	22,804
30,001-42,000	13,172	1.2%	36,096	16,418	1.2%	36,413
above 42,000	28,155	2.6%	99,017	43,692	3.1%	97,357
Total	1,068,925	100.0%		1,422,966	100.0%	

Table 7: Estimated number of private formal sector workers, its distribution and mean income of each income bracket before and after the expansion for 2016

Simulation II

	Before expansion			After expansion		
	Number	Distribution	Mean income	Number	Distribution	Mean income
0-2,592	45,048	11.3	1,756	160,431	11.3	1,756
2,593-6,000	135,508	34.1	4,318	482,589	34.0	4,318
6,001-12,000	92,299	23.2	8,338	328,708	23.2	8,338
12,001-24,000	61,395	15.4	16,181	218,648	15.4	16,181
24,001-60,000	48,783	12.3	36,176	173,733	12.3	36,176
60,001-120,000	9,214	2.3	69,921	32,814	2.3	69,921
above 120,000	5,740	1.4	227,996	20,442	1.4	227,996
Total	397,988	100%		1,417,364	100.0%	

Simulation III

	Before expansion			After expansion		
	Number	Distribution	Mean income	Number	Distribution	Mean income
0-2,592	45,048	11.3	1,756	129,623	11.3	1,756
2,593-6,000	135,508	34.1	4,318	389,915	34.0	4,318
6,001-12,000	92,299	23.2	8,338	265,584	23.2	8,338
12,001-24,000	61,395	15.4	16,181	176,660	15.4	16,181
24,001-60,000	48,783	12.3	36,176	140,370	12.3	36,176
60,001-120,000	9,214	2.3	69,921	26,513	2.3	69,921
above 120,000	5,740	1.4	227,996	16,516	1.4	227,996
Total	397,988	100%		1,145,180	100.0%	

Figure 1: Income tax revenue from public and private sectors
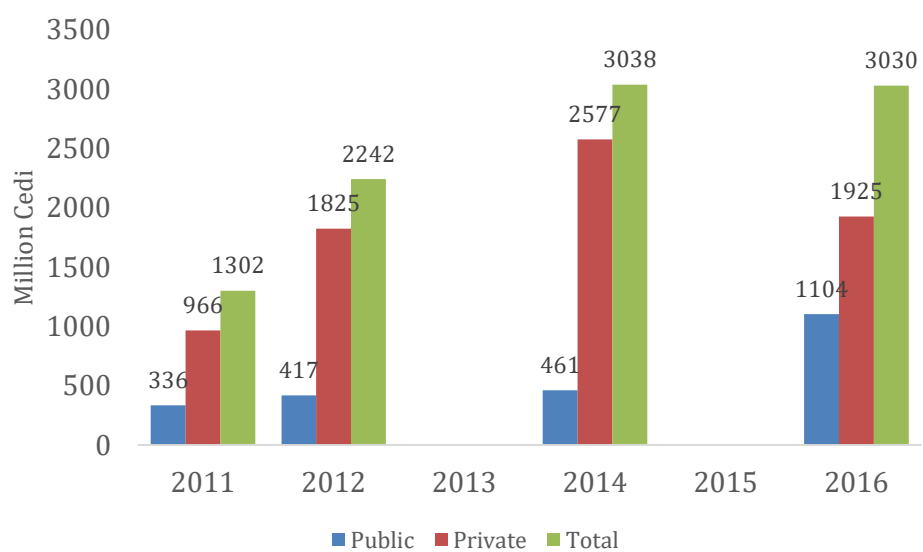
Figure 2: Comparison of public sector and formal private sector workers in income tax data and GLSS 6 data

Public
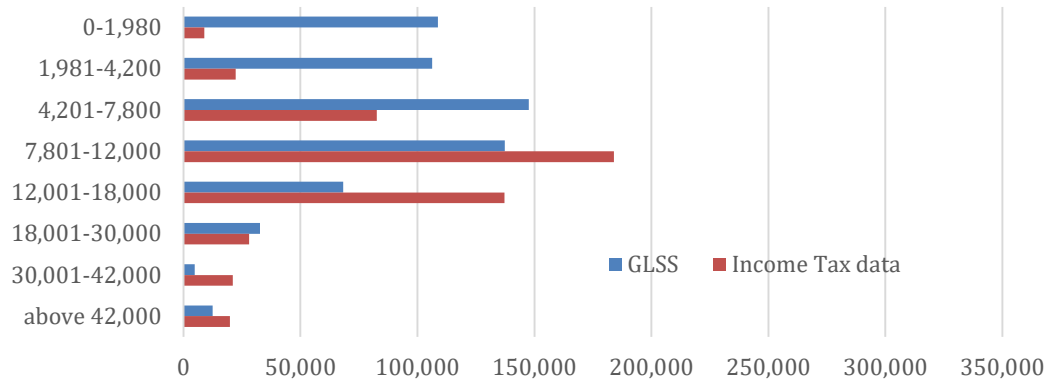
Private
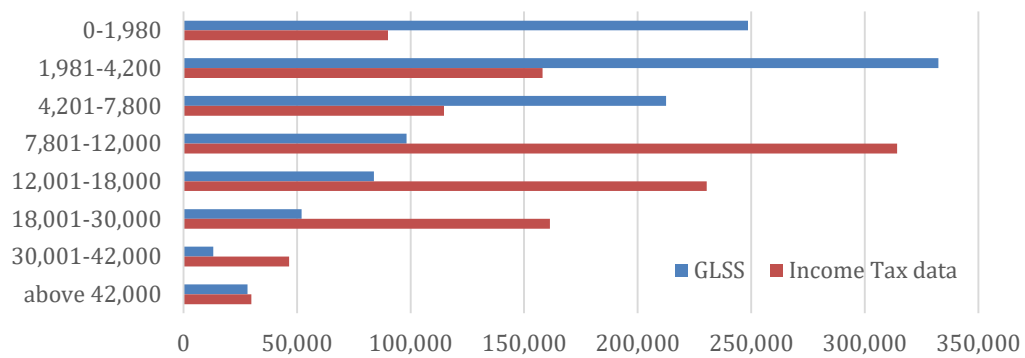
Figure 3: Comparison of public sector and formal private sector workers in income tax and LFS data
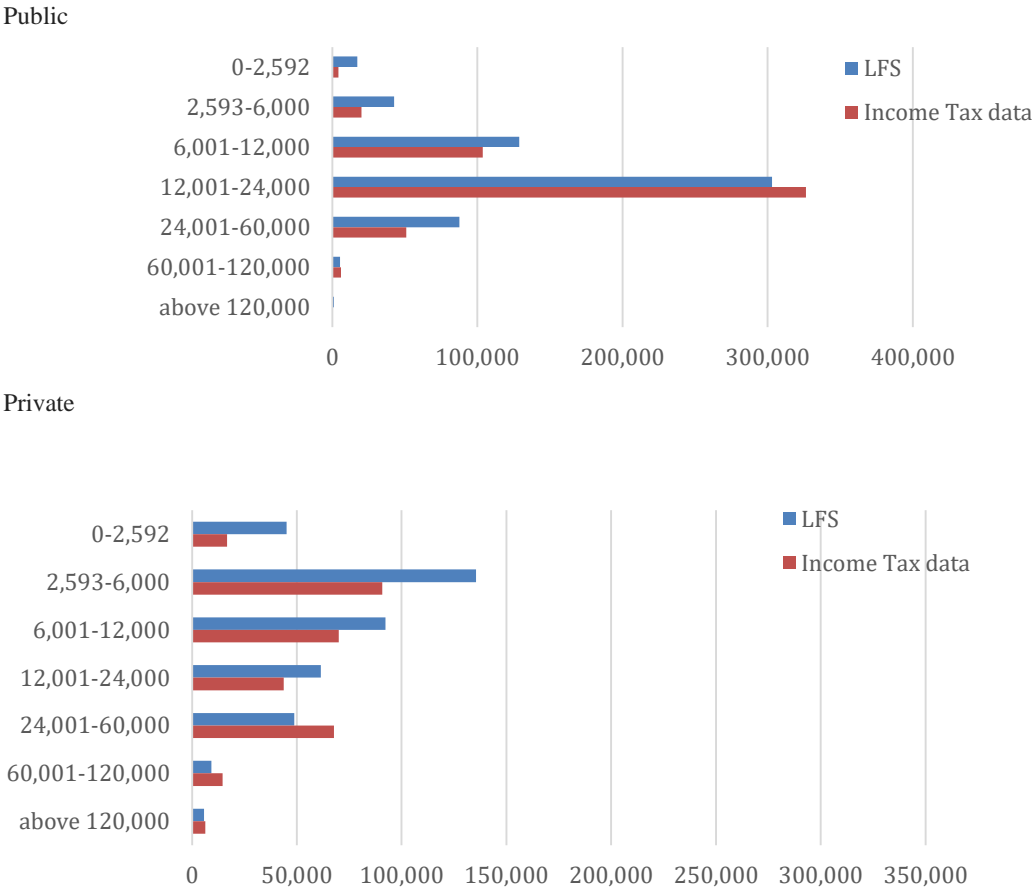
Public

Private

Figure 4: Distributions of wage earners for 2014 simulations
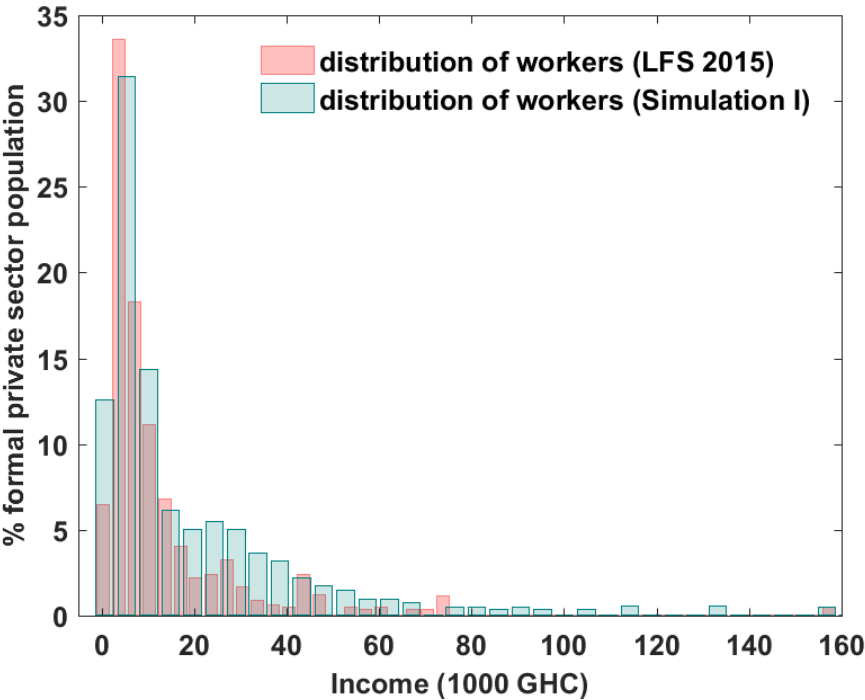
Public
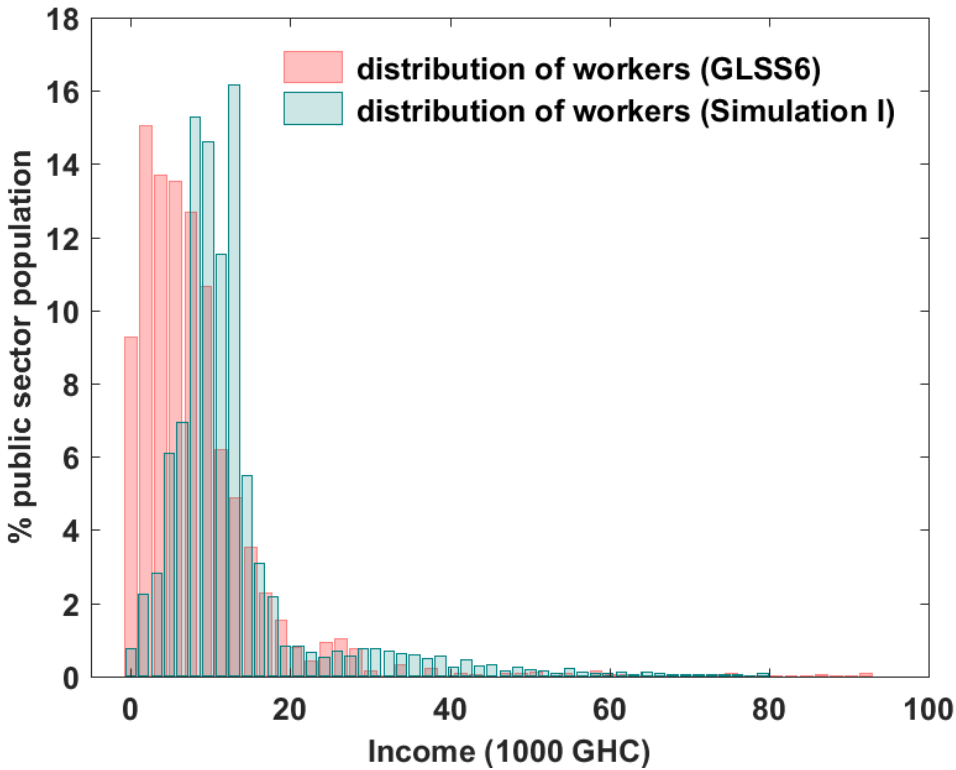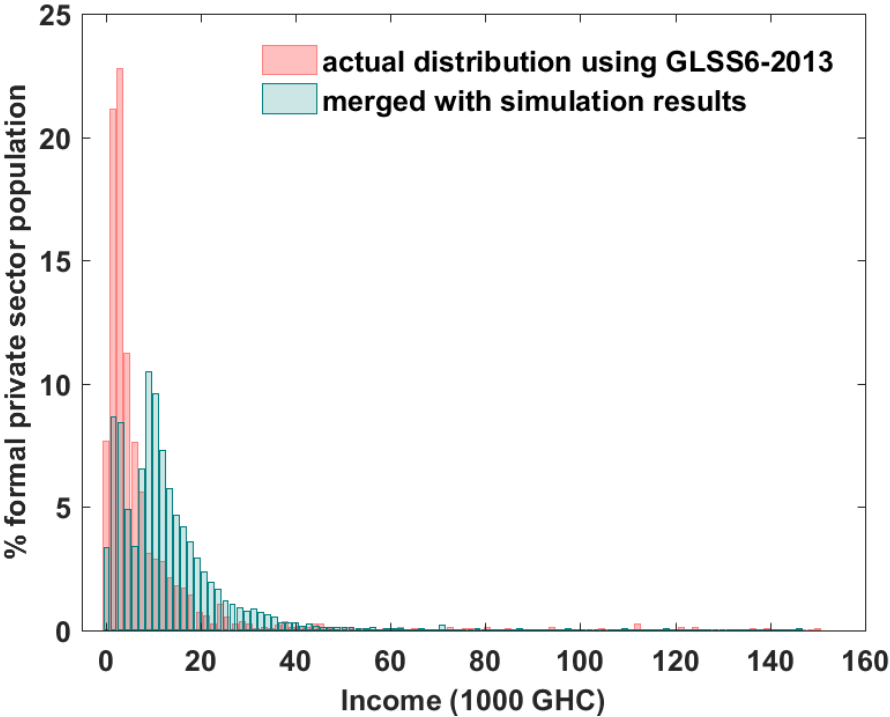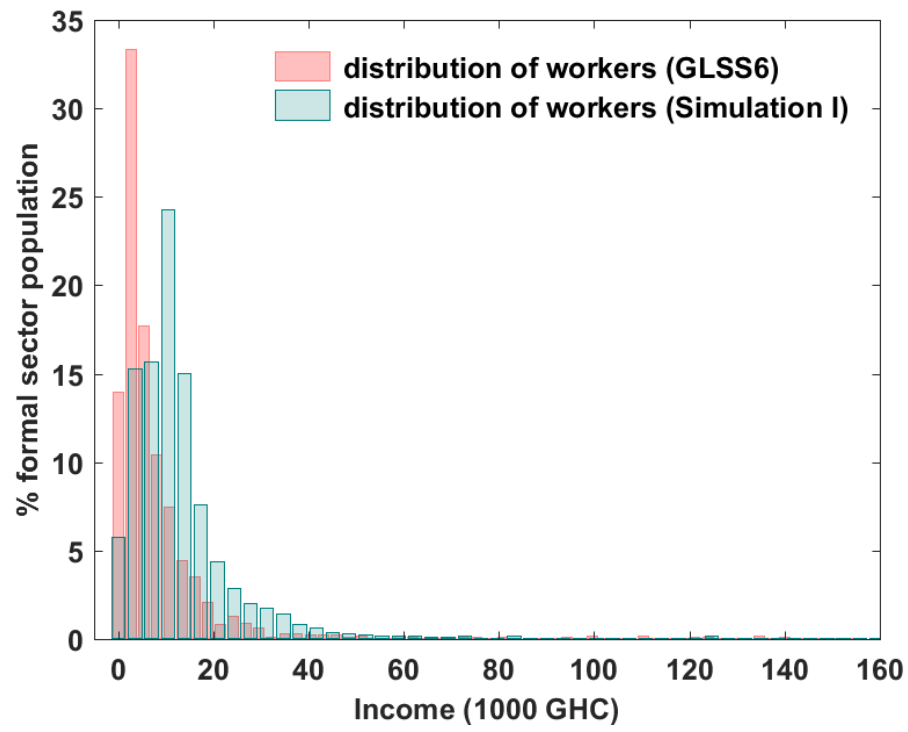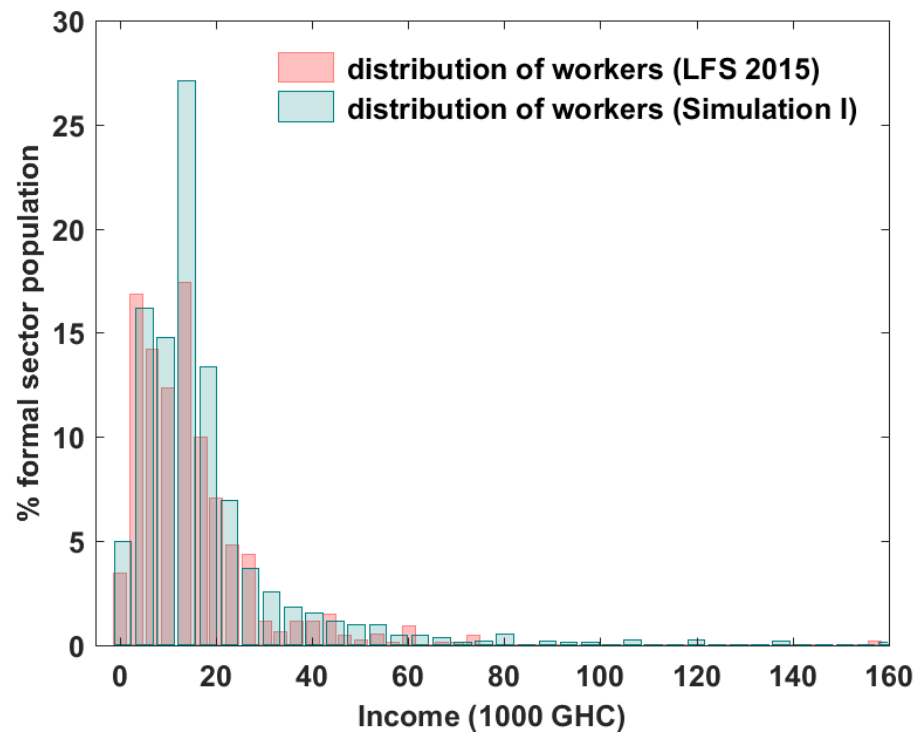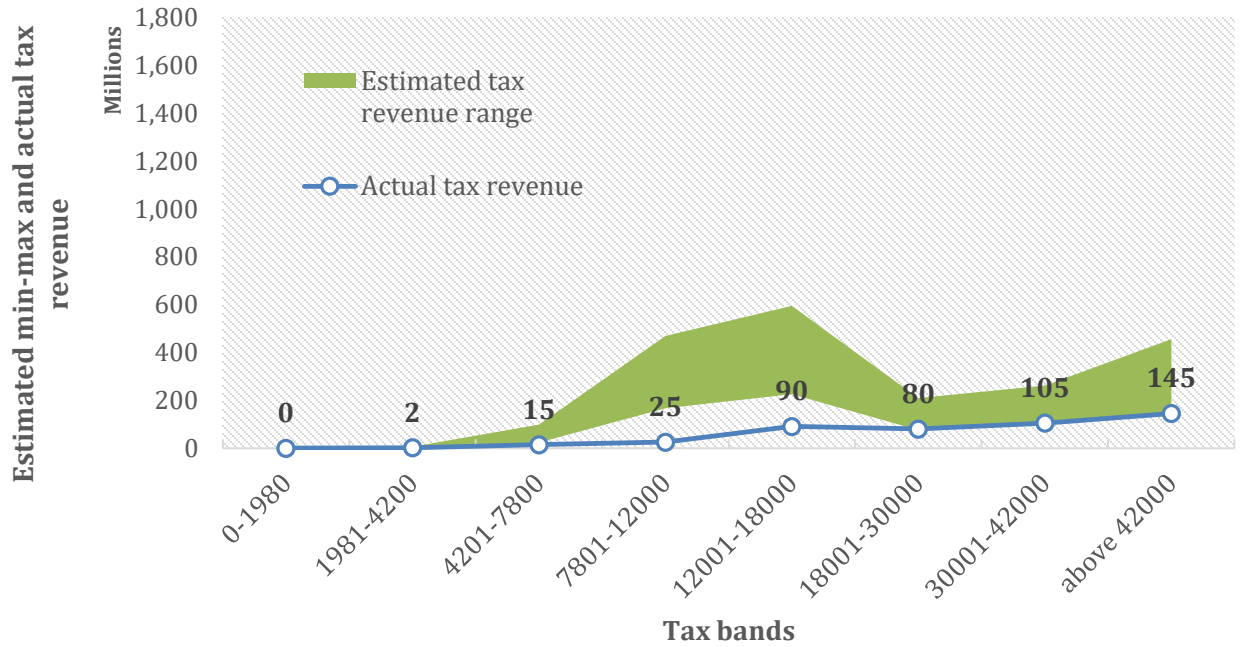

Private

Total

Figure 5: Income distributions of wage earners for 2016 simulations

Public
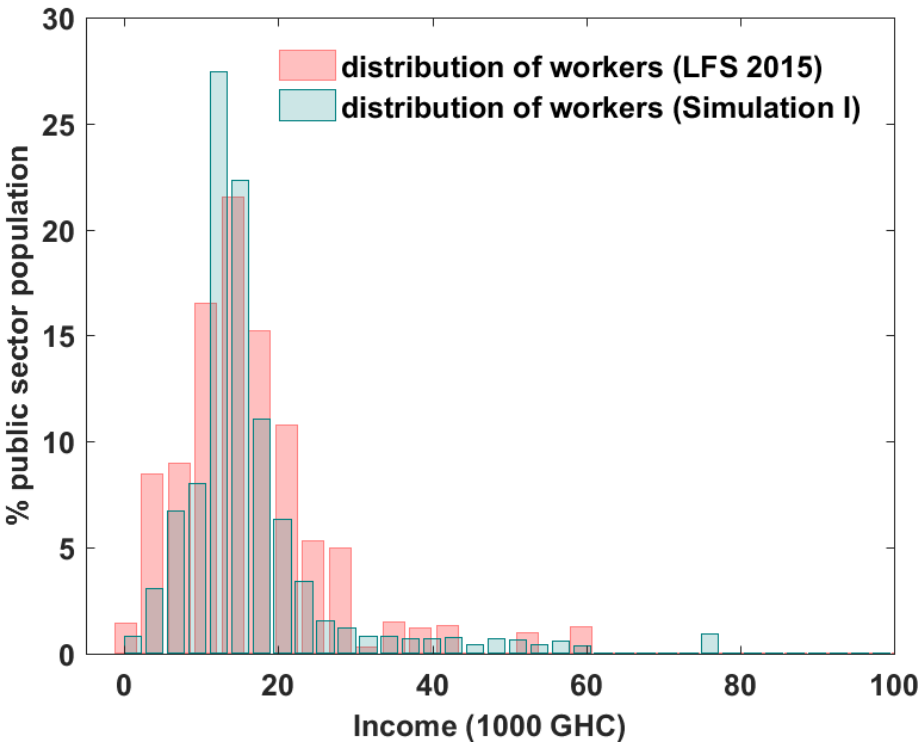
Private

Total

Figure 6: Comparison of actual revenue and estimated revenues under the assumption of min and max income within brackets (2014)
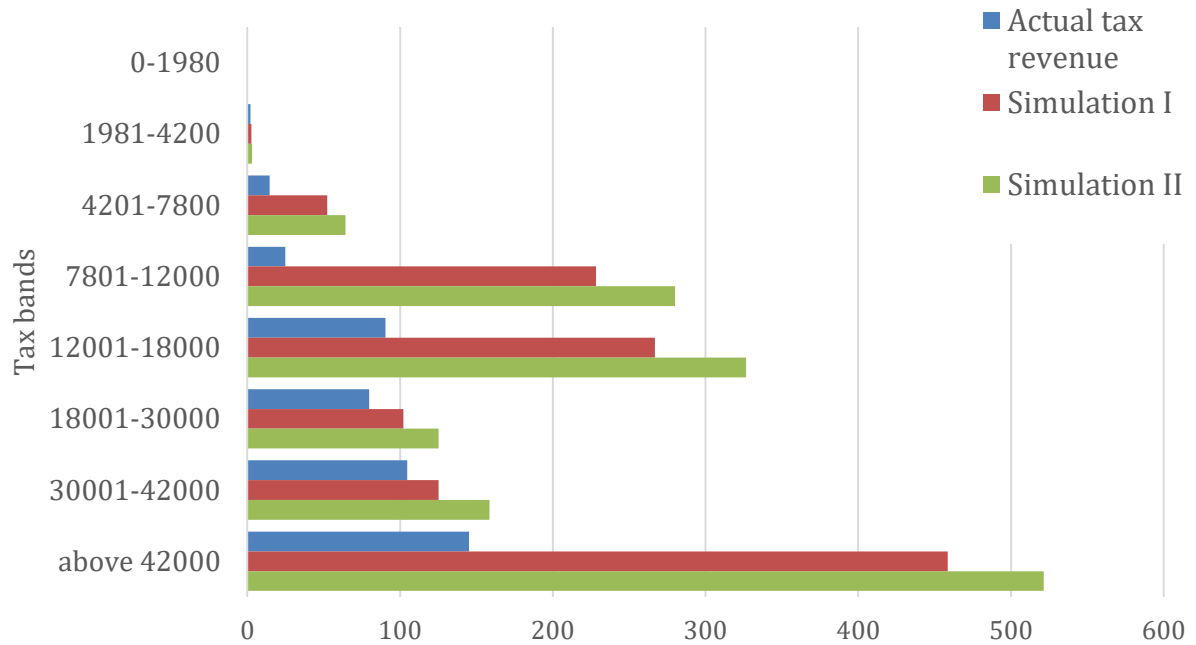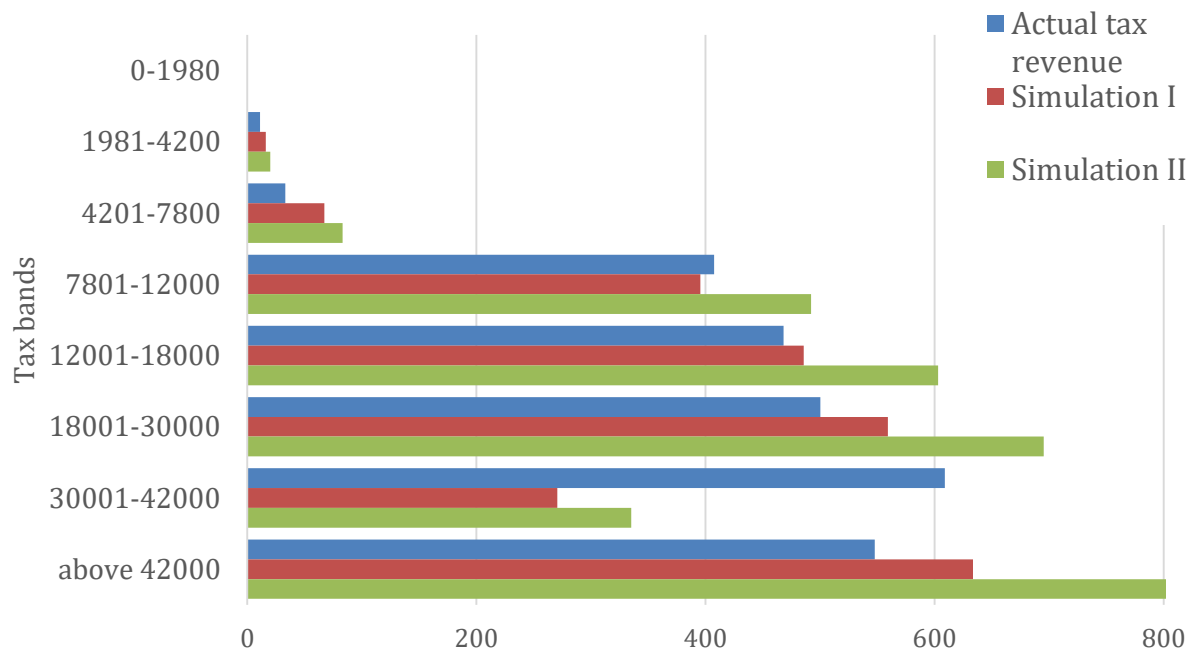
Public

Private
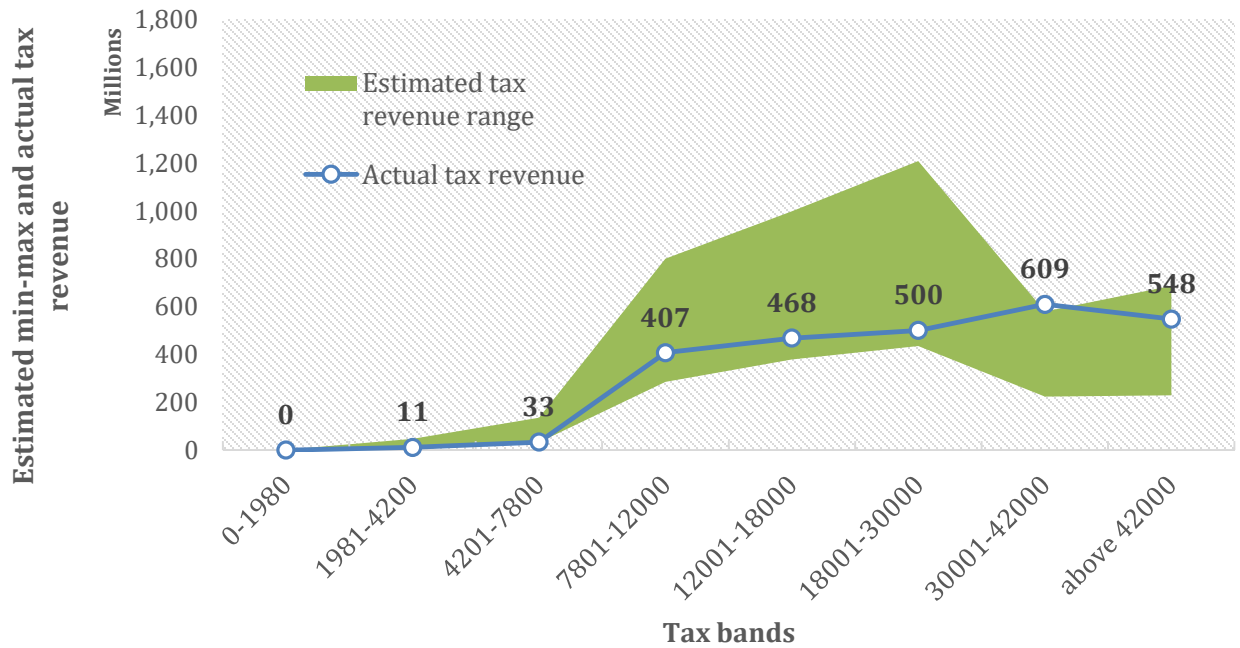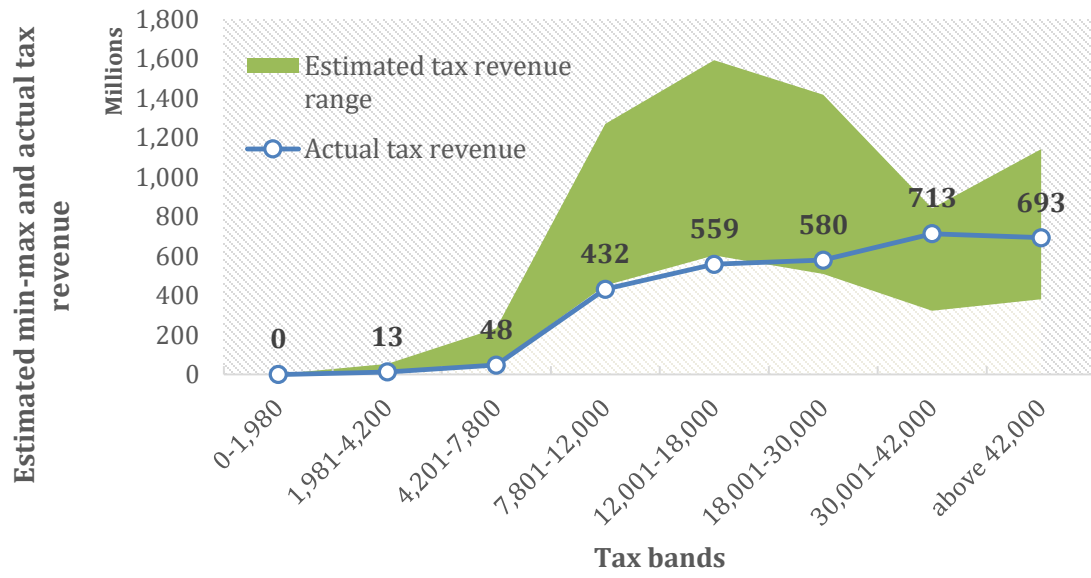
Total

Figure 7: Comparison of actual revenue and estimated revenues under the min and max income within brackets (2016)

Public
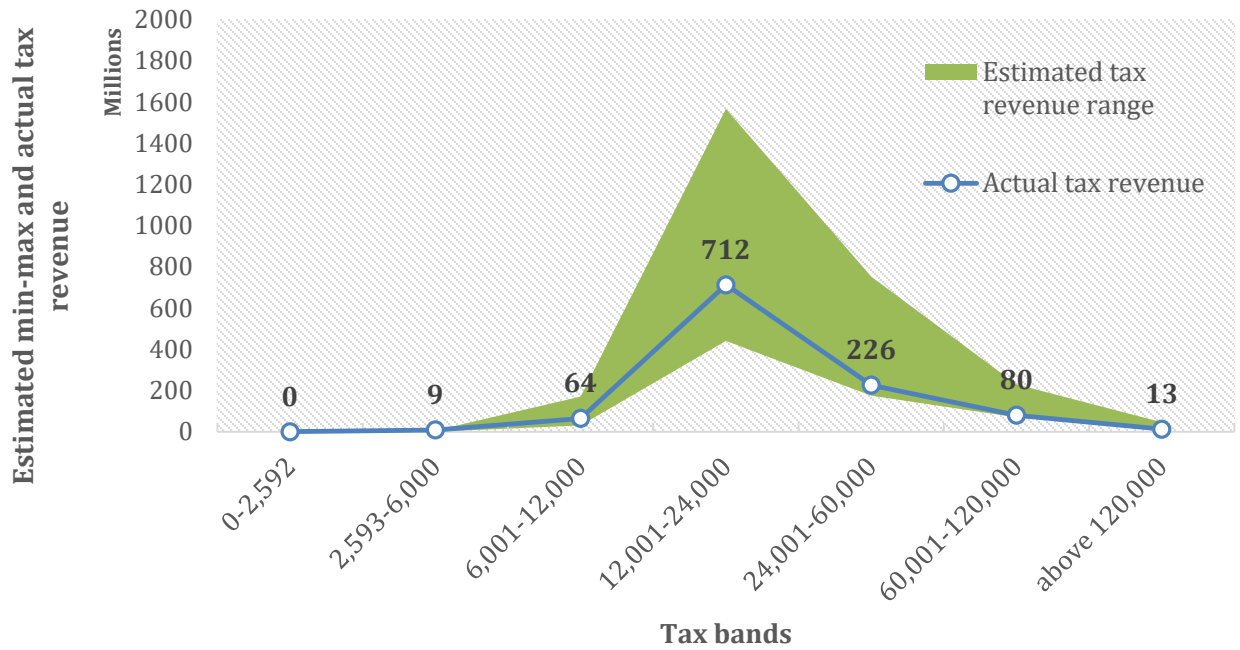
Private
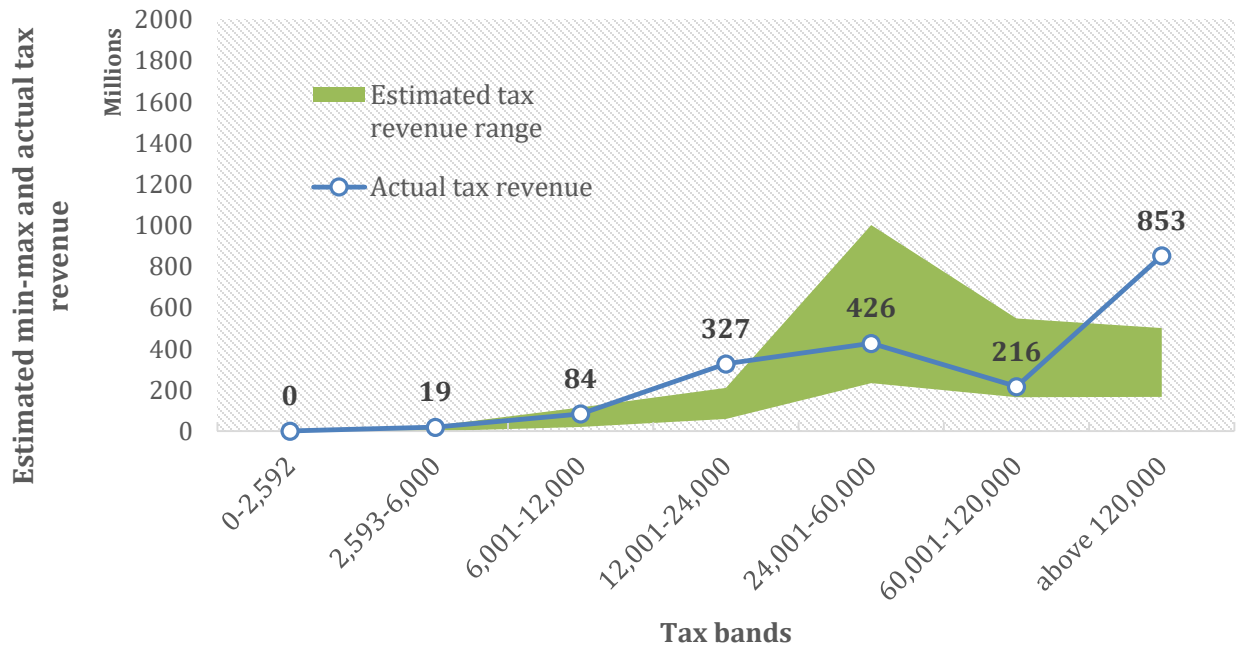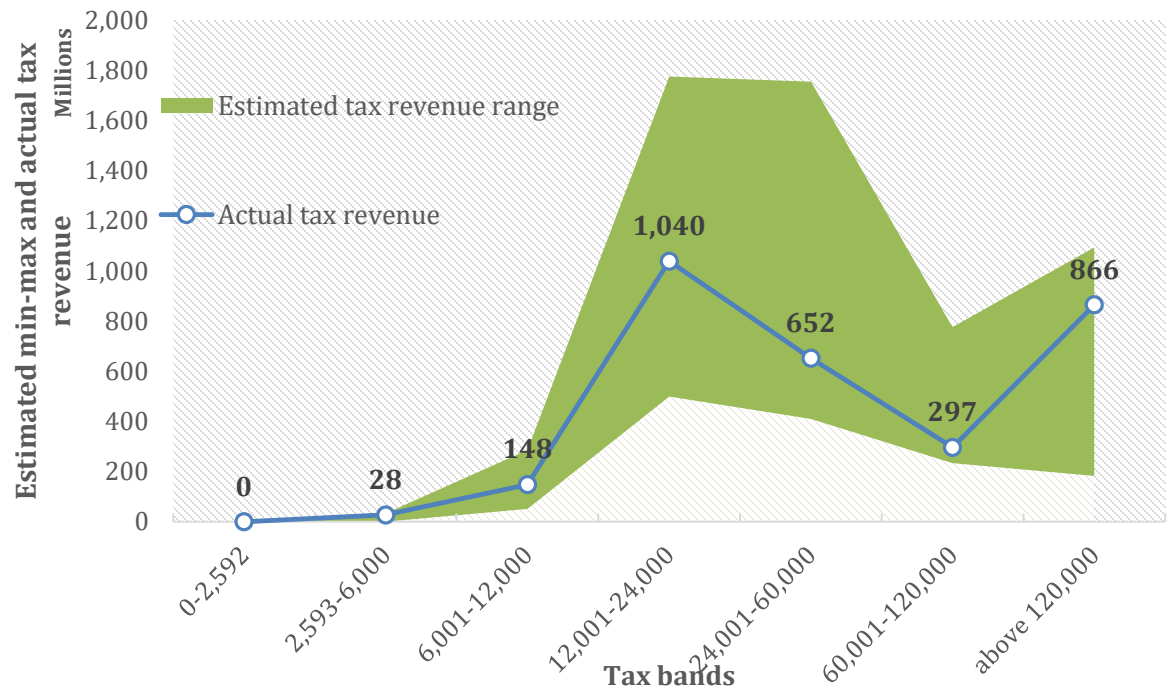
Total

Figure 8: Comparison of actual revenue and estimated revenues under simulations I and II (2014)
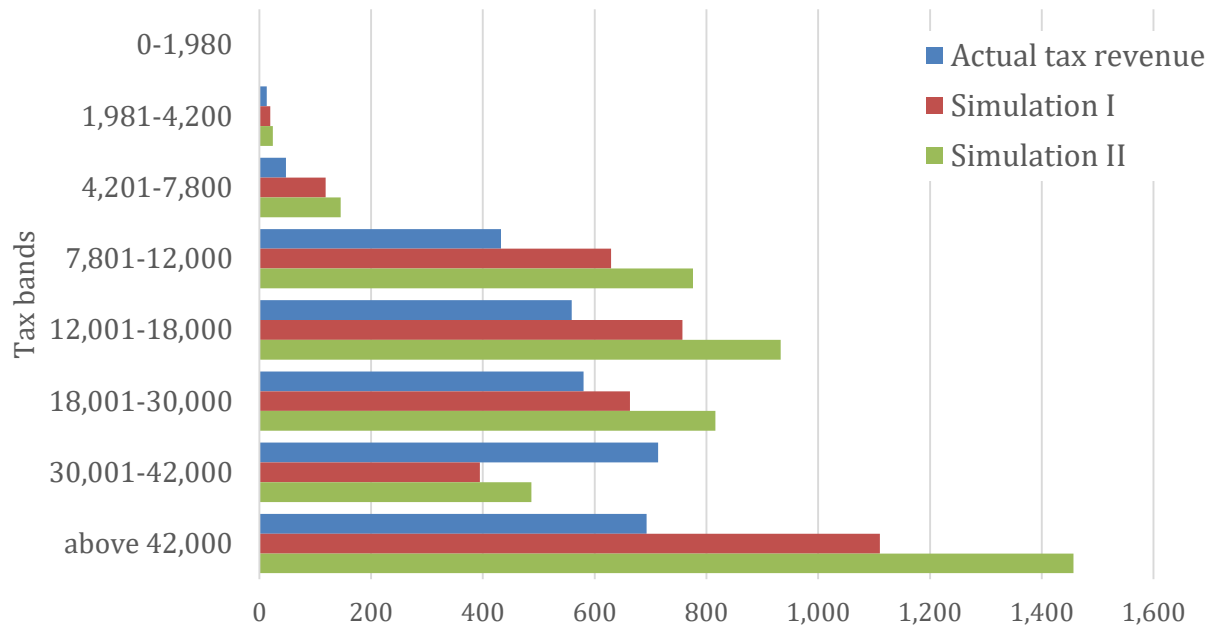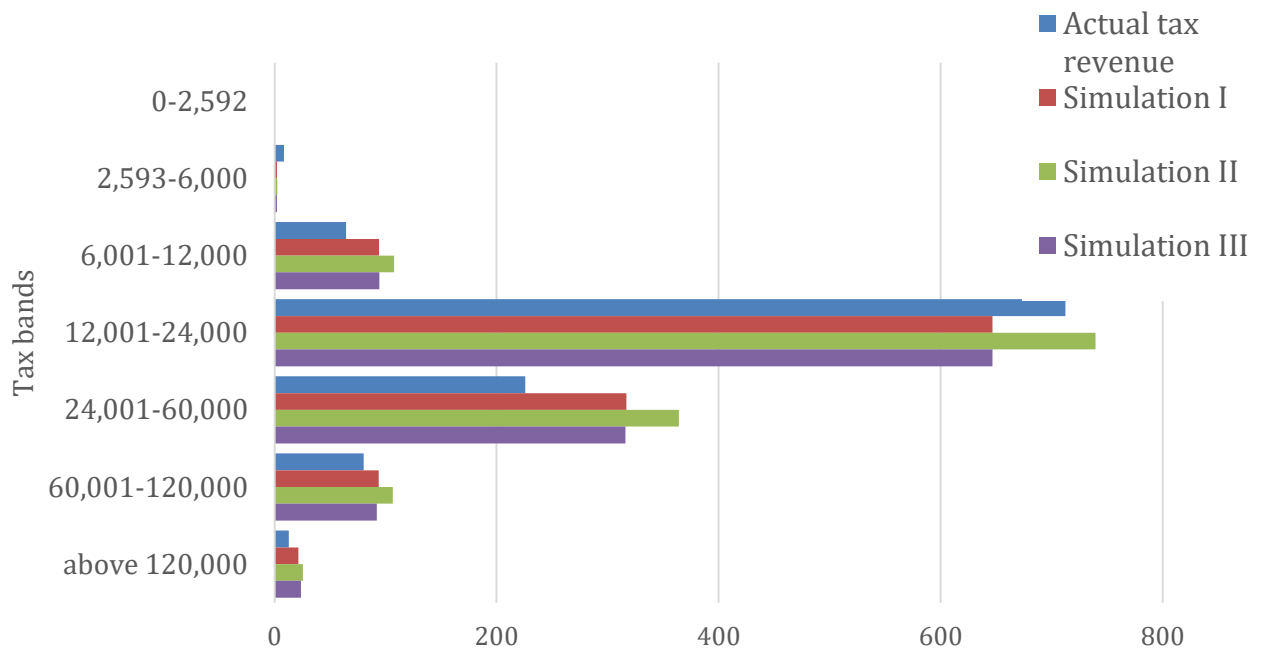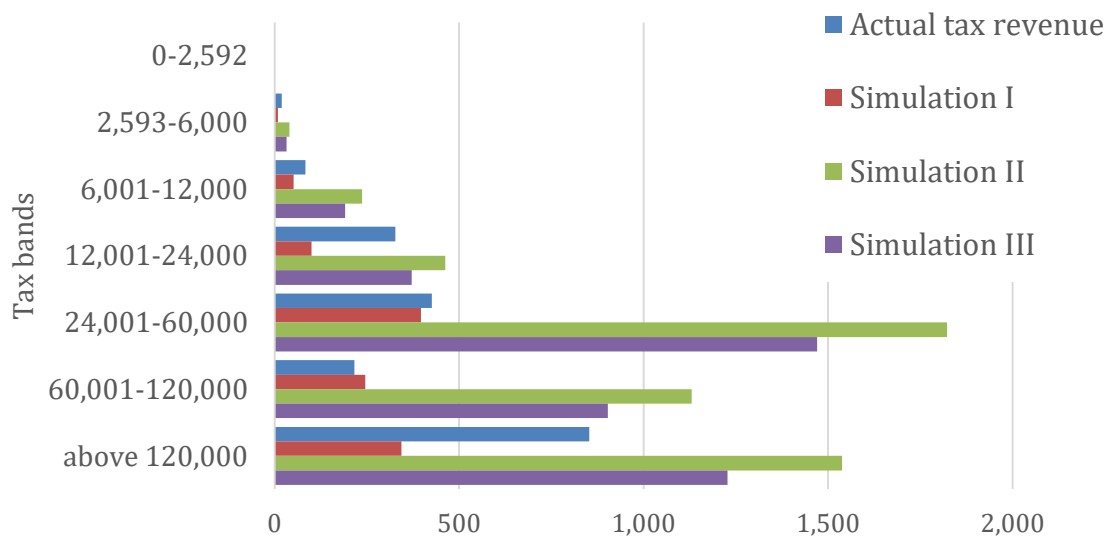
Public

Private

Total

Figure 9: Comparison of actual revenue and estimated revenues under simulations I, II and III (2016)
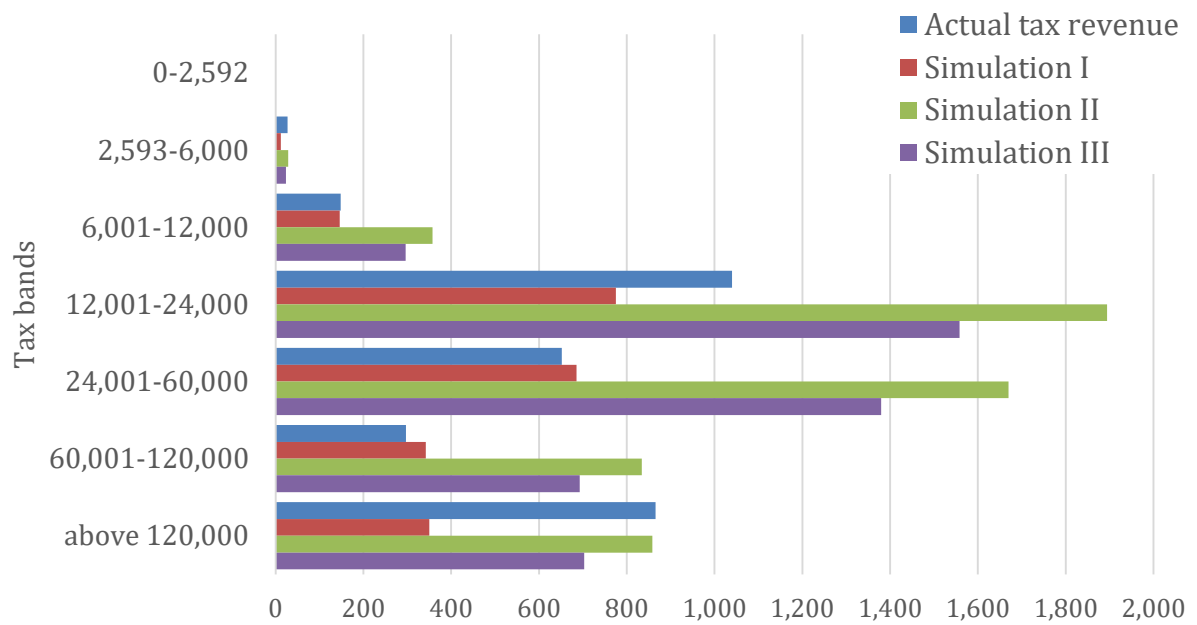
Public

Private

Total

Appendix

Table A.1: Number of establishments and permanent workers in the private sector by region in November 2013

Region	Number of establishments	Of which: registered with GRA (%)	Number of employees (Thousand)	Of which: Working for firms registered with GRA (%)
Western	59,978	33.7	239.1	56.6%
Central	50,198	39.6	151.4	48.6%
G. Accra	175,111	42.7	1,067.2	76.5%
Volta	39,217	18.2	102.8	27.9%
Eastern	55,464	30.5	168.3	42.2%
Ashanti	119,492	31.4	363.7	47.3%
Brong Ahafo	46,370	34.1	140.5	46.3%
Northern	36,760	14.7	108.1	21.5%
Upper East	15,418	22.9	50.3	37.3%
Upper West	12,400	22.7	35.4	31.1%
Total	610,408	33.4	2,427	58.3%

Source: IBES (2014)