TECHNICAL ASSISTANCE REPORT

REPUBLIC OF ARMENIA

Corporate Income Tax Gap Prediction 2023
Based on Corporate Income Tax Returns

SEPTEMBER 2024

Prepared By
Tobias Gabel Christiansen

CCAMTAC

MEMBERS
Armenia, Azerbaijan, Georgia, Kazakhstan, Kyrgyz Republic, Mongolia, Tajikistan, Turkmenistan, Uzbekistan

PARTNERS
Switzerland, Russia, China, Korea, United States, Asian Development Bank, Poland

DISCLAIMER

The contents of this report constitute technical advice provided by the staff of the International Monetary Fund (IMF) to the authorities of the Armenia (the "TA recipient") in response to their request for technical assistance. This report (in whole or in part) or summaries thereof may be disclosed by the IMF to IMF Executive Directors and members of their staff, as well as to other agencies or instrumentalities of the TA recipient, and upon their request, to World Bank staff, and other technical assistance providers and donors with legitimate interest, including members of the Steering Committee of CCAMTAC, unless the TA recipient specifically objects to such disclosure (see Operational Guidelines for the Dissemination of Technical Assistance Information). Publication or Disclosure of this report (in whole or in part) or summaries thereof to parties outside the IMF other than agencies or instrumentalities of the TA recipient, World Bank staff, other technical assistance providers and donors with legitimate interest, including members of the Steering Committee of CCAMTAC, shall require the explicit consent of the TA recipient and the IMF's Fiscal Affairs Department.

MEMBERS


Armenia


Azerbaijan


Georgia


Kazakhstan


Kyrgyz Republic


Mongolia


Tajikistan


Turkmenistan


Uzbekistan

DEVELOPMENT PARTNERS (DONORS)

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Economic Affairs SECO

Ministry of Economy and Finance

EUROPEAN UNION

ADB

Ministry of Foreign Affairs
Republic of Poland

Contents

Figures

Tables

Annexes

Abbreviations and Acronyms

AMD	Armenian Dram is the currency of Armenia
CCAMTAC	The Caucasus, Central Asia, and Mongolia Regional Capacity Development Center
CD	Capacity Development
CIT	Corporate Income Tax
CRM	Compliance Risk Management
FAD	IMF Fiscal Affairs Department
GDP	Gross Domestic Product
IMF	International Monetary Fund
NACE	Nomenclature statistique des Activites economiques dans la Communaute Europeenne (Statistical nomenclature of economic activities in the European Community)
PIT	Personal Income Tax
SRC	Armenia State Revenue Committee
STX	IMF Short Term Expert
TIN	Taxpayer Identification Number
VAT	Value Added Tax

Preface

In response to a request from the Armenia State Revenue Committee (SRC), a capacity development (CD) mission team comprising Mr. Soren Pedersen (FAD) and Mr. Tobias Gabel Christiansen (FAD short-term expert), carried out a duty-based station mission during the period June 30 – July 04, 2024. The purpose of this mission, financed by CCAMTAC, was to predict the Corporate Income Tax (CIT) gap based on operational audits for the income year 2023. This is a follow up on a mission in March 2024 that estimated the CIT gap for 2020-2022 based on operational audits.

Prior to the mission SRC had provided the required data consisting of CIT returns for income year 2023. The report's findings are based on the data provided by the SRC.

The IMF team expresses its sincere appreciation to SRC for the excellent cooperation and the excellent support provided before this mission. The team particularly acknowledges the excellent support provided by Mr. Arsen Sarikyan, Head of Development and Administration Strategy Programs Department; Mr. Mkhitar Ayvazyan, Deputy Head of Development and Administration Strategy Programs Department; Mr. Martin Sandoyan, Head of Revenue Assessment and Analysis Division; and Grigor Hakobyan, Chief Specialist.

This report represents the final version of the draft report that was submitted to Mr. Ashot H. Muradyan, Deputy Chairman and Mr. Arsen Sarikyan, Head of Development and Administration Strategy Programs Department, on July 10, 2024. It consists of an Executive Summary and the following sections: (I) Introduction (II) Key Findings; and (III) Next Steps.

Executive Summary

This mission, financed by CCAMTAC, predicted[1] the corporate income tax (CIT) gap in Armenia based on a bottom-up approach using data from the State Revenue Committee's (SRC) operational audits (2020-2022) and CIT returns for 2023.[2] The CIT gap was estimated in a previous mission in March 2024 based on completed audits for income years 2020-2022. The CIT gap is predicted to be 34.1 percent (or 25.5 percent)[3] of potential CIT liability for the income year 2023 based on non-audited CIT returns.

For the years 2020, 2021, 2022 the CIT gap is estimated at 38.8, 35.2, 32, respectively.[4] Excluding audit adjustments with no immediate impact on revenue (where the audit adjustment resulted in reducing a loss or carry forward of losses), the gap is 29.2, 26.4, 24.6 in 2020, 2021, and 2022.

The CIT gap is predicted to be 1.5 percent of GDP in 2023. For the years 2020, 2021, and 2022, the CIT gap was estimated to be 1.3, 1.4, and 1.5 percent of GDP, respectively. Excluding audit adjustments with no immediate impact on revenue the corresponding numbers are 0.9 percent for 2020, 0.9 percent for 2021, 1.0 percent for 2022, and predicted to be 1.0 percent for 2023.

In absolute terms, the CIT gap is predicted to be 121.2 billion AMD in 2023, while it was estimated to be 83.1 billion AMD in 2020, 90.3 billion AMD in 2021, 108.5 billion AMD in 2022. Excluding audit adjustments with no immediate impact on revenue, the CIT gap in 2023 is predicted to be 80.1 billion AMD, while it was estimated to be 54.0 billion AMD in 2020, 59.6 billion AMD in 2021, and 72.4 billion AMD in 2022.

The CIT gap in terms of potential CIT is trending downward when the largest corporations are excluded. However, when considering all corporations, the predicted CIT gap as mentioned above shows a slight increase in 2023. This suggests that the overall rise is predicted to be driven by the influence of large corporations.

As the next step, it is recommended to measure the CIT gap once audits for 2023 have been completed. When audits are completed, likely in 2025, it is recommended to estimate the CIT gap based on these completed audits and compare the results to the predicted CIT gap for 2023. The SRC can use the provided material (R code with detailed instructions) to conduct this analysis with their own data analysts or seek IMF capacity development (CD) for an assisted self-assessment measure of the CIT gap.

---

[1] The term "predicted" is used here to show that the CIT gap is based on CIT returns before they have been audited. The term "estimated" is used where the CIT gap has been estimated based on audited CIT returns.

[2] This analysis measures the CIT compliance gap, i.e., the policy gap is excluded. For simplicity, "gap" and "compliance gap" will be used indistinguishably in this report.

[3] Excluding audit adjustments with no immediate impact on revenue (due to tax losses before and after audit). The estimates do not account for undetected noncompliance which could lead to underestimation of the CIT gap.

[4] Potential CIT liability is defined as self-reported CIT plus the estimated CIT gap.

Recommendations

Estimate the CIT gap when audits are completed in 2025		Due data
1	It is recommended that the SRC carry out the CIT gap estimation once the audits are completed and compare the results to the predicted CIT gap for 2023. Data analysts in the SRC can perform this analysis using the provided R code and detailed instructions. Alternatively, the SRC can seek IMF capacity development (CD) for an assisted self-assessment measure of the CIT gap based on completed audits.	May 2025

I. Introduction

1. This mission, financed by the CCAMTAC, predicted the CIT gap based on a bottom-up approach using CIT returns for 2023 and a previously developed model to estimate the CIT gap from operational audits for 2020-2022[5]. In a mission in March 2024 a model was developed to measure the CIT gap for income years 2020-2022. This model has been used in this mission to predict the CIT gap for 2023 based on non-audited CIT returns for income year 2023. The prediction was specially requested by SRC to get a fast forecast of the CIT gap for 2023 based on the previous developed model.

2. The previously developed model used operational audits to measure the CIT gap. Using operational audits to estimate a tax gap needs to account for the non-random selection of audited corporations. Without correction for cases not selected randomly, the CIT gap would be overestimated. This bias, known as "Sample Selection Bias," arises from the non-random selection of cases. The method devised by James J. Heckman is utilized to correct for Sample Selection Bias.[6]

3. The CIT gap for 2020-2022 was thus estimated using the Heckman Sample Selection model. The Heckman method is a two-stage procedure. In the first stage, it estimates the probability that a company is selected for audit. This is done using a probit model[7] and 25 separate risk scores used by the SRC to target CIT audits. The second stage models the audit outcome (i.e., tax uncovered) using company characteristics (i.e., lines from the CIT return, sector, number of employees etc.) and a regressor that accounts for the selection process (see Annex I for more details).

4. The Heckman two-step estimator's suitability in estimating the CIT gap depends on audit practices. Narrowly focused audits may miss undisclosed taxes, leading to an underestimated CIT gap. When audits target specific sectors or types of firms, reliable estimates become challenging due to increased extrapolation. Additionally, some noncompliance may remain undetected, potentially underestimating the CIT gap.

5. CIT gap estimates based on operational audits were obtained for three consecutive years. These estimates cover the latest available audit data for the income years 2020, 2021, and 2022. Due to the limited annual number of audits, the yearly estimates are more uncertain.

6. The previous developed model based on audit cases was used to predict the CIT gap for income year 2023. The prediction for 2023 was based on non-audited CIT declarations. The deadline for corporations to declare their income tax for 2023 was April 20, 2024.

7. The predicted CIT gap should be close to the actual estimate of the CIT gap but only under certain conditions. The audit selection strategy should remain unchanged. Additionally, there should be no changes to tax regulations or new, unseen methods of evasion that significantly affect compliance

[5] "ARMENIA. CIT Gap Estimation Based on Operational Audits. MAY 2024". Soren Pedersen and Tobias Gabel Christiansen. IMF report.

[6] James J. Heckman (1979). "Sample Selection Bias as a Specification Error". Econometrica. vol. 47(1), pp. 153-161.

[7] In a probit model the outcome is binary (0 or 1) – in this case whether a company has been audited or not.

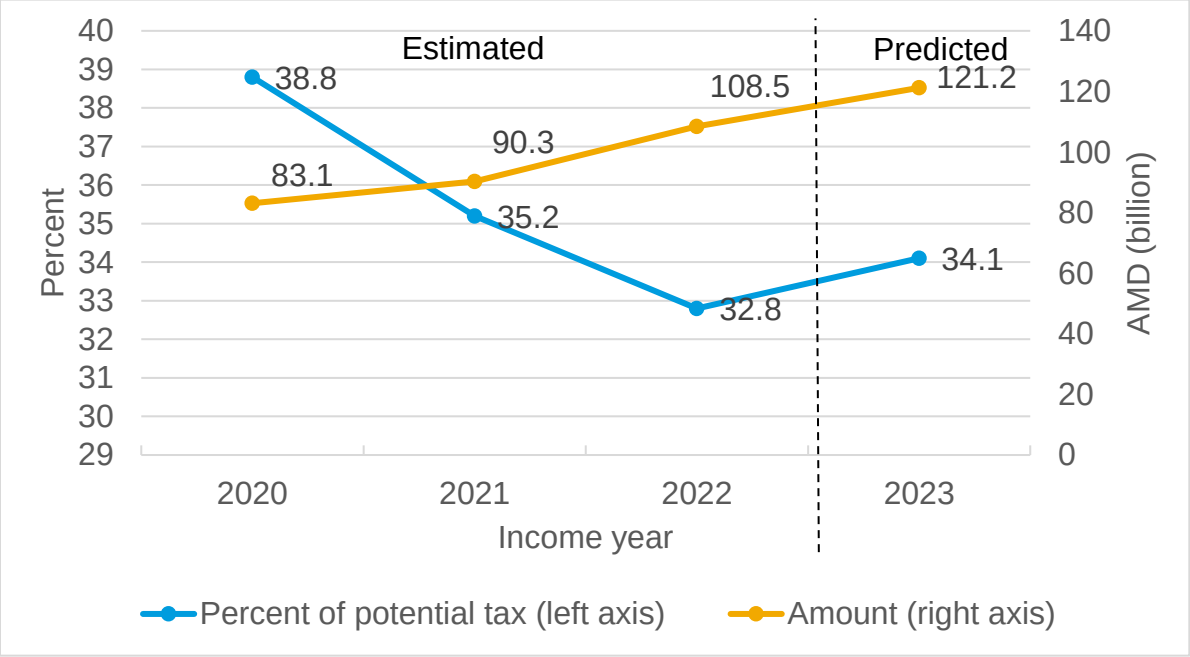
compared to previous years. An actual estimate of the CIT gap can be obtained once the 2023 audits are finished.

II. Key Findings

A. CIT gap results

8. The CIT gap in 2023 is predicted to be 121.2 billion AMD[8] or 34.1 percent of potential CIT liability (Figure 1). The predicted CIT gap corresponds to 1.5 percent of GDP in 2023. The CIT gap prediction is based on a Heckman model using operational audits conducted for the income years 2020, 2021, and 2022. A total of 4,432 comprehensive audits, covering all aspects of the company and carried out by the SRC over the three-year period, were used to estimate the Heckman model. The results have been adjusted to account for the non-random selection of operational audits, addressing what is known as "sample selection bias."

Figure 1. Estimated and predicted CIT gap



Source: IMF calculations based on data from SRC.

9. The predicted median CIT gap is highest in the sector with "Extraterritorial" (Figure 2).[9] However, it should be noted that this sector is small, so its contribution to the total CIT gap is limited. As seen in Figure 2, the second-highest median CIT gap is in the mining sector. The lowest median CIT gap among the top-10 sectors is found in the accommodation sector.

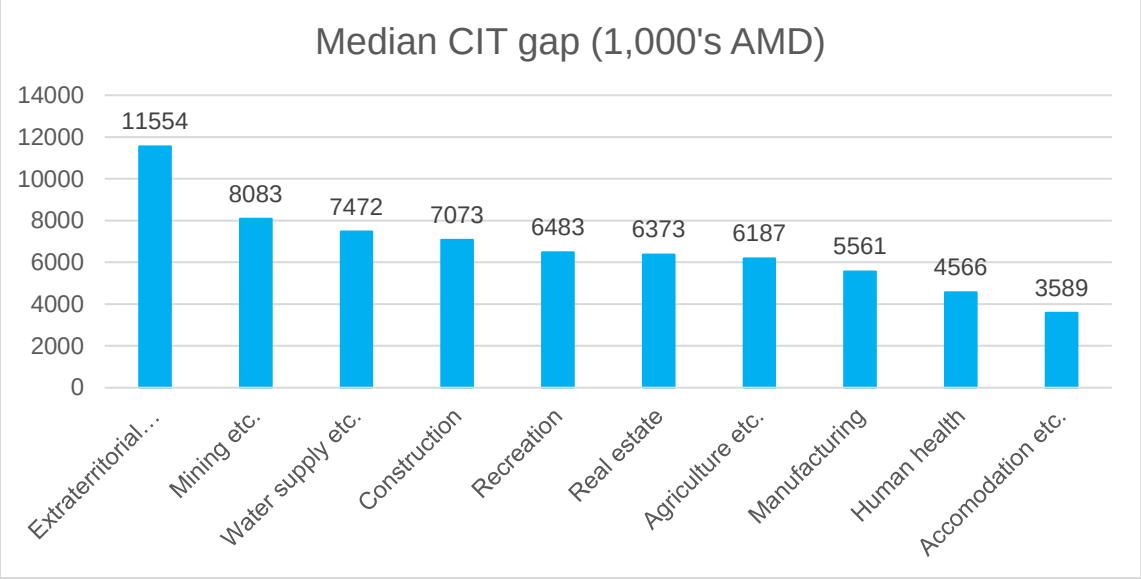
10. It is important to examine the reasons for non-compliance more closely in high-risk sectors such as the mining sector. It will give valuable information for a compliance improvement plan

[8] Measured in 2020-prices.

[9] Sector classification follows NACE coding (Nomenclature statistique des Activites Economiques'-Statistical classification of economic activities in the European Community).

to understand if non-compliance is due to deliberate evasion, lack of knowledge of the tax law or complicated legislation.

Figure 2. Predicted median CIT gap by top-10 sectors[10]



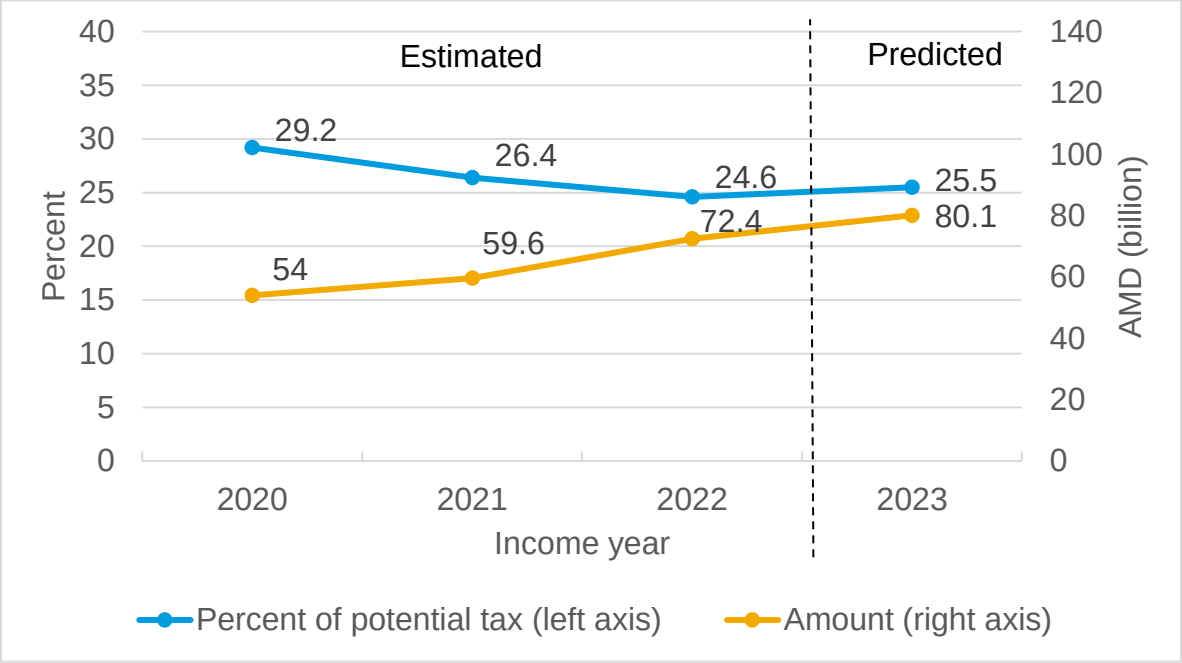
Source: IMF calculations based on data from SRC.

11. If a Heckman model that excludes adjustments with no immediate effect on revenue is used, the CIT gap in 2023 is predicted to be 80.1 billion AMD or 25.5 percent of potential CIT (Figure 3). The predicted CIT gap corresponds to 1.0 percent of GDP in 2023. The immediate tax liability of companies with a net loss before and after an audit remains unchanged, and whether to include it in the tax gap or not is subject to debate. On one hand, the audit adjustment has no immediate effect on revenues. On the other hand, if the losses of a company are reduced following an audit, the losses that the company can carry forward to future years are reduced, leading to higher future tax revenues (except in cases where the company goes bankrupt, is liquidated, or continues to evade taxes etc.).[11]

[10] This sector is considered the primary sector for a corporation if they operate in more than one sector. Sector codes are self-reported by corporations. The chart only displays the top 10 sectors due to space limitations.

[11] In Denmark, for example, all corrections are included in the estimation of the tax gap, including those with no immediate impact on revenue (source: https://www.ft.dk/samling/20171/almdel/SAU/bilag/92/1839723.pdf, available in Danish only). On the other hand, the Inland Revenue Service (IRS) in the United States (US) does not include corrections in their tax gap estimation it the audit adjustment results in no payment – see page 29 in Publication 5784 (10-2022) (irs.gov).

Figure 3. CIT gap: Excluding adjustments with no immediate effect on revenue



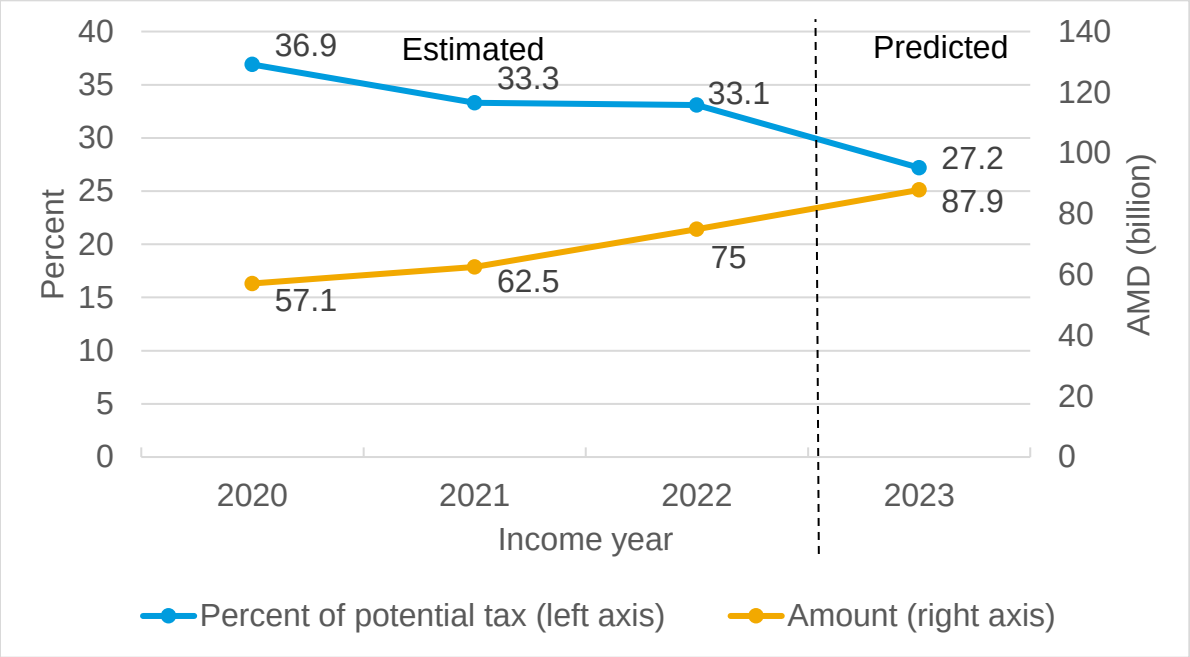
Source: IMF calculations based on data from SRC.

12. Using a Heckman model that excludes the largest corporations, the CIT gap in 2023 is predicted to be 87.9 billion AMD and 27.2 percent of potential CIT (Figure 4). The predicted CIT gap corresponds to 1.1 percent of GDP in 2023.

13. Excluding the largest corporations likely provides a more accurate estimate of the CIT gap. Due to the limited number of large companies, the associated tax gap entails more uncertainty. Excluding these companies can reduce 'noise' and provide more accurate tax gap estimate but comes at the cost of reducing the population for which the tax gap is estimated.[12] To balance the influence of outliers and still account for most of the population, the top 0.1 percent largest companies, measured by turnover, were excluded before estimating the Heckman model and making predictions.

[12] The IRS CIT gap methodology excludes Large Businesses from the scope of the Heckman methodology.

Figure 4. CIT gap: Excluding large corporations

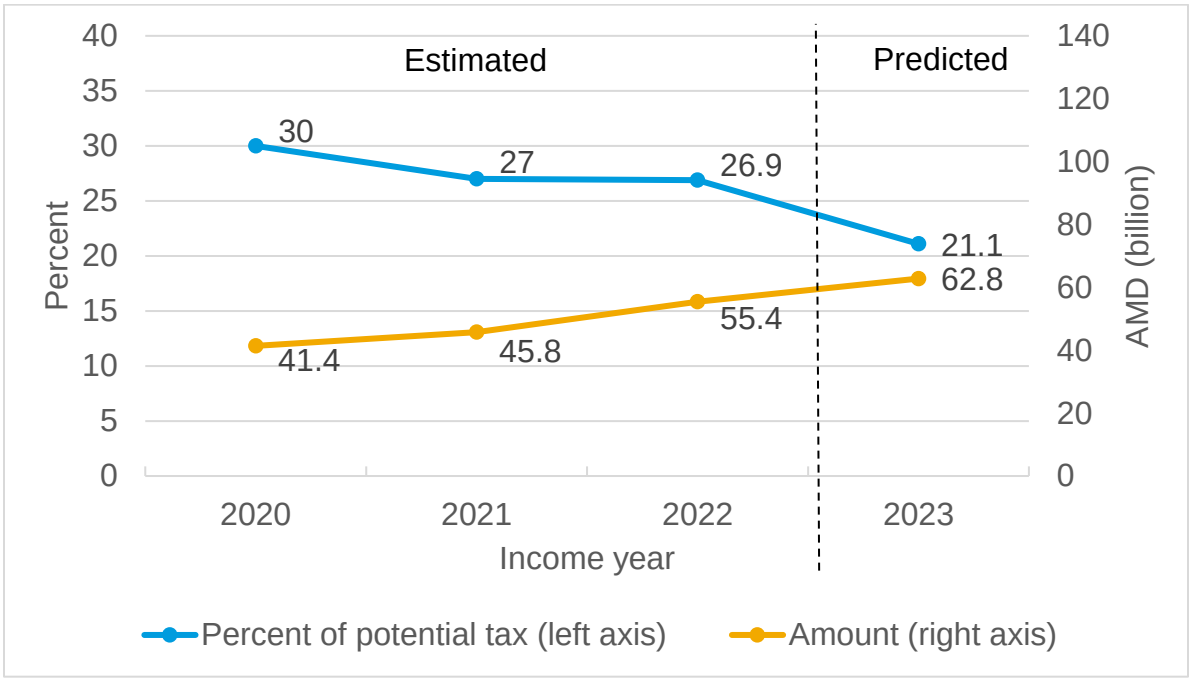


Source: IMF calculations based on data from SRC.

14. Using a Heckman model that excludes the largest corporations and adjustments with no immediate effect on revenue, the CIT gap in 2023 is predicted to be 62.8 billion AMD or 21.1 percent of potential CIT (Figure 5). The predicted CIT gap corresponds to 0.8 percent of GDP in 2023.

15. The CIT gap in terms of potential CIT is trending downward when the largest corporations are excluded (Figures 4 and 5). However, when considering all corporations, the predicted CIT gap shows a slight increase in 2023 (Figures 1 and 3). This suggests that the rise is predicted to be driven by the influence of large corporations, although the large differences in scale and size among large corporations introduce noise, which could be driving the observed pattern.

Figure 5. CIT gap: Excluding large corporations and adjustments with no immediate effect on revenue



Source: IMF calculations based on data from SRC.

III. Next Steps

A. Set up a team of data analysts and apply the developed model

16. Based on the previous mission it was recommended that SRC appoints 2-3 data analysts with responsibility to compile future CIT Gap estimations. To sustain the CIT gap model, it is crucial that SRC invests in data analytics. It was recommended that at least 2-3 individuals be trained in data analytics. This will enable SRC to independently conduct CIT Gap analysis based on operational audits developed by this mission.

17. It was also recommended that the appointed team should invest around 80 percent of their time on data analytics. It is important to invest sufficient time to be able to perform good data analytics. The team should learn the necessary programming tools to be able to carry out CIT gap analysis.

B. Recommendations

18. SRC should carry out the CIT Gap estimation after income year 2023 audits are completed in 2025. The predicted CIT gap for 2023 determined in this mission is based on CIT declarations before audit. When audits are completed in 2025 the CIT Gap should be estimated and compared to the predicted CIT Gap for 2023 presented here.

19. If SRC is not able to carry out the estimation in 2025 itself, they could seek IMF assistance for an assisted self-assessment of the CIT Gap. This approach will help the SRC to develop the required data analytical skills to use the model in the future and become self-sufficient.

Annex I. Supplementary material.

Tax gap estimates from non-random risk-based audits are prone to sample selection bias due to the selection process being influenced by the perceived risk of non-compliance. Put differently, the audited companies are not representative of the general population of companies since they are selected as being more prone to risk of tax non-compliance based on several indicators. Hence the tax gap estimate based purely on such operational risk-based audits does not reflect that of the general population. A common approach to account for this is through the Heckman 2-step estimator.[13] The Heckman 2-step estimator corrects for sample selection bias by estimating both the selection process and the level of non-compliance (i.e., the tax uncovered from audit) in the same model. Following Wooldridge (2010)[14] the Heckman 2-step estimator is given by an outcome equation and a selection equation:

$$Y_i = X_i\beta + u_i \qquad (1)$$

$$S_i = 1[Z_i\delta + v_i > 0] \qquad (2)$$

Here equation (1) is the outcome equation, where Y_i measures the tax uncovered from audit and X_i is company characteristics (i.e., lines from the CIT return, sector, number of employees etc.)[15]. Next, equation (2) is the selection equation, where S_i is an indicator of audit, with $S_i = 1$ denoting company i was audited and $S_i = 0$ denoting company i was not, while Z_i are factors that determine whether a company is audited or not. This includes 25 separate risk scores used by the SRC to target CIT audits.[16] Importantly, the value of Y_i is only observed if company i was selected for an audit ($S_i = 1$). Finally, u_i and v_i are independent of X_i and Z_i, with $v_i \sim N(0,1)$ and $E(u_i|v_i) = \gamma v_i$.[17]

What we are interested in estimating is $E[Y_i|X_i]$. However, since Y_i is observed only when $S_i = 1$, what we can estimate is $E[Y_i|X_i, S_i = 1]$. Using equation (1) and (2) this can be rewritten as:

$$E[Y_i|X_i, S_i = 1] = X_i\beta + E[u_i|v_i > -Z_i\delta] = X_i\beta + \gamma\lambda(Z_i\delta) \qquad (3)$$

Here $\lambda(\cdot) = \phi(\cdot)/\Phi(\cdot)$ where $\phi(\cdot)$ and $\Phi(\cdot)$ are the probability density function (pdf) and cumulative distribution function (cdf) of a standard normal distribution, respectively. The form of $\lambda(\cdot)$ follows from the assumption that $v_i \sim N(0,1)$ and it is labeled the inverse Mills ratio. Equation (3) presents a way to

---

[13] James J. Heckman (1979). Sample Selection Bias as a Specification Error. *Econometrica*. vol. 47(1), pp. 153-161.

[14] Jefrey M. Wooldridge (2010). Econometric Analysis of Cross Section and Panel Data. The MIT Press

[15] A total of 25 variables are used. These variables correspond to the 25 variables with the highest variable importance in the machine learning model developed to predict audit adjustments. The top 25 were chosen because adding more variables led to limited increases in the explained variation of the audit adjustments and the risk of multicollinearity and loss of degrees of freedom.

[16] The 25 factors that determine each of the risk scores are confidential. Hence, the estimation of equation (2) was based on the assigned scores and not the underlying factors. Line item A?91 in section 1 of the corporation tax return "Total amount of profit tax deducted due to the privilege of deduction of profit tax" was not one of the 25 variables. This line item was included in a separate analysis after request from SRC but had not impact on the results due to the very limited number of corporations that filed an amount in line item A91. Similarly, after a request from SRC, an indicator defining whether a company is included in the group of resident taxpayers of the Republic of Armenia carrying out a program, and thereby subject to the privilege of lower tax rates, was included in a separate analysis. The inclusion also did not affect the results.

[17] We only require v_i to be normally distributed. It is sufficient to assume that the conditional expectation of u_i given v_i is linear, which does not require u_i to be normally distributed.

consistently estimate β. Following Heckman (1979) we can consistently estimate β and γ by regressing Y_i on X_i and $\lambda(Z_i\hat{\delta})$ using OLS, where $\hat{\delta}$ is obtained by estimating equation (2) using a probit model. Once an estimate of β has been obtained using the 2-step Heckman estimator, it can be used to construct an estimator of the unconditional expectation of non-compliance (not conditioning on $S_i = 1$), given by $E[\widehat{Y_i|X_i}] = X_i\hat{\beta}$.[18] This can be applied to obtain predicted values of non-compliance for all companies in the population, and thereby the overall CIT gap. Table A1 presents estimates of the selection and outcome models using all data on CIT returns and operational audits from 2020, 2021 and 2022. The selection model obtains an R^2 of 0.26.[19] This should be viewed in the context of the current audit strategy. When a company gets a high-risk score, it undergoes an audit. However, it's not just the income year when the company got the high score that gets audited; all previous non-audited CIT returns are also audited. Consequently, a company might have an income year audited even if it received a low score during that period, making risk-scores non-perfect predictors of audits at the TIN/income year level. Turning to the outcome model, it obtains an R^2 of 0.14. Ideally, we aim for this value to be as high as possible. However, due to the considerable diversity among companies and audit adjustments, reaching this goal is challenging. Interestingly, the coefficient on the inverse Mills-ratio is significant (IMR in Table A2), indicating the presence of sample selection bias.

Two important points need to be highlighted. First, when using the Heckman 2-step estimator to predict non-compliance instead of inferring causal relationships, the accuracy of predictions depends on how well the selection model and the outcome model explain the data. Second, it is best to avoid using the same variables in both models. Doing so makes the outcome model's identification rely on the non-linearity of the inverse Mills-ratio, which can cause unstable results due to high multicollinearity. To prevent this, the selection model should include at least one variable that determines whether a company gets audited but doesn't affect non-compliance levels (known as an exclusion restriction).[20] However, finding such a variable can be tricky if audits are solely based on estimated non-compliance. In this context, we use 25 risk scores to predict whether a company undergoes an audit. Table 3 in the main text shows that the relationship between the SRC's risk scores and average audit adjustments is non-monotone. This implies that some risk scores may have little to no connection to non-compliance levels, which satisfies the exclusion restriction.[21] However, if the SRC updates their targeting strategy and implements a MLM designed to identify large audit adjustments based on a large array of variables associated with the company, identifying exclusion restrictions may become more challenging.

[18] Standard errors are wrong when manually estimating the 2-step Heckman estimator. Correct standard errors can be obtained using bootstrap.

[19] This is McFadden's Pseudo R^2

[20] In other words, exclusion restriction means that there must be at least one variable appearing with a non-zero coefficient in the selection equation but not in the equation of interest.

[21] Regressing the audit adjustment on the risk scores reveals that 16 out of the 25 risk scores are not statistically significantly related to the tax uncovered from audit.

Table A1: Regression results

Selection Model			Outcome model		
Variables	**Coefficients**	**Std. Error**	**Variables**	**Coefficients**	**Std. Error**
Intercept	-1,115***	0,030	Intercept	6153,822**	3109,093
Risk Score 1	-0,005***	0,001	Tax line 85	-0,006	0,021
Risk Score 2	0,022***	0,001	Tax line 76	0,001	0,005
Risk Score 3	-0,006***	0,001	Tax line 50.5	-0,066	0,048
Risk Score 4	-0,001*	0,000	Turnover pr. emp.	0,003	0,002
Risk Score 5	-0,003***	0,000	Tax line 84	0,010	0,022
Risk Score 6	-0,006***	0,001	Number of emp.	11,705**	4,015
Risk Score 7	-0,001**	0,001	Tax line 40	-0,001	0,001
Risk Score 8	0,007***	0,001	Tax line 83	-0,005	0,006
Risk Score 9	-0,007***	0,001	Tax line 19	0,183	0,127
Risk Score 10	0,005***	0,002	Tax line 7	0,001	0,001
Risk Score 11	-0,001***	0,000	Tax line 62	-0,005	0,006
Risk Score 12	0,002***	0,000	Tax line 50	-0,001	0,001
Risk Score 13	0,002	0,002	Tax line 50.4	-0,150**	0,061
Risk Score 14	0,000	0,000	Total compensation	-0,031*	0,016
Risk Score 15	-0,004***	0,001	Tax line 11	0,000	0,001
Risk Score 16	-0,014***	0,001	Tax line 59	0,005	0,004
Risk Score 17	-0,002	0,002	Tax line 59	0,006	0,012
Risk Score 18	-0,010***	0,001	Tax line 51	0,000	0,002
Risk Score 19	0,005***	0,001	Tax line 58	-0,248	0,283
Risk Score 20	0,005***	0,001	Tax line 16	0,007	0,007
Risk Score 21	0,009***	0,002	Tax line 41	0,005	0,021
Risk Score 22	-0,010***	0,001	Tax line 43	0,001	0,001
Risk Score 23	0,041***	0,003	Tax line 9	0,000	0,001
Risk Score 24	-0,026	0,058	Tax line 45	0,000	0,001
Risk Score 25	-0,021***	0,002	IMR	-1032,997*	578,905
Includes Sector dummies	No			Yes	
Number of observations	71,374			4,324	
R^2	0.26			0.14	

Source: IMF calculations based on data from SRC. Note: [1)] In 2020-prices. 0.1 pct. of audit results were trimmed in top and bottom for each year. Re-audits are discarded. Only corporations with a reported CIT return. The outcome model also includes sector dummies. Standard errors are computed using bootstrap, with resampling done at the TIN-level. For the selection model R^2 corresponds to McFadden's Pseudo R^2. *, **, *** denotes p < 0.01, p < 0.05, p<0.10.

A: Supplementary results

20. At the request of the SRC, Table A2 shows the predicted values in 2023 prices (non-deflated). The Heckman model used to make the predictions utilized CIT returns and audits from multiple years (2020, 2021, and 2022). For this reason, prices were deflated to 2020-prices. To ensure coherence, the CIT returns for 2023 were also deflated to 2020 prices before making predictions in the main results. The SRC requested the CIT gap prediction in 2023 prices (non-deflated).

Table A2. Predicted CIT GAP in 2023 (non-deflated).

	Predicted CIT gap in 2023			
	All	**No loss**	**No large**	**No large / No loss**
CIT gap	143.9 billion AMD[1)]	95.1 billion AMD[1)]	104.3 billion AMD[1)]	74.5 billion AMD[1)]
Percent of potential CIT	34.1	25.5	27.2	21.1
Percent of GDP	1.5	1.0	1.1	0.8

Source: IMF calculations based on data from SRC. Note: 1) In 2023-prices. Predictions made using a Heckman model estimated on audits from 2020-2022. 0.1 pct. of audit adjustments were trimmed in top and bottom for each year. Re-audits are discarded. "All" refers to the Heckman model that used all remaining audits. "No loss" excluded audit adjustments with no immediate impact on revenue. "No large" excluded the top 0.1 percent of the largest companies, as measured by turnover. The predictions only corporations with a reported CIT return. Potential CIT liability is defined as self-reported CIT plus the estimated CIT gap.

B: Materials Left with Armenia State Revenue Committee

- R-code used to clean 2023-CIT return data.
- R-code used to make CIT gap predictions for 2023.
- Powerpoint presentation with key findings.
- CSV-file with tax gap predictions using the model based on all audits.
- CSV-file with tax gap predictions using the model that excludes audit adjustments with no immediate impact on revenue.