Tomasz Tratkiewicz [iD] [✉]
Institute of Finance, University of Lodz, Poland

Mehmet Burak Turgut [iD] [✉]
Department of Macroeconomics and Foreign Trade Theory, University of Warsaw, Poland

Estimation of the Personal Income Tax (PIT) Gap in Poland[*]

Oszacowanie luki w podatku dochodowym od osób fizycznych (PIT) w Polsce

Abstract

The literature on estimating the tax gap in direct taxes generally emphasises that using a bottom-up method produces much better results than using a top-down method. The existing methods of estimating the PIT gap using the bottom-up approach rarely take advantage of risk-based audits, which can be successfully used when tax gap estimation based on random audits cannot be applied. An important issue with risk-based audits is the potential selection bias. To address this issue, we use the Heckman selection model. We estimate the size of the PIT gap in Poland in 2017 by using data from the tax filings of taxpayers earning income from business activities taxed at a flat 19-percent tax rate and risk-based audits. Our findings show that the predicted PIT gap corresponds to 4.5 percent of the total tax liability of this group of taxpayers. We also identify the regions and sectors with the highest predicted PIT gap per taxpayer. Our results show that the Heckman model can be used as a tool by tax authorities to improve the efficiency of risk-based audits.

Streszczenie

W literaturze dotyczącej szacowania luki w podatkach bezpośrednich podkreśla się, że oddolna metoda szacunku (ang. *bottom-up approach*) daje znacznie lepsze wyniki niż metoda odgórna (ang. *top-down approach*). Istniejące metody szacowania luki w PIT przy użyciu metody oddolnej stosunkowo rzadko wykorzystują kontrole bazujące na analizie ryzyka, które administracja podatkowa stosuje, aby zapewnić w szczególności przestrzeganie obowiązków podatkowych. Ważną kwestią, która – w przeciwieństwie do kontroli losowych – występuje w kontrolach bazujących na analizie ryzyka jest potencjalne obciążenie próby. Aby rozwiązać tę kwestię, wykorzystujemy model selekcji Heckmana. Oszacowaliśmy wielkość luki w PIT w Polsce w 2017 r. na podstawie danych z zeznań podatkowych podatników uzyskujących dochody z działalności gospodarczej opodatkowanej liniową 19-procentową stawką podatku oraz wyników kontroli bazujących na analizie ryzyka.

Nasze wyniki pokazują, że przewidywana luka w PIT odpowiada 4,5% całkowitego zobowiązania podatkowego tej grupy podatników. Zidentyfikowaliśmy również regiony i sektory o najwyższej przewidywanej luce w PIT na podatnika. Stwierdzamy, że model Heckmana może być wykorzystywany przez organy podatkowe jako narzędzie poprawiające skuteczność kontroli bazujących na analizie ryzyka.

Introduction

The personal income tax (PIT) gap is the difference between the amount of PIT due the fiscal authority and the amount of PIT collected [**Feige, 1990**].[1] The PIT gap would be zero if all taxpayers were fully compliant with their tax obligations. Hence, the gap is the loss of PIT revenue owing to non-compliance that could cost governments a great deal of revenue.[2] Moreover, the gap influences the distribution of income and wealth and tax morale among taxpayers. Therefore, it is crucial to estimate the PIT gap and assess the extent of non-compliance (the tax gap), its components, and the underlying causes in order to improve tax policy and enhance the tax administration's collection efficiency.

Countries apply a variety of methods to measure their PIT gap [**IMF, 2021**]. These methods can be divided into two main categories: top-down and bottom-up. In top-down methods, the data used to estimate the tax gap usually comes from a macroeconomic model or national accounts and covers the entire tax base, whereas the bottom-up methods use one or more data sources at the individual taxpayer level such as audits, surveys or questionnaires, and extrapolate them to a larger population [**FISCALIS, 2017**].

In this paper, we follow the bottom-up approach and estimate the PIT gap in Poland for a group of taxpayers earning income from business activities taxed at a flat rate of 19% using data from the PIT declarations and risk-based audits.[3] The content of these two data sources is different. The PIT declarations are submitted by each taxpayer; hence, they represent the whole population of taxpayers who have declared their revenues to the tax administration. However, risk-based audits only represent a part of the population since they are conducted on a sample of taxpayers selected after a risk assessment performed by the tax authorities. Therefore, a sample from risk-based audits cannot be considered representative of the whole population due to selection bias because the selection of taxpayers for audits is based on risk assessment, as opposed to random controls. Moreover, the results cannot be directly extrapolated to taxpayers who have not been selected for audit to obtain an overall estimate.

Nonetheless, data from risk-based audits can be extrapolated to the taxpayer population if the problem of selection bias is addressed. We use the **Heckman** [**1979**] sample selection model (hereafter, Heckman model) to correct for the selection bias inherent in risk-based audits. To the best of our knowledge, this study is the first to apply the Heckman model to PIT gap estimation, representing its primary contribution to the literature.[4] In addition to estimating the total PIT gap, we also estimate its size by region (voivodeship) and sectoral activity using the Polish Classification of Activities (PKD). In our study, we only estimate the overall PIT gap by region and sectoral activity; we do not decompose it by type of non-compliance. Such a decomposition would require detailed audit data disaggregated by specific forms of non-compliance. We did not have

---

[1]   The literature and practical estimates detail the concept of the gross tax gap as the difference between the tax liability in a given year and the amount of tax that is paid voluntarily and on time [**Toder, 2007**].

[2]   The tax gap for Income Tax, National Insurance contributions, and Capital Gains Tax in the UK was £12.7 billion from 2020 to 2021. This was the largest share of the total tax gap when viewed by type of tax (39.5%). Of this amount, £7.4 billion was attributable to Income Tax Self Assessment taxpayers [**HMRC, 2022**].

[3]   We will further justify why we have selected the self-employed for the PIT gap study, and out of the entire group of self-employed, those paying PIT at a flat rate of 19%.

[4]   The tax administrations of Belgium, Denmark and Italy used or intended to use the Heckman model in their attempts to estimate the gap with risk-based audit data [**FISCALIS, 2018**]. However, we do not know the results of these attempts and whether they concerned the estimation of the PIT gap. Referring to income taxes in general, the so-called Censored Heckman approach was applied to estimate the corporate income tax (CIT) gap in Slovakia [**Chudý et al., 2020**].

access to such data and, to our knowledge, it remains difficult–if not impossible–to identify specific types of non-compliance from available audit reports.

The data for the estimation comes from two sources. The first source is PIT declarations in Poland for 2017, containing various financial, geographical, and tax- related information about taxpayers.[5] However, PIT declarations are not uniform in the Polish tax system and depend on the nature of the income (revenue) source. There are four main types of PIT declarations in Poland: (i) for self-employed individuals who can pay a flat rate of 19% out of business income, (ii) for self-employed taxpayers who can pay one of the 3%, 8.5%, 10%, 17%, and 20% rates[6] out of self-employment revenue, (iii) for taxpayers with the majority of income from self-employment or rent, and (iv) for taxpayers with the majority of income from employment or pensions.[7] This is a very simplified picture of the system under which PIT is paid in Poland because some taxpayers are required to file PIT declarations under two different schemes and taxpayers paying PIT under the general rules can file jointly with their spouses.

An important caveat related to the PIT declarations is that the variables providing information about the taxpayers are different across different PIT types; hence, it is not possible to pool different types of PIT declarations for the estimation. Therefore, one needs to estimate the PIT gap separately for taxpayers across different PIT types. Self-employed individuals can pay PIT using three forms: the general rules (progressive taxation with two rates and a tax-free amount), a flat rate of 19% or a lump sum on registered revenue (taxpayers cannot deduct costs, seven rates applied in 2017). Among taxpayers paying PIT under the general rules and independently submitting annual tax returns (i.e. without employer withholding), numbering nearly 20 million individuals, approximately 10% reported income from business activity. Among those taxed at the flat 19% rate (more than 572,000 taxpayers), almost all reported income from business activity, including non-agricultural business activity and special divisions of agricultural production. Of the more than 1.1 million taxpayers paying PIT under the lump-sum tax on registered revenue, approximately 50% earned income from business activity. In this study, we use data from PIT declarations filed by taxpayers earning income from business activity taxed at the flat 19% rate and estimate the PIT gap only for these taxpayers, labelled as *type I*. This group was selected as the primary focus of the analysis because the study centres on the self-employed, among whom the risk of tax non-compliance is generally the highest. This focus is motivated by previous research demonstrating that non-compliance is most prevalent among taxpayers whose income is not subject to third-party information reporting or withholding requirements [**Gemmell, Hasseldine, 2012**]. Evidence of the difficulty in detecting non-compliance in this area is provided by research from the United States, which shows that for every dollar of underreported income detected by Taxpayer Compliance Measurement Program examiners without the assistance of third-party information, an additional USD 2.28 remained undetected [**Feldman, Slemrod, 2007**]. These findings provide strong justification for concentrating on the self-employed.

The self-employed are best represented within the flat 19% tax regime, which in principle applies exclusively to this group. The choice of taxpayers taxed under this scheme is further justified by audit data: 8.3% of taxpayers paying PIT at the flat 19% rate in 2017 were subject to audit activities, compared with only 1% of taxpayers taxed under the progressive tax scale. Moreover, only individual taxation is permitted under the flat-rate regime, as joint taxation is not allowed, eliminating potential differences in compliance behaviour arising from joint filing [**Andreoni et al., 1998**].

Taxpayers taxed at the flat 19% rate also report the highest tax bases among the self-employed. On average, they reported nearly 38 times more revenue from business activity than taxpayers taxed under the lump-sum regime, and just over 15 times more than those earning business income under the general rules. Revenue is used for comparison because taxpayers under the lump-sum regime are not permitted to deduct costs.

---

[5]   A more detailed information about fiscal data is given in Section 3.
[6]   The tax rates in force in 2017.
[7]   The latter two types of declaration may also include income from sources other than those indicated.

The substantially higher tax bases among flat-rate taxpayers likely reflect the fact that, regardless of income level, the marginal tax rate does not exceed 19%, whereas under the progressive tax scale income above PLN 85,528 is taxed at 32%. As a result, owners of profitable one-person businesses are more likely to choose flat-rate taxation over the general rules. Although lump-sum taxation on registered revenue offers relatively low statutory rates depending on the type of activity, it does not allow cost deductions. In addition, in 2017 this form of taxation was subject to revenue thresholds (not exceeding EUR 250,000) and excluded certain types of activity, such as pharmacy operations and the trade of motor vehicle parts and accessories. Consequently, developing a model to detect non-compliance among type I taxpayers is expected not only to improve the effectiveness of audits for this group, but also to enhance overall compliance and ultimately increase PIT revenues.

The second data source consists of risk-based audits providing quantitative information on the tax compliance of selected taxpayers in Poland in 2017. These audits include checking activities, tax and customs controls, and tax inspections carried out by tax authorities,[8] i.e. activities using risk analysis rather than random sampling.[9] The choice of 2017 was deliberate: it was the first full year for which data from checking activities were centrally available, and both the quality and volume of these activities were sufficiently robust for empirical analysis. At the time of the study, no more recent data were available for use in the model.

The audits capture various forms of behaviour leading to tax non-compliance, including, but not limited to, ordinary errors, failures to exercise reasonable care, non-payment of tax, tax evasion, tax avoidance, and income concealment. All of these behaviours contribute to the tax gap, albeit to differing extents. However, owing to the limited level of detail in the audit reports, it is not possible to estimate the size of individual components of the gap using this approach; instead, we estimate the overall PIT gap. The literature identifies a wide range of determinants of tax evasion, which can broadly be grouped into economic factors (such as income level and income sources as well as the marginal tax rate), demographic factors (including age, gender and education), and behavioural factors (such as tax fairness, tax system complexity, and tax morale [**Wu, Teng, 2005**; **Christie, Holzner, 2006**; **Richardson, 2006**]. Due to the large number of determinants that can potentially affect tax non-compliance, no unified theory exists that specifies a definitive set of variables driving the PIT gap. Accordingly, we follow a hybrid approach that combines theory with statistical modelling to identify variables explaining tax non-compliance and the PIT gap within the Heckman model framework. Applying the Heckman model to the combined dataset reveals the presence of sample selection bias, indicating that using OLS would produce biased estimates.

We apply the estimated model to predict the PIT gap for all type I taxpayers since the Heckman model allows us to extrapolate the results for the whole population. The findings show that the total PIT gap in Poland for the fiscal year 2017 for type I taxpayers is predicted to be roughly PLN 1.2 billion, approximately 4.5% of the total tax liability of this group of taxpayers. We conduct various sensitivity analyses which confirm the robustness of our predictions. Among our results, we find that the declaration of loss in PIT returns has an impact on the PIT gap but also on the probability of selection into the audit.

The paper is organised as follows. The next chapter reviews the literature. Chapter 3 explains the methodology, and Chapter 4 summarises the data. Chapter 5 describes the empirical methodology and discusses the results. Chapter 6 concludes.

---

[8]   In this paper, we use the term "audit" in a broad sense, encompassing checking activities, tax controls, tax and customs controls, and fiscal inspections, the last of which are no longer conducted. The details of the different types of controls and their description can be provided on request (text volume restrictions).

[9]   Poland's National Revenue Administration, unlike its counterparts in some other countries, does not examine the size of the tax gap using random sampling. Such analysis is perceived as costly and sometimes has a negative impact on taxpayers' compliance with tax obligations [**Gemmell, Hasseldine, 2012**].

Literature review

The literature on tax gap estimation in direct taxes generally emphasises that the use of the top-down estimation method is subject to high uncertainty. The uncertainty of top-down estimates is such a serious problem that it can render top-down estimates meaningless [**Rubin, 2011**].

In estimating the PIT gap, more reliable results are generally achieved using bottom-up methods based on random audits. This type of audit for estimating the PIT gap seems to be recommended by the **IMF** [**2021**] and has been successfully used for many years by the tax administrations of the United States, Britain and Denmark, among others. For example, the US Internal Revenue Service's Taxpayer Compliance Measurement Program (TCMP) consists of tax audits conducted on a stratified random sample of personal income tax returns between 1963 to 1988. TCMP data shows that tax evasion is higher for income categories without third-party reporting [**IRS, 1996**; **Bloomquist, 2003**; **Slemrod, 2007**]. **Kleven et al.** [**2011**] carried out a field experiment in cooperation with the Danish tax collection agency (SKAT) by imposing different audit regimes on randomly selected taxpayers. Their findings reinforce the results from the TCMP, in which the tax evasion rate is close to zero for third-party income while being substantial for self-reported income. They also found that past audits and audit threats decrease tax evasion for self-reported income. In identifying the benefits of random audits, attention should primarily be paid to the production of scientifically relevant data that can be used in the process of improving risk analysis and better allocating tax administration resources (for more on the benefits and drawbacks of random audits see: **OECD** [**2004**]). Using a randomised control programme, the impact of controls on long-term compliance behaviour was recently studied by **Advani et al.** [**2021**]. However, one must also keep in mind the costs and effects associated with random audits [**Gemmell, Hasseldine, 2012**]. Therefore, alternative ways of estimating and decomposing the PIT gap are being studied, such as the use of risk-based audit data.

Risk-based audits focus on taxpayers selected with specific risk analysis criteria. Although the cost of risk-based audits is much lower than that of random audits, there is a selection bias in risk-based audits since the available sample does not represent the whole population [**FISCALIS, 2018**]. The literature uses post-stratification, Heckman selection [**Heckman, 1979**], and diff-in-diff approaches to address the selection bias. **Chudý et al.** [**2020**] used Gaussian and Censored Heckman models to estimate the corporate income tax (CIT) gap with firm-level data in Slovakia. They predicted that the CIT gap amounts to approximately 35% of total tax liabilities (both declared and undeclared) under both methods in 2016. **D'Agosto et al.** [**2018**] employed the diff-in-diff (DID) method to analyse the impact of audits on tax compliance behaviour. Their main assumption is constant selection bias, and they find a positive impact of audits on tax compliance behaviour. The latest available study using the Heckman model to estimate the CIT gap was presented by the IMF, whose team made such an estimate in Armenia [**IMF, 2024**].

We are aware of some of the limitations of the Heckman model. One major limitation of the classical Heckman model is its reliance on linear models for describing both the selection and the outcome processes. This limitation can be resolved by various semiparametric estimation techniques that have been increasingly used in combination with the Heckman model [**Marra, Radice, 2013**; **Wojtyś, Marra, 2015**]. Another approach is the adoption of a non-parametric machine learning tool to improve on the prediction accuracy of individual tax bases [**Alaimo Di Loro et al., 2023**]. However, we stick with the Heckman model because, to our knowledge, the audit selection process was more homogeneous in the studied year than in subsequent years, where the targeting of PIT audits tended to be based on local risk analysis, without taking into account the central risk analysis that was mainly prepared for VAT.

Random selection of audit cases and risk-based selection are not considered to be complementary rather than mutually exclusive strategies. Most countries, even those with advanced risk management systems, use outcomes from random audit selection to calibrate and fine-tune risk-based methods of audit selection [**World Bank, 2011**].

Another bottom-up approach relies on survey data. A prominent example is the consumption method developed by **Pissarides and Weber** [**1989**], which estimates the size of income underreporting by the self-employed in the UK based on consumption propensities between the self-employed and wage earners. This method has been applied to different countries under various modifications and the results mostly point out the larger income underreporting by the self-employed relative to wage earners (see **Kukk et al.** [**2020**] for a review).

Finally, the advanced binary selection model developed by **Feinstein** [**2001**] and **Erard and Feinstein** [**2010**] and randomised controlled trials (RCTs) conducted by the State of Minnesota Department of Revenue in cooperation with **Slemrod et al.** [**2001**] and by **Bérgolo et al.** [**2018**] on Uruguayan firms are further examples of the bottom-up method.

So far, few studies have examined the scale of the PIT gap or underreporting of income in Poland.[10] **Klonowska** [**2017**] estimates both the total PIT gap and the PIT gap among self-employed individuals for the 2003–2012 period using a top-down method. Notably, the estimated gap values are very similar for the two groups, and in some years the gap among the self-employed is even lower than that among taxpayers earning income from other sources. This result may appear questionable in the context of findings from studies using bottom-up methods of estimation.

Other studies have employed indirect estimation approaches to quantify the tax gap using less detailed data. **Adamczyk** [**2020**] used a microsimulation model and found that income underreporting among the self-employed ranged from 10% to 11% of actual income, resulting in lost tax revenues of between 8.3% and 9.4% for the period 2013–2016. In another study, **Turgut and Tratkiewicz** [**2023**], using the consumption method developed by **Pissarides and Weber** [**1989**], estimated that self-employed individuals in Poland failed to report one-quarter of their total income over the period 2005–2017. At the same time, some studies have highlighted limitations associated with the consumption method for estimating underreported income. Analyses of fiscal and survey data show that estimates of underreported income derived from fiscal data may be up to twice as high as those obtained using survey-based methods [**Cabral, Gemmell, Alinaghi, 2021**], implying that measurement error may be less problematic when fiscal data are used.

These results identify the need to look for alternative methods of estimating the PIT gap in Poland, particularly approaches that rely more extensively on fiscal data, either to improve accuracy or to complement existing methods. In our view, a bottom-up method using data from risk-based audits offers such an alternative. However, we acknowledge that this method is unlikely to capture the portion of the gap attributable to completely undeclared income. The irregularities identified in audit data relate only marginally to undeclared income and predominantly to issues such as improper expense deductions. This is not surprising as detecting undeclared income is far more difficult than verifying the correct application of the provisions of the PIT Act.

Our analysis does not account for potential differences in auditors' ability to detect irregularities in PIT compliance, an issue widely discussed in the literature, e.g. **Erard and Feinstein** [**2011**]. The audit data available to us contained no information on auditors' experience or effectiveness, nor on the durability of audit findings, i.e. whether they were subsequently overturned in administrative or judicial proceedings. Our results indicate (see Chapter 5.2) that audit effectiveness varies substantially across voivodeships. This variation may reflect differences in auditors' expertise, but it could also stem from other factors, such as differing risk-analysis guidelines issued by supervisory units or varying levels of taxpayer preparedness in fulfilling tax obligations.

Nevertheless, our estimates complement those derived from studies focusing on undeclared income and may contribute to more comprehensive and precise assessments of the overall PIT gap, while also identifying areas where tax administrations may wish to concentrate their audit efforts. By applying the Heckman model to estimate the PIT gap in Poland, this study contributes to the literature in several ways. First, to the best of our knowledge, it is the first study to apply the Heckman model to estimate the PIT gap using data from tax returns combined with risk-based audits. Second, the framework we developed for specifying the Heckman

---

[10]　　We omit studies on the shadow economy.

model may serve as a guide for variable selection in contexts where theory cannot be used as a yardstick. Finally, despite the informational constraints of the audit data, we successfully attempted to decompose the PIT gap by type of activity and by region, thereby providing additional insights into its structure.

Methodology

For the estimation, we define the PIT gap as the tax loss due to the difference between the potential (audited plus declared) tax liability and the declared tax liability of individual taxpayers. This gap is not always positive. It may be zero (if there is no additional tax after the audit) and even negative in some cases (if there is less tax to pay after the audit). The personal income tax liability can be different among taxpayers depending on the source of income.

The Heckman model suits our purposes since we use risk-based audit data that may involve selection bias due to the non-random nature of audit selection. By using the Heckman model, we can assess whether selection bias is present, identify the factors contributing to it, and control for this bias when estimating the PIT gap. Moreover, by correcting for selection bias, we can predict the total PIT gap, which is the amount of uncollected tax among all taxpayers included in this study.

There are two main equations in the Heckman model, the selection equation given in Eq. (1) and the outcome equation given in Eq. (2):

$$z_i = \begin{cases} 1, & if\ w_i'\gamma + u_i \geq 0 \\ 0, & otherwise \end{cases}, \tag{1}$$

$$y_i = x_i'\beta + \varepsilon_i. \tag{2}$$

In the above system of equations, y_i is the PIT gap (difference between audited tax liability and declared tax liability); x_i is the vector of explanatory variables (e.g. revenue, location, and industry); z_i is a binary variable that takes value 1 if the taxpayer is selected into audit and the PIT gap is observed and zero otherwise; w_i is a vector that contains all variables in the vector x_i plus additional variables; and ε_i and u_i are the error terms.

There are two observations worth mentioning related to Eqs. (1) and (2). First, the PIT gap, y_i, is not observed for all taxpayers since tax audits are conducted by the tax authorities on selected taxpayers. Second, vector w_i contains at least one exogenous variable not contained by x_i that only affects the audit selection but does not affect the PIT gap in the outcome equation. In other words, at least one exclusion restriction is imposed such that $w_i = (x_i,\ m_i)$, where m_i is the variable contained only in vector w_i but not in x_i.[11]

The assumptions to estimate Eqs. (1) and (2) are the following:

1. $(w_i,\ z_i)$ are always observed. Observability means that the variables in the vector w_i are observable for every taxpayer i regardless of whether he or she is selected for audit or not. y_i is observed only for taxpayers selected for audit, when $z_i = 1$.
2. The error terms $(\varepsilon_i,\ u_i)$ have zero conditional mean given w_i: $E[\varepsilon_i\,|\,w_i] = 0$; $E[u_i\,|\,w_i] = 0$. This implies that the explanatory variables in vector w_i are uncorrelated with the error terms.
3. u_i is normally distributed with zero mean and unit standard deviation: $u_i \sim N(0,\ 1)$.
4. The conditional expectation of ε_i on u_i is linear: $E[\varepsilon_i\,|\,u_i] = \rho\sigma u_i$, where ρ is the correlation between ε_i and u_i and σ is the standard deviation of ε_i.

If ρ is larger than zero, then there is a correlation between the error terms and the standard OLS estimation of Eq. (1) suffers from sample selection (omitted variable) bias. It is convenient to define the covariance term between ε_i and u_i as $\beta_\lambda = \rho\sigma$. **Heckman [1979]** shows that in the selected sample (in our case audited

---

[11] Imposing at least one exclusion restriction is not necessary for the estimation of the Heckman model; however, it mitigates the possible collinearity problem in Eq. (3). This is why we opt to impose an exclusion restriction to avoid collinearity issues during the estimation.

taxpayers), the actual y_i is written as the sum of the expected y_i (conditional on w_i and selection) and a mean-zero residual:

$$y_i = x'_i \beta + \beta_\lambda \lambda (w'_i \gamma) + \eta_i, \qquad (3)$$

where $\lambda (.)$ is the inverse Mills ratio. Hence, the Heckman model corrects for the sample selection (omitted variable) bias by including the inverse Mills ratio as an additional explanatory variable in the structural equation.

Data

We use fiscal and audit data for type I taxpayers who pay a 19% flat tax rate out of their income earned through their business activities, as reported by the Ministry of Finance (MoF) for the fiscal year 2017, which includes two different types of audits performed by tax authorities: checking activities (CZSP) and tax controls (SPR).[12] Table 1 summarises the audit data sets.[13]

Table 1. Summary of audit data for type I taxpayers

Audit type	No. of audits	No. of positive hits	Total PIT gap (in PLN)
CZSP – Checking activities	46,487	12,534	88,095,590
SPR – Tax controls	936	653	39,395,251

Source: Authors' own elaboration.

Checking activities constitute almost 98% of audits conducted for type I taxpayers in Poland. However, the PIT gap assessed after checking activities constitutes 69% of the total PIT gap. The ratio of positive hits (a positive hit occurs when an audit finds a positive tax loss) to the number of all audits is around 70% for tax controls, compared with 27% for checking activities.

The fiscal data includes financial, geographical and sectoral information about all taxpayers carrying out business activity and paying the 19% flat tax. A total of 572,298 taxpayers fell into this category in the fiscal year 2017.

We merge checking activities and tax controls in our analysis and treat them as a single audit activity. The main advantage of this approach is that it allows us to utilise all available information from controls. The drawback is that we do not account for potential heterogeneity in the selection criteria between checking activities and tax controls, if such differences exist.[14] However, the Heckman model can identify selection into audit but cannot distinguish between different audit types into which a taxpayer might potentially be selected, since the audit type is unobservable for unaudited taxpayers. Therefore, it is not possible to have different types of audit controls in the Heckman model.[15]

We then merge the audit dataset with the fiscal dataset and drop observations with missing values for the *powiat* and PKD variables[16] (here, "variable" refers to the information contained in the dataset column with the heading "variable;" e.g. the *powiat* variable specifies each taxpayer's place of residence at the *powiat* level

---

[12] We adopt the following simplifications in our research. First, we assume that the audit results are definitive, although this does not fully correspond to reality. While this assumption generally holds for checking activities, in the case of standard tax controls, some findings may be disputed by taxpayers. Such disputes may move to administrative tax proceedings or even to court, where the positive audit outcome may be challenged in part or in full. Second, we assume that the auditors' skills are the same, i.e. we do not account for potential differences in experience or ability, especially on a regional basis.

[13] The numbers of audits and positive hits are aggregated at the taxpayer level.

[14] We estimated the Heckman model using each audit dataset separately under various model specifications; however, the estimation did not produce reliable results, especially for tax controls (SPR) where the rho term was always close to 1 with a statistically not significant lambda coefficient.

[15] The literature uses the double selection method when there are two stages of selection [**Estache, Iimi, 2011**]. However, this method is not suitable for distinguishing between different types of control in our case as the selection is made in the first stage. There is no second-stage selection because some taxpayers selected for audit are not eliminated from the sample.

[16] A *powiat* is a local self-governing community and a unit of basic territorial division covering the area of several municipalities or the entire area of a city with *powiat* rights (i.e., a municipality with city status that has been granted *powiat* rights). In 2017, there were

in the column labelled "*powiat*"). This results in the master dataset used for the regression analysis.[17] Table 2 summarises the mean values of the selected variables and the distribution of taxpayers across voivodeships and PKD categories in the master dataset.

Table 2 shows that, on average, the tax base, revenue and profit are higher for audited taxpayers than for unaudited ones. It also indicates that the distribution of audited and unaudited taxpayers across voivodeships is not perfectly proportional. For example, the share of audited taxpayers is the second highest in Dolnośląskie, whereas the proportion of unaudited taxpayers in this voivodeship is the fifth highest. A similar pattern can be observed across sector groups, where the proportion of audited taxpayers exceeds that of unaudited taxpayers in construction, trade and transportation. Overall, Table 2 suggests the presence of voivodeship- and PKD-related risk characteristics perceived by the tax authorities as relevant in audit selection. We aim to capture this geographic and sectoral heterogeneity in our estimation using these variables.

Table 2. Summary of master data

Variables / Sample type	Total	Audited	Unaudited
No. of observations	473,764	39,341	434,423
Tax base (in PLN)	223,671	304,355	216,364
PIT gap (in PLN)	2,657	2,657	.
Revenue from business activity (in PLN)	1,710,001	2,684,738	1,621,730
Profit from business activity (in PLN)	229,251	310,626	221,881
Loss from business activity (in PLN)	5,178	5,014	5,192
Distribution by voivodeship (region)			
Dolnośląskie	8.3%	11.9%	8.0%
Kujawsko-pomorskie	4.4%	8.2%	4.1%
Lubelskie	3.5%	5.4%	3.3%
Lubuskie	2.2%	2.1%	2.2%
Łódzkie	5.7%	5.2%	5.7%
Małopolskie	8.4%	7.0%	8.6%
Mazowieckie	24.2%	18.2%	24.7%
Opolskie	1.9%	3.0%	1.8%
Podkarpackie	3.0%	3.4%	3.0%
Podlaskie	2.3%	2.0%	2.4%
Pomorskie	6.9%	7.5%	6.8%
Śląskie	10.2%	6.8%	10.5%
Świętokrzyskie	2.0%	2.7%	2.0%
Warmińsko-mazurskie	2.8%	3.5%	2.7%
Wielkopolskie	9.8%	6.4%	10.1%
Zachodniopomorskie	4.4%	6.7%	4.1%
Distribution by selected sector codes (PKD)			
Construction	9.2%	11.5%	9.0%
Whole and retail trade; repair of motor vehicles and motorcycles	18.7%	20.3%	18.5%
Transportation and storage	5.9%	7.3%	5.8%
Accommodation and food service activities	2.3%	2.8%	2.3%
Information and communication	8.2%	5.7%	8.4%
PST activities (professional, scientific and technical activities)	19.9%	16.1%	20.2%

Notes: The values presented in PLN are mean values per variable.

Source: Authors' own elaboration.

---

314 powiats. PKD (*Polska Klasyfikacja Działalności*), or the Polish Classification of Activities, is a division of the types of economic activities carried out by business entities operating on the market.

[17]   The selection of the explanatory variables is discussed in the next section.

Empirical model and results

We estimate the following system of equations given in Eqs. (4) and (5) – the selection and outcome equations respectively – using the Heckman model:

$$z_i^* = w_i'\gamma + u_i, \tag{4}$$

$$z_i = \begin{cases} 1, & if \ z_i^* \geq 0 \\ 0, & otherwise \end{cases},$$

$$y_i = x_i'\beta + \beta_\lambda \lambda(w_i'\gamma) + \eta_i, \tag{5}$$

where z_i is a binary variable equal to 1 if the taxpayer is selected for audit; y_i is the PIT gap; x_i is the vector of explanatory variables in the outcome equation including revenue, the tax base, a loss dummy (equal to 1 if the cost from business activity is larger than revenue from business activity and zero otherwise), an excess tax payment dummy (takes the value of 1 if the difference between the advance payment and the tax due is positive); *powiat* of residence and business sector dummies; and $\lambda(w_i'\gamma)$, the inverse Mills ratio. w_i is the vector of explanatory variables in the selection equation that includes all the variables in vector x_i and a VAT dummy equal to 1 if the taxpayer is liable for VAT (value-added tax) and zero otherwise.

The VAT dummy was included to impose at least one exclusion restriction and thus avoid collinearity. The validity of the exclusion restriction rests on the assumption that VAT liability has a direct effect on the probability of being selected for audit but no direct effect on the PIT gap. The intuition behind using VAT liability as an exclusion restriction is based on informational asymmetry. Tax authorities may possess more financial information on VAT-liable taxpayers than on non-liable taxpayers, enabling them to assess PIT compliance risk more effectively for the former group. However, there is no documented evidence of a link between VAT liability and personal income tax evasion. Therefore, VAT liability is unlikely to directly influence the PIT gap, which supports the validity of the VAT dummy as a variable to satisfy the exclusion restriction.[18]

Before discussing the results, we outline the procedure used to select the variables. The master dataset contain 53 potential explanatory variables for each taxpayer.[19] These variables were identified based on an analysis of the taxpayer characteristics typically considered by the tax administration when selecting taxpayers for audit. This simplified approach was due to the lack of detailed data for each audit in terms of the specific tax risks and their magnitude that justified the audit decision. In a set of potential explanatory variables, we first calculate the information value (IV) of each variable.[20] The IV is calculated by splitting continuous variables into bins according to the following formula:

$$IV = \sum_{k=1}^{N}\left(\frac{unaudit_k}{total\ unaudit} - \frac{audit_k}{total\ audit}\right) \times \ln\left(\frac{unaudit_k\ /total\ unaudit}{audit_k\ /\ total\ audit}\right), \tag{6}$$

where k splits the variable into bins, N is the total number of bins; $unaudit_k$ and $audit_k$ denote the total number of unaudited and audited taxpayers in the kth bin respectively; *total unaudit* and *total audit* are the total number of audited and unaudited taxpayers in the sample; and ln is the natural logarithm.[21] IV measures the ability of the explanatory variables to predict the dependent variable (in our case z_i, the probability of selec-

---

[18]   There is no formal statistical test to determine whether the VAT dummy satisfies the exclusion restriction assumption. Therefore, we rely on intuition and data about the ratios of positive PIT audit hits among taxpayers registered and not registered for VAT purposes. The positive-hit ratios are nearly identical for VAT-registered and non-registered taxpayers, at 28.63% and 28.83% respectively.

[19]   In total, the master dataset includes 53 variables: 50 derived from fiscal data, one from audit data, and two additional dummy variables constructed for this study – the loss dummy and the excess tax payment dummy.

[20]   As there is no unified theory of tax evasion that can guide an empirical analysis, we apply the information value (IV) technique to select appropriate variables for the analysis.

[21]   A bin refers to an interval into which the original data values are grouped. For example, k = 1 may represent values of a variable ranging from 0 to 100. The number of bins, in our case k and N, depends on the size of the data.

tion for audit). Following **Siddiqi** [**2012**], we exclude variables with an IV below 0.02.[22] This eliminates 29 of 53 potential explanatory variables. Then, we conduct a multi-collinearity test among the remaining 18 variables and remove the variables with high variance inflation factors (VIF).[23,24] This eliminates additional 11 potential explanatory variables. Finally, we run a preliminary estimation of the Heckman model and drop variables that are not statistically significant at 1% level in one of the equations.[25] This three-step procedure leaves us with explanatory variables in vectors w_i and x_i.

We estimate Eqs. (4) and (5) using the Heckman approach with a two-step estimator. We winsorize the PIT gap and other continuous variables at the 1st and 99th percentiles to remove the impact of outliers in the regression. Table 3 displays the results of the estimation.

Table 3. Heckman estimation results

Explanatory variables	Coefficient	Standard error
Outcome (primary) equation		
Revenue	−220.3998***	49.2846
Excess payment dummy	−2,995.4281***	406.8896
Tax base	1,616.4079***	333.1207
Loss dummy	3,214.6069***	632.8101
Selection equation		
VAT dummy	0.0531***	0.0109
Revenue	0.0205***	0.0011
Excess payment dummy	0.1693***	0.0057
Tax base	0.1285***	0.0104
Loss dummy	−0.2531***	0.0123
Inverse Mills Ratio		
Lambda (β_λ)	−8,592.7536*	2,872.3657
Observations	473.764	
Censored (non-selected) observations	434.423	
Chi-sq	2,222.25	
Degree of freedom	383	
Rho	−0.8260	
Sigma	10,403.30	

Notes: Coefficients of *powiat* and PKD dummies are suppressed. The stars denote * p < 0.01, ** p < 0.001, *** p < 0.0001. The revenue and tax base variables are expressed in million units (nominal value divided by a million).

Source: Authors' own elaboration.

Table 3 shows that all estimated coefficients are statistically significant at the 0.01% level. The coefficient on the VAT dummy in the selection equation is positive, which means that the probability of selection for audit increases if the taxpayer is liable for VAT. This result is consistent with our intuition. Moreover, the estimated coefficient of the VAT dummy is statistically significant at the 0.01% level in the selection equation,

---

[22] **Siddiqi** [**2012**] shows that variables with an IV below 0.02 are not useful predictors. The full list of variables and their corresponding IVs is available on request.

[23] We calculate the VIF for each variable and drop variables with a VIF value higher than 10. When several variables exhibit similar VIF values, we drop the one with the lower IV. However, this can be only done for continuous variables and binary categorical variables. Therefore, this analysis does not include *powiat* (two variables), voivodeship (two variables), and sector code variables since they are categorical variables with more than two values. Moreover, we drop the loss and excess tax payment variables because we use loss and excess tax payment dummies to avoid estimation problems. Finally, we incorporate the VAT dummy into the VIF analysis, which explains why the remaining variables number 18 rather than 24.

[24] We drop the business *powiat* variable (indicating where the taxpayer conducts business activity) from the explanatory vector because it has a 0.92 correlation coefficient with the residential *powiat* variable and has a lower IV. Moreover, we exclude both residential and business voivodeship variables, as they are collinear with the *powiat* variable.

[25] We adopt a 1% significance level as the threshold due to a large number of observations in our sample.

indicating the existence of sample selection bias (see **Certo et al.** [**2016**]). Higher revenue, a larger tax base, and excess tax payments increase the likelihood of selection for audit, whereas declaring a loss decreases it.

We also report the estimated correlation between the residuals of the selection and outcome equations, rho (ρ), the standard error of the residual in the outcome equation, sigma (σ_ϵ), and the coefficient on the inverse Mills ratio, $\lambda(.)$, that is β_λ in the outcome equation. The ρ with a negative coefficient indicates a negative correlation between the unobservable determinants of audit selection and the PIT gap, implying that running OLS would produce biased results. The Heckman model corrects for this bias and produces consistent estimates. This is further confirmed by the negative and statistically significant β_λ at the 1% level in the outcome equation.

We assess the fit of the model by calculating the Area Under ROC Curve (AUC) for the selection (probit) equation. The ROC (Receiver Operating Characteristics) curve shows sensitivity (the true positive rate), the number of correctly predicted audited taxpayers by the model relative to the actual number of audited taxpayers, and 1-specificity (the 1-true negative rate), the number of incorrectly predicted audited taxpayers by the model relative to the actual number of unaudited taxpayers. Figure 1 plots the ROC curve where sensitivity is displayed on the y-axis and 1-specificity on the x-axis. AUC can be considered as a measure of the accuracy of the model. Figure 1 gives the value of AUC equal to 72.60, meaning that the average value of sensitivity for all possible values of specificity is equal to 72.60%.[26] In other words, the probability that the model will correctly predict the selection of an audited taxpayer into the audit is 72.60% on average. This value is within the limits of the acceptance zone.[27] While some may argue that an AUC close to 73% is only moderately informative, we contend that the model provides valuable insights into audit selection and PIT compliance. Its practical usefulness for the tax administration can ultimately be evaluated through implementation and testing in actual audit procedures.

Figure 1. ROC Curve
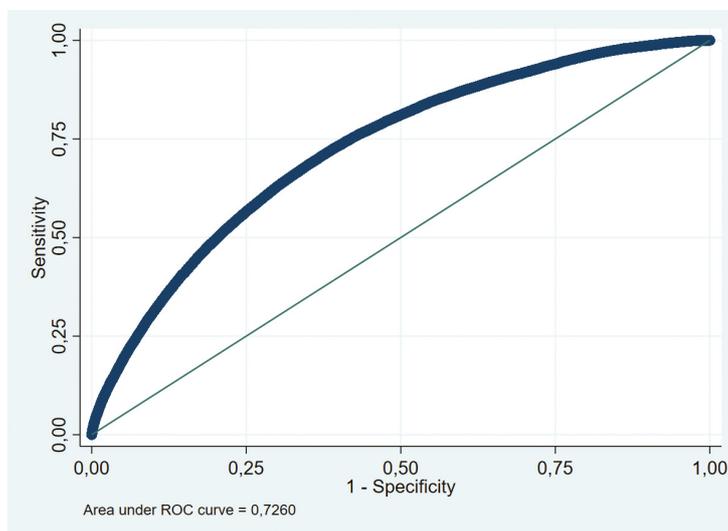
Source: Authors' own elaboration.

---

[26]   We also estimate the Heckman model including all variables with an IV greater than 0.02 in order to check how the AUC varies across models. We find that the AUC is 0.7265, which is almost identical to that of the baseline model. This indicates that dropping variables with similar VIFs does not reduce the predictive power of the model for the selection process.

[27]   An AUC of 1 means the model has perfect accuracy and an AUC of 0.5 means the model has no better accuracy than chance. An AUC between 0.7 and 0.8 is generally viewed as an acceptable zone [**Mandrekar, 2010**].

Prediction

In this section, we use the estimated model to forecast the total PIT gap in Poland for the fiscal year 2017 for type I taxpayers. First, we look at the aggregate prediction and then at the predictions at the voivodeship and PKD levels.[28] We apply the following equation to predict the PIT gap for unaudited taxpayers:[29]

$$E[PIT\ Gap_i \mid x_i,\ z_i = 1] = x_i'\hat{\beta} + \hat{\beta}_\lambda \lambda\,(w_i'\gamma), \tag{7}$$

where the hatted terms denote the estimated coefficients from Table 3. Table 4 presents the model-predicted PIT gap where the mean predicted PIT gap amounts to PLN 1,880.44. Multiplying this figure by the number of unaudited taxpayers (525,016) yields a total predicted PIT gap of PLN 987,261,087 for the fiscal year 2017. Adding the actual PIT gap to the predicted PIT gap gives the total PIT gap for type I taxpayers in 2017 of PLN 1,119,699,808:

$$Type\ I\ Total\ PIT\ Gap = \text{PLN } \mathbf{1{,}119{,}699{,}808{,}}$$

i.e. approximately PLN 1.12 billion, or around 4.5% of the total tax liability of this group of taxpayers. Table 4 also presents summary statistics for the actual (winsorized) and model-predicted PIT gap among audited taxpayers. We see that the mean of the predicted PIT gap for unaudited taxpayers is slightly higher than the mean of the actual PIT gap (winsorized).

Table 4. Model-predicted vs. actual PIT gap

	Observation	Mean	Std. Dev.	Min.	Max.
Predicted PIT Gap – Unaudited	434,423	1,880.44	3,264.67	−24,683.6	30,682.4
Actual PIT Gap – Winsorized	39,341	1,768.88	7,184.83	−4,098.0	51,677.0
Predicted PIT Gap – Audited	39,341	1,768.88	1,883.99	−3,499.7	26,529.4

Source: Authors' own elaboration.

We now turn to the prediction of the PIT gap for unaudited taxpayers by voivodeship and sector. This exercise aims to help tax authorities identify the regions and industries with the greatest potential undiscovered PIT gap and thus improve audit targeting.

Prediction by voivodeship

Table 5 presents the mean actual (winsorized) PIT gap for audited taxpayers and the predicted PIT gap for unaudited taxpayers. It also reports the number of observations for both series at the voivodeship level, along with the absolute and relative differences between the two measures. The model predicts the largest mean PIT gap in the Lubuskie and Wielkopolskie regions, and the smallest difference in the Łódzkie and Zachodniopomorskie voivodeships. These results show that audits targeting taxpayers in voivodeship with the largest predicted PIT gap can potentially result in the largest tax collection per person. In other words, the highest expected tax collection from one additional audit would, on average, be achieved by auditing a taxpayer residing in the Lubuskie voivodeship. The absolute and relative difference columns compare the model-predicted PIT gap with the actual (winsorized) PIT gap. For example, Łódzkie has a negative difference in levels, suggesting that the tax authorities have targeted the riskiest taxpayers in that voivodeship. An additional audit of an unaudited taxpayer would yield PLN 343.40 on average, whereas the audits already conducted generated PLN 1,708.20 on average.

---

[28]   By "aggregate" we mean the total population of type I taxpayers.
[29]   This is the expected value of the PIT gap from the underlying distribution of the regression model.

The model's predictions also provide insights relevant to tax policy and audit management in Poland. The results indicate that the risk analysis conducted for the purpose of risk-based audits is most effective in the Łódzkie, Warmińsko-mazurskie and Zachodniopomorskie voivodeships. While these are not the voivodeships with the highest mean PIT collection from audits, the combination of positive audit outcomes and model predictions indicates that additional audits would result in the lowest mean tax recovery compared with those already carried out. This means that risk-based audits in these three voivodeships are the most optimal.

We have also identified voivodeships where risk-based audits are least effective, including Wielkopolskie, Małopolskie and Opolskie. In these voivodeships, additional audits carried out in accordance with the model's predictions would result in significantly higher mean tax collection than those carried out on the basis of risk analysis. This means that the model predicts risky taxpayers better than the tax authorities themselves do. These findings warrant further investigation into why risk analysis is carried out so inefficiently in these voivodeships and what needs to be done to improve the performance of tax inspectors. Overall, our model provides a useful tool for evaluating the efficiency of tax administration.

Table 5. Actual vs. predicted PIT gap by voivodeship

Voivodeship	Actual PIT gap (audited taxpayers)		Predicted PIT gap (unaudited taxpayers)		Level difference	Relative difference
	Obs.	Value (in PLN)	Obs.	Value (in PLN)	(in PLN)	
Dolnośląskie	4,694	2,211.4	34,579	2,239.0	27.63	1.01
Kujawsko-pomorskie	3,208	1,791.8	17,677	2,111.8	320.06	1.18
Lubelskie	2,108	1,427.8	14,492	1,171.8	−256.00	0.82
Lubuskie	842	2,349.8	9,663	2,625.5	275.71	1.12
Łódzkie	2,045	1,708.2	24,703	343.4	−1,364.86	0.20
Małopolskie	2,761	1,539.5	37,156	2,142.2	602.71	1.39
Mazowieckie	7,154	1,650.6	107,493	1,864.5	213.87	1.13
Opolskie	1,193	989.7	7,906	1,443.9	454.26	1.46
Podkarpackie	1,346	2,174.6	13,049	2,037.6	−137.01	0.94
Podlaskie	769	2,164.1	10,242	2,322.1	157.94	1.07
Pomorskie	2,961	2,043.3	29,635	2,321.1	277.80	1.14
Śląskie	2,658	1,644.4	45,525	1,805.3	160.87	1.10
Świętokrzyskie	1,061	1,999.9	8,612	1,989.1	−10.75	0.99
Warmińsko-mazurskie	1,384	1,748.6	11,710	1,323.0	−425.54	0.76
Wielkopolskie	2,520	1,734.4	44,018	2,394.9	660.52	1.38
Zachodniopomorskie	2,637	1,446.8	17,963	1,141.4	−305.33	0.79

Source: Authors' own elaboration.

Prediction by PKD

Table 6 presents the mean actual PIT gap and the mean predicted PIT gap, along with the number of observations for both groups at the PKD level, as well as the absolute and relative differences between them. The construction sector exhibits the largest predicted mean PIT gap, whereas the human health and social work activities sector shows the lowest. This implies that, on average, the largest tax collection from an additional audit could be achieved by targeting self-employed taxpayers in the construction sector. Our findings are consistent with the long-standing view of the tax administration that the construction sector is characterised by elevated tax risk [**NIK, 2019**]. Notably, this sector was among those in which the reverse-charge mechanism was introduced in 2017 to counter fictitious service transactions, which had generated losses not only in VAT but also in income taxes.

We have also identified sectors where risk-based controls are least effective. These include electricity, gas, steam and air conditioning supply; arts, entertainment and recreation; wholesale and retail trade; and repair

of motor vehicles and motorcycles. In these sectors, additional audits carried out in accordance with the model's predictions would result in significantly higher tax collection than audits carried out on the basis of risk analysis. This shows that the model predicts risky taxpayers better than tax authorities do.

This finding warrants further examination of why risk analysis is inefficient in these sectors, particularly in those with a low number of audits, such as electricity, gas, steam and air conditioning supply. Our model provides a useful framework for evaluating the effectiveness of tax administration activities across different PKD sectors.

Table 6. Actual vs. predicted PIT gap by PKD

PKD	Actual PIT gap		Model-predicted PIT gap		Absolute difference	Relative difference
	Obs.	Value (in PLN)	Obs.	Value (in PLN)	(in PLN)	
Agriculture	455	2,659.6	3,428	2,593	−66.61	0.97
Mining and quarrying	77	1,793.6	679	1,506.5	−287.13	0.84
Manufacturing	4,161	2,232.1	37,593	2,424.3	192.21	1.09
Electricity, gas, steam and air conditioning supply	41	265.6	370	1,019.3	753.73	3.84
Water supply, sewerage, waste management and remediation services	170	2,189.9	1,520	2,224.2	34.34	1.02
Construction	4,508	2,860.0	39,173	3,037.7	177.71	1.06
Wholesale and retail trade; repair of motor vehicles and motorcycles	7,993	1,802.0	80,372	2,115.4	313.47	1.17
Transportation and storage	2,863	2,004.7	25,296	2,281.3	276.52	1.14
Accommodation and food service activities	1,085	2,375.4	9,896	2,558.1	182.73	1.08
Information and communication	2,235	1,403.8	36,577	1,561.7	157.94	1.11
Financial and insurance activities	1,364	1,312.4	15,123	1,463.3	150.93	1.12
Real estate activities	1,098	1,200.9	12,996	1,343.9	142.91	1.12
PST activities (Professional, scientific and technical activities)	6,337	1,302.8	87,800	1,482.8	180.03	1.14
Administrative and support service activities	1,087	2,491.5	11,288	2,625.7	134.21	1.05
Education	961	1,261.6	11,252	1,360.9	99.28	1.08
Human health and social work activities	4,177	776.3	52,523	969.2	192.88	1.25
Arts, entertainment activities	370	1,776.5	4,457	2,095.9	319.43	1.18
Other service activities	359	1,769	4,080	1,950	180.96	1.10

Source: Authors' own elaboration.

Conclusions

In this paper, we estimated the PIT gap in Poland for self-employed taxpayers who pay a 19% flat tax rate on business income, using income tax returns and risk-based audit data. We addressed the sample selection problem inherent in risk-based audit data by applying the two-step Heckman model. We then extrapolated the estimation results to predict the total PIT gap for the targeted taxpayer population. Our estimates indicate an average PIT gap of PLN 1,880 per taxpayer and a total PIT gap of PLN 1.12 billion in 2017. Recorded revenue from this source of income tax in 2017 amounted to PLN 23.99 billion, meaning that the estimated PIT gap represents 4.5% of the total tax liabilities of this group of taxpayers. It is difficult to relate this figure to available international estimates, as the PIT gap estimate applies only to a subset of self-employed taxpayers in Poland and only to personal income tax (excluding social security contributions and capital gains tax). For reference, the income tax, NICs and capital gains tax gap arising from total self-assessment in the UK for the 2020–2021 tax year was estimated at 13.9% [**HMRC, 2022**], three times higher than in Poland.

Our results show that the Heckman model can serve as a useful tool for tax authorities seeking to improve the efficiency of risk-based audits. For example, we find that the average PIT gap among all taxpayers (audited and unaudited) is higher than the average gap among audited taxpayers alone. This implies that some non-compliant taxpayers were not selected for audit and that there is scope to improve the audit selection process. We also find that declaring a loss in PIT returns affects both the size of the PIT gap and the probability of being selected for audit, which has direct policy implications. In addition, our results indicate that conducting additional audits in the construction sector could yield the greatest benefits for tax authorities, as this sector has the highest predicted PIT gap on average. Finally, based on the model, we can determine in which voivodeships an audit will have the best effect. At the regional level, the model predicted the largest mean PIT gap in Lubuskie and Wielkopolskie, while the smallest differences between the predicted and actual gaps occur in Łódzkie and Zachodniopomorskie. These findings suggest that targeting taxpayers residing in voivodeships with the highest predicted PIT gap could result in the largest tax collection per person. The highest expected yield from one additional audit would, on average, be achieved in the Lubuskie voivodeship.

This information may assist the National Revenue Administration in better targeting audits of taxpayers subject to the 19% flat PIT rate. Beyond identifying regions with the highest potential tax recovery from additional audits, our model also enables a comparison between predicted returns from additional audits and the returns achieved by audits already conducted. This makes it possible to identify voivodeships where audit efficiency is relatively high or low, thereby providing a basis for managerial action to improve performance where necessary. Similar analyses can be conducted across PKD sectors.

The model assumes that the selection process is adequately captured by the variables considered. We are aware that actual tax audits involve discretionary and unpredictable factors that may not be fully accounted for by the model. However, we assume that such factors affect only a limited number of audits, particularly given that tax inspectors relied more heavily on general audit guidelines during the studied period than they do now. In order to develop the Heckman model in the future, this assumption should be revisited, especially as risk analysis in income tax audits has significantly improved in recent years. The robustness of our results is supported by alternative specifications that include different sets of variables (excluded from the final model due to high VIF values), which yield qualitatively similar findings.[30]

Although our methodology corrects for sample selection bias, further improvements are possible. First, information on the specific reason for each audit should be systematically recorded and included in the results; this would allow a more detailed decomposition of the PIT gap, which was not feasible for the period under study. Second, data on the experience and skills of tax inspectors should be incorporated. The absence of such information may bias the results both geographically and across taxpayer characteristics. For example, more experienced inspectors in certain regions may be more effective in selecting taxpayers for audit and detecting irregularities in complex tax records.

Overall, the model presented here has considerable potential not only for estimating the PIT gap but also for improving the efficiency of selecting taxpayers for audits. The results can be used for preventive and informational measures, which play a crucial role in reducing the tax gap. It would also be valuable to develop a similar model for self-employed taxpayers using forms of PIT other than the 19% flat rate. This would make it possible to assess compliance across different tax regimes and support a more effective allocation of resources by the National Revenue Administration in auditing taxpayers.

References

Adamczyk A. [2020], *Zjawisko zaniżania dochodu przez samozatrudnionych – istota, determinanty, pomiar*, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.

Advani A., Elming W., Shaw J. [2021], The dynamic effects of tax audits, *Review of Economics and Statistics*, 105: 1–45.

---

[30]   Due to text volume restrictions, alternative estimates can be provided on request.

Alaimo Di Loro P., Scacciatelli D., Tagliaferri G. [2023], G. 2-step Gradient Boosting approach to selectivity bias correction in tax audit: an application to the VAT gap in Italy, *Statistical Methods & Applications*, 32: 237–270, https://doi.org/10.1007/s10260-022-00643-4.

Andreoni J., Erard B., Feinstein J. [1998], Tax compliance, *Journal of Economic Literature,* 36(2): 818–860.

Bérgolo M., Ceni R., Cruces G., Giaccobasso M., Perez-Truglia R. [2018], Misperceptions about tax audits, *American Economic Association Papers and Proceedings*, 108: 83–87.

Bloomquist K. M. [2003], Tax evasion, income inequality and opportunity costs of compliance, in: *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*: 91–104, National Tax Association.

Cabral A. C. G., Gemmell N., Alinaghi N. [2021], Are survey-based self-employment income underreporting estimates biased? New evidence from matched register and survey data, *International Tax and Public Finance*, 28(2): 284–322, https://doi.org/10.1007/s10797-020-09611-8.

Certo S. T., Busenbark J. R., Woo H. S., Semadeni M. [2016], Sample selection bias and Heckman models in strategic management research, *Strategic Management Journal*, 37(13): 2639–2657.

Christie E., Holzner M. [2006]. *What explains tax evasion? An empirical assessment based on European data*, Wiener Institut für Internationale Wirtschaftsvergleiche Working Papers No. 40.

Chudý M., Gábik R., Bukovina J., Šrámkova L. [2020], *Searching for gaps: Bottom-up approach for Slovakia*, Economic Analysis – Policy Paper No. 54, https://www.mfsr.sk/sk/financie/institut-financnej-politiky/publikacie-ifp/ekonomicke-analyzy/54.html, (accessed on 25.12.2024).

D'Agosto E., Manzo M., Pisani S., D'Arcangelo F. M. [2018], The effect of audit activity on tax declaration: Evidence on small businesses in Italy, *Public Finance Review*, 46(1): 29–57.

Departament Podatków Dochodowych [2018], *Informacja dotycząca rozliczenia podatku dochodowego od osób fizycznych za 2017 rok*, Warszawa, https://dane.gov.pl/pl/dataset/191, informacje-dotyczace-rozliczenia-podatku dochodowego-od-osob-fizycznych (accessed on: 25.12.2024).

Erard B., Feinstein J. [2011], *The Individual Income Reporting Gap: What We See and What We Don't*, Recent Research on Tax Administration and Compliance Selected Papers Given at the 2011 IRS-TPC Research Conference: New Perspectives on Tax Administration, https://www.irs.gov/pub/irs-soi/11rescon.pdf, (accessed on 25.12.2024).

Erard B., Feinstein J. [2010], Econometric models for multi-stage audit processes: an application to the IRS national research program, in: Aim J., Martinez-Vazquez J., Torgler B. (eds.), *Developing Alternative Frameworks for Explaining Tax Compliance*: 113–137, Routledge, London–New York.

Estache A., Iimi A. [2011], Bidders' entry and auctioneer's rejection: applying a double selection model to road procurement auctions, *Journal of Applied Economics*, 14(2): 199–223.

Feige E. L. [1990], Defining and Estimating Underground and Informal Economies: The New Institutional Economics Approach, *World Development*, 18(7): 989–1002.

Feinstein J. S. [1999], Approaches for estimating noncompliance: examples from federal taxation in the United States, *The Economic Journal*, 109(456): 360–369.

Feldman N. E., Slemrod J. [2007], Estimating tax noncompliance with evidence from unaudited tax returns, *The Economic Journal*, 117 (March): 327–352.

FISCALIS [2017], *The concept of tax gaps. Report on VAT gap estimations, prepared by FISCALIS Tax Gap Project Group* (FPG/041), Brussels, March 2016, https://taxation-customs.ec.europa.eu/system/files/2016-09/tgpg_report_en.pdf (accessed on 25.12.2024).

FISCALIS [2018]. *Corporate Income Tax Gap Estimation Methodologies*, EU Commission Taxation Papers WP No. 73, https://taxation-customs.ec.europa.eu/system/files/2018-11/taxation_papers_73_en.pdf (accessed on 25.12.2024).

Gemmell N., Hasseldine J. [2012], *The tax gap: a methodological review*, Working Paper 20, Emerald Group Publishing Limited.

Heckman J. J. [1979], Sample selection bias as a specification error, *Econometrica: Journal of the Econometric Society*, 47: 153–161.

Hill R. C., Adkins L. C., Bender K. A. [2003], *Test statistics and critical values in selectivity models*, in: Fomby Th. B., Hill R. C. (eds.), *Maximum Likelihood Estimation of Misspecified Models: Twenty Years Later*, Emerald Group Publishing Limited.

HMRC [2022], *Measuring tax gaps 2022 edition: Tax gap estimates for 2020 to 2021*, https://webarchive.nationalarchives.gov.uk/ukgwa/20230206161139/https://www.gov.uk/government/statistics/measuring-tax-gaps (accessed on 25.12.2024).

Hyndman R. J., Koehler A. B. [2006], Another look at measures of forecast accuracy, *International Journal of Forecasting*, 22(4): 679–688.

IMF (International Monetary Fund) [2021], *The Revenue Administration Gap Analysis Program: An Analytical Framework for Personal Income Tax Gap Estimation*, International Monetary Fund Technical Notes and Manuals, TNM/2021/009.

IMF (International Monetary Fund) [2024], *Fiscal Affairs Dept. "Republic of Armenia: Technical Assistance Report-Corporate Income Tax Gap Prediction 2023 Based on Corporate Income Tax Returns"*, Technical Assistance Reports 2024, 088 (2024), https://doi.org/10.5089/9798400291579.019.

IRS (Internal Revenue Service) [1996], *Federal Tax Compliance Research: Individual Income Tax Gap Estimates for 1985, 1988, and 1992*, IRS Publication 1415 (Rev. 4–96), Washington, DC.

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

Klonowska A. [2017], *Luka podatkowa. Skutki dla finansów publicznych*, C. H. Beck, Warszawa.

Kukk M., Paulus A., Staehr K. [2020], Cheating in Europe: underreporting of self-employment income in comparative perspective, *International Tax and Public Finance*, 27 (2): 363–390.

Mandrekar J. N. [2010], Receiver operating characteristic curve in diagnostic test assessment, *Journal of Thoracic Oncology*, 5 (9): 1315–1316.

Marra G., Radice R. [2013], Estimation of a regression spline sample selection model, Computational Statistics & Data Analysis, 61: 158–173.

NIK (Najwyższa Izba Kontroli) [2019], *Informacja o wynikach kontroli, Nadzór Ministra Finansów nad poborem podatku od towarów i usług*, KBF.430.002.2019, https://www.nik.gov.pl/plik/id,21004, vp,23636.pdf (accessed on 25.12.2024).

OECD [2004], *Compliance Risk Management: Use of random audit programs*, https://www.oecd.org/content/dam/oecd/en/topics/policy-issues/tax-administration/compliance-risk-management-audit-case-selection-systems.pdf (accessed on 25.12.2024).

Pissarides C., Weber G. [1989], An Expenditure-Based Estimate of Britain's Black Economy, *Journal of Public Economics*, 39 (1): 17–32.

Richardson G. [2006], Determinants of tax evasion: A cross-country investigation, *Journal of International Accounting, Auditing and Taxation*, 15: 150–169.

Rubin M. [2011], *The Practicality of a Top Down Approach to the Direct Tax Gap*, The IRS Research Bulletin, Recent Research on Tax Administration and Compliance: Selected Papers Given at the 2011 IRS-TPC Research Conference: New Perspectives on Tax Administration.

Siddiqi N. [2012], *Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring*, SAS Institute Inc.

Slemrod J. [2007], Cheating ourselves: The economics of tax evasion, *Journal of Economic Perspectives*, 21 (1): 25–48.

Slemrod J., Blumenthal M., Christian C. [2001], Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota, *Journal of Public Economics*, 79 (3): 455–483.

The World Bank [2011], *Risk-Based Tax Audits, Approaches and Country Experiences*, https://documents.worldbank.org/pt/publication/documents-reports/documentdetail/490491468159916971/risk-based-tax-audits-approaches-and-country-experiences (accessed on 25.12.2024).

Toder E. [2007], What is the tax gap, *Tax Notes*, 117 (4): 367–378.

Turgut M. B., Tratkiewicz T. [2023], Estimate of the Underground Economy in Poland Based on Household Expenditures and Incomes, *Central European Journal of Economic Modelling and Econometrics*, 15: 1–29.

Wojtyś M., Marra G. [2015], *Copula based generalized additive models with non-random sample selection*, arXiv preprint arXiv:1508.04070.

Wu S.-Y., Teng M. J. [2005], Determinants of tax compliance – A cross-country analysis, *Finanzarchiv/Public Finance Analysis*, 61 (3): 393–417.