RUCH PRAWNICZY, EKONOMICZNY I SOCJOLOGICZNY
Rok LXXXVII – zeszyt 1 – 2025
https://doi.org/10.14746/rpeis.2025.87.1.12

MARCIN GORAZDA,[a] TOMASZ KWARCIŃSKI[b]

THE RELIABILITY OF TAX DATA FOR MEASURING DISPOSABLE INCOME INEQUALITY IN POLAND

WIARYGODNOŚĆ DANYCH PODATKOWYCH W POMIARZE NIERÓWNOŚCI DOCHODÓW ROZPORZĄDZALNYCH W POLSCE

This paper critically examines the suitability of tax data for measuring disposable income inequality, with a focus on the Polish income tax system. It assesses the limitations of relying solely on tax data and proposes an alternative approach that combines survey data with administrative records for improved accuracy. The analysis draws on data from tax reports, national accounts, and a review of relevant literature on taxpayer behaviour. By analysing issues such as discrepancies between disposable income and total income, misreporting, and adaptive taxpayer responses to tax regulations, the study highlights the complexity and volatility of tax systems. The findings suggest that researchers need a thorough understanding of both tax law and taxpayer behaviour to avoid substantial errors in measuring income inequality accurately.

Keywords: income inequality; tax data; survey data; disposable income; Polish income tax system
JEL: D31, C83, H24, H26

Artykuł krytycznie analizuje przydatność danych podatkowych do pomiaru nierówności dochodów do dyspozycji, koncentrując się na polskim systemie podatkowym. Celem artykułu jest ocena ograniczeń wynikających z wykorzystania wyłącznie danych podatkowych oraz zasugerowanie alternatywnego podejścia, które łączy dane z badań ankietowych z rejestrami administracyjnymi w celu poprawy wiarygodności badania. Analiza opiera się na danych z deklaracji podatkowych, rachunków narodowych oraz przeglądu literatury dotyczącej zachowań podatników. Analizą takich kwestii, jak rozbieżności między dochodem do dyspozycji a dochodem całkowitym, zaniżanie dochodów oraz adaptacyjne reakcje podatników na regulacje podatkowe, zwracamy uwagę na złożoność i zmienność systemów podatkowych. Na podstawie przeprowadzonych analiz wnioskujemy, że badacze potrzebują dogłębnego zrozumienia zarówno prawa podatkowego, jak i zachowań podatników, aby uniknąć istotnych błędów w rzetelnym pomiarze nierówności dochodowych.

Słowa kluczowe: nierówności dochodowe; dane podatkowe; dane ankietowe; dochód dyspozycyjny; polski system podatkowy
JEL: D31, C83, H24, H26

[a] Jagiellonian University in Krakow, Poland / Uniwersytet Jagielloński w Krakowie, Polska
marcin.gorazda@gsw.com.pl, https://orcid.org/0009-0002-9122-8440
[b] Krakow University of Economics, Poland / Uniwersytet Ekonomiczny w Krakowie, Polska
kwarcint@uek.krakow.pl, https://orcid.org/0000-0002-9474-4216

RPEiS 87(1), 2025: 217–237. *© WPiA UAM, 2025.* *ISSN (Online) 2543-9170 ISSN (Print) 0035-9629*
Open Access article, distributed under the terms of the CC licence (BY, https://creativecommons.org/licenses/by/4.0/deed.en).

I. INTRODUCTION

Income distribution analysis typically relies on two primary data sources: household budget surveys and income tax data. Statistical agencies in developed countries often use surveys to estimate income, providing a foundational element for economic research on inequality. These surveys, conducted annually on representative population samples, provide comprehensive data that is harmonized in certain regions (such as the European Union) but may require adjustments elsewhere to ensure comparability.

Despite their limitations, surveys generally offer ample information for discussions on welfare improvements (Okun, 2015; Ostry et al., 2019), political decision-making (Stoetzer, 2023), and perceptions of fair redistribution. The analysis of income inequality serves multiple purposes, each affecting the choice of measurement approach. We identify three primary concerns highlighted in the literature: welfare, political stability, and social justice.

The first and most widely acknowledged by economists is welfare, as inequalities can impact overall societal welfare, either positively or negatively, depending on distributional dynamics. Secondly, income inequality may influence political choices, with higher inequality often linked to shifts toward populism and challenges to liberal democracy. Lastly, societal perceptions of fairness are closely tied to inequality, as greater disparities are often viewed as unjust, motivating policies that aim to address income imbalances.

However, survey data are often criticized for potential inaccuracies, especially at the extreme ends of the income spectrum. At the upper end, detailed questionnaires demand significant effort from respondents, leading to high rates of non-response or underreporting among wealthier households (Angel et al., 2019; Frenette et al., 2007; Korinek et al., 2006; Yonzan et al., 2022). This issue, known as the 'missing rich' problem (Lustig, 2019), affects the top 1% of the population in developed countries (Burkhauser et al., 2018; Yonzan et al., 2022). Research shows that higher incomes correlate with lower survey response rates (Hlasny & Verme, 2018). Although statistical agencies attempt to correct for this, the significant underrepresentation of the richest individuals remains a concern, particularly when survey results are compared with tax data.

Similar issues arise at the lower end of the income distribution, where the poorest individuals often overestimate their incomes. This discrepancy becomes evident when survey data is compared with tax or consumption data (Szulc, 2022). Social factors, such as the reluctance to admit poverty due to stigma or cognitive biases that skew income perceptions towards the median, contribute to this overestimation (Fernández-Albertos & Kuo, 2018). These concerns drive researchers to investigate income tax data collected by tax administrations for a more accurate assessment.

This paper critiques the use of tax data to measure disposable income inequality, using the Polish income tax system as a case study. The unique aspects of the Polish tax system are frequently overlooked or misinterpreted, leading to inaccurate results. We suggest that for measuring disposable in-

come (as opposed to total income), tax data should not be the sole source of information. Instead, survey data should be used and adjusted with administrative data, including but not limited to tax data. The volatility of tax systems and taxpayers' adaptive behaviours necessitate comprehensive knowledge of both tax laws and taxpayer responses, knowledge which is often inaccessible or neglected due to the effort and cost required to obtain it. Even when adjustments are made, they are often based on estimates and simplifications that significantly affect the results.

The structure of this paper is as follows: Section II outlines the well-known limitations of tax data, focusing on three critical issues: differences in income definitions and population coverage, the distinction between tax units and households, and the misreporting of taxable income. Section III examines the peculiarities of the Polish tax system that impact the reliability of tax data, particularly the prevalence of entrepreneurs among high-income earners, the treatment of retained earnings, and the taxation of pass-through entities. Section IV critiques the common oversimplifications and misapplications of tax data in Polish inequality research, while Section V concludes and summarizes our findings.

II. LIMITATIONS OF TAX DATA

Tax data have been an important source of information about inequality since Simon Kuznets' (1955) pioneering work. These data provide a straightforward measure of inequality for several reasons. First, tax data are considered reliable because they are based on official registers that follow strict protocols. They also capture the incomes of high earners who may not participate in surveys or who underreport their income. Another advantage of tax data is their historical depth, often spanning several decades, which allows for long-term analysis of inequality trends (Atkinson et al., 2011; Bukowski & Novokmet, 2021). Furthermore, tax data enable the measurement of both pre-tax and post-tax inequality, which is useful for analysing the redistributive effects of tax systems.

However, many authors rightly raise intense critiques in this regard (Galbraith, 2018). Despite the above-mentioned advantages, tax data also have significant limitations that economists have long recognized. These limitations include differences in income definitions and population coverage, the distinction between tax units and households, and the misreporting of taxable income.

1. Differences in income definitions and population coverage

Richard Titmuss (1962) noted tax data limitations, particularly the exclusion of non-taxable incomes like capital gains and in-kind benefits. He argued that official reports reflect taxable rather than disposable income – the latter being essential for measuring inequality (Atkinson 1975).

The concepts of disposable income and total taxable income differ significantly. Disposable income is typically the focus of surveys, with statistical agencies collecting data on the income available for individual consumption within the surveyed period, regardless of whether it is taxed. If the income is taxed but not disposable, it is excluded from the survey. Non-taxable or tax-exempt income, as well as taxable income that is not disposable (e.g. unrealized capital gains, retained earnings, company's undistributed profits), reveal significant differences between tax and survey data.

In Poland, taxable and disposable incomes are influenced by unique tax laws, such as lump-sum taxes for farmers, requiring income imputation based on agricultural land quality. For many small businesses, the Polish tax system allows income tax calculation based on annual proceeds/turnover rather than net income. This form of taxation became popular among professional service providers after the 2020 tax reform, which increased the upper turnover threshold to EUR2 million and reduced tax rates to 17%, 15%, and 12%, depending on the services rendered. Small traders, craftsmen, and rental income earners are taxed at lower rates (3%, 5.5%, 8.5%, and 11%).

Entrepreneurs must include gains from disposing of tangibles, intangibles, financial assets, and interests in business profits. Individuals declare these gains in separate tax returns (PIT-38, PIT-39) though only a fraction are reported. Sales of agricultural and real estate held for longer than five years are tax-exempt, as are movables held longer than six months, creating incentives to realize gains after the taxable period expires. In Poland, dividends and interest received are generally taxed at source by the payer, without indicating the taxable person.

Retained earnings and unrealized capital gains complicate inequality research, especially in systems like Poland's, where businesses can choose between different taxation methods (CIT or PIT). These choices influence income reporting and must be considered in inequality calculations.

2. Differences in income subjects: Tax units vs households

Regarding the issue of different income subjects, surveys typically analyse households, while tax data analyse tax units or tax filers. This distinction is significant for at least three reasons.

First, since not all types of income are subject to taxation, relying on tax data may exclude some earners. These exclusions usually pertain to incomes exempt due to their nature or magnitude. In most countries with personal income taxation, there is a threshold below which income is tax-exempt. The higher this threshold, the more income earners are excluded. According to Hümbelin and Farys (2016), the influence of non-tax units on the Gini coefficient in Switzerland amounts to 0.12 points. This issue is also relevant for historical studies when tax systems were less developed and fewer people were required to pay taxes.

Second, one earner may file more than one tax return. In Poland, for instance, entrepreneurs choosing a flat income tax rate (PIT-36L) often submit a standard tax return for much lower ancillary income (PIT-36). If not adjusted, this can create the impression of a significant number of low-income earners.

Third, if a household consists of one high-income tax filer and two to three non-earners or non-tax filers, analyses based on tax data tend to overestimate measured inequalities. According to Hümbelin and Farys (2016), the Gini coefficient for tax units in Switzerland is 0.06 points higher than that for households (0.45 vs 0.39).

3. Misreporting of taxable income

Misreporting of taxable income is another critical issue. Both nonreporting and underreporting can introduce significant biases, particularly at the upper and lower ends of the income distribution (Medalia et al., 2019). For middle-income earners, income is often derived from employment where taxes are deducted at the source, minimizing evasion opportunities (Bukowski et al., 2023). However, at the upper end, business profits dominate and are more susceptible to various forms of tax avoidance and evasion. This includes unregistered sales, overreported costs, and shifting profits to lower-tax jurisdictions.

Despite numerous studies on the 'grey economy', some income inequality researchers using tax data have ignored this effect. Taxable income at the upper end is often higher than disposable income reported in surveys, leading researchers to assume that tax data are closer to actual income (Hümbelin & Farys, 2016). As Carranza et al. (2022) state, 'in all cases, tax data can be at least considered a reliable lower bound of the upper tail of the distribution' (p. 5). However, this assumption may not hold true, especially at the lower end where under-the-table income from unregistered employment is prevalent. In Poland, unregistered employment is estimated to account for 24–34% of the 'grey economy' (Pasternak-Malicka, 2019).

Researchers use various strategies to mitigate biases in inequality measurement. One approach is adjusting the highest income distribution decile based entirely on survey data (Jordá & Niño-Zarazúa, 2019). Another is combining tax data with household surveys to enhance the accuracy of information on high-income earners (Bartels & Metzing, 2019; Burdín et al., 2022; Medalia et al., 2019). Macro data from national accounts can also be used to reconcile discrepancies between tax or survey data and macroeconomic indicators (Blanchet et al., 2022; Bukowski et al., 2023; Zwijnenburg, 2022). Additionally, constructing fiscal income from surveys, rather than the other way around, can help align tax units with households (Yonzan et al., 2022). Surveys provide detailed information on household composition and income sources, allowing for assumptions about who files taxes within the household and adjustments to the definition of income accordingly.

III. PECULIARITIES OF THE POLISH INCOME TAX SYSTEM

The Polish income tax system has several unique characteristics that affect the measurement of disposable income inequality. This section discusses three main peculiarities: the dominance of business income among the wealthiest taxpayers, the treatment of retained earnings in pass-through entities, and the tax wedge between employees and the self-employed.

1. Entrepreneurs as the wealthiest taxpayers

In Poland, the wealthiest individuals are predominantly business owners. Most of the richest are entrepreneurs actively involved in managing their private, usually non-public, companies (Bukowski et al., 2023; Bukowski & Novokmet, 2021). The share of business income among the top 1% earners has been steadily growing since the beginning of the economic transformation (1990), reaching almost two-thirds of all income, while within the 5-1 upper percentiles, non-business incomes (labour, pensions, and others) still dominate (Bukowski et al., 2023; Bukowski & Novokmet, 2021).

According to tax data, those who declare business income (at least after 2004) usually choose the flat tax rate (PIT-36L). This tax regime may not be applied to dividends, but to some extent may be applied to capital gains. In 2020, 763,681 taxpayers submitted PIT-36L, declaring an aggregate taxable income of PLN212,899 million. Those who decided to tax their business income with progressive tax rates declared only PLN40,173 million. The average income declared by PIT-36L taxpayers was PLN279,200. Examining the sources of declared income more closely, typical passive income does not constitute a significant amount, either in terms of the number of taxpayers or taxable income. In 2020, PIT-38 (mostly capital gains from the disposition of financial assets) was submitted by 379,022 taxpayers (189,849 revealed taxable income) with an average income of PLN35,541. PIT-39 (capital gains from real estate disposition) was submitted by 69,525 taxpayers with an average income of PLN89,800 (Ministerstwo Finansów, 2022). Another report by the same group of researchers confirmed that 'In the Polish case, the addition of the dividend has a negligible effect on the level and evolution of income inequality' (Bukowski et al., 2023, p. 9).

The dominance of entrepreneurs taxed transparently (pass-through taxation) under personal income tax (PIT) instead of corporate income tax (CIT) is also confirmed by national statistics. Of Poland's 2,261,856 non-financial enterprises, 87% are natural-person businesses (including partnerships of natural persons).

2. Retained earnings in pass-through entities and CIT payers

We have emphasized the crucial distinction between disposable income and retained earnings. Once a researcher defines the type of income to be con-

sidered for measuring inequality, they must use this definition consistently. This means that, for economically similar sources of income, the same type of income should be included in the measure. Maintaining this consistency can be challenging when using tax data to analyse business income, primarily due to the different taxation rules and tax bases for pass-through entities versus corporate taxpayers.

Business owners have considerable flexibility in choosing the most advantageous form of taxation, unlike employees, whose form of employment (labour contract, service agreement, or self-employment) is typically determined by their employer. For small and medium enterprises (SMEs), the choice of legal form is often driven by tax considerations, with other factors being secondary. Kopczuk rightly noted that in the USA, following the 1986 Reagan tax reform, most SMEs converted to S-corporations, which are tax-transparent and face lighter cumulative tax burdens compared to C-corporations (Kopczuk & Zwick, 2020). C-corporations are subject to 'double taxation', where business income is first taxed at the corporate level (both federal and state) and then taxed again on dividends distributed to shareholders. In most inequality calculations, only profits distributed to natural persons are considered. Consequently, due to the tax reform, corporate income increasingly became personal income. Without significant changes in overall income, natural persons and business owners began to report significantly higher taxable income, which was mistakenly interpreted as a rapid income rise among top earners.

A similar situation occurred in Poland after the 2004 tax reform and the introduction of a 19% flat tax rate for entrepreneurs (PIT-36L; Kopczuk, 2012). Starting that year, two trends became apparent. Firstly, there was an increase in the number of sole proprietors, which likely included some hidden employment of the wealthiest earners. Between 2012 and 2021, the number of sole proprietors rose from 2,917,272 to 3,551,893. Secondly, there was a noticeable shift from corporate income tax (CIT) to personal income tax (PIT), especially among the most profitable enterprises. This transition is relatively straightforward; in the USA, it requires checking a box on the annual tax return, while in Poland, it involves a formal restructuring procedure under company law or transferring the business title between different companies.

However, the second trend is less evident. During the investigated period, the Polish economy experienced almost uninterrupted dynamic growth, reflected in the growing number of businesses in all legal forms. The choice between a tax-transparent and a non-tax-transparent entity depends on various ancillary circumstances. For instance, the largest enterprises listed on the stock exchange, as well as banks, or insurers, are unable to choose due to regulatory requirements. Enterprises intending to invite institutional or financial investors or those engaged in high-risk businesses (like start-ups or heavily leveraged businesses) avoid tax-transparent forms due to the unlimited liability of members. Similarly, businesses that do not plan to distribute profits in the foreseeable future lack any incentive to use PIT-36L.

Conversely, privately owned and highly profitable businesses were particularly responsive to tax reform changes. This shift is evident through a coun-

terfactual comparison. Figure 1 shows the actual and counterfactual tax revenues from PIT-36L and CIT from 2008–2021. The counterfactual inflows represent what revenues would have been if the ratio of both tax streams had remained fixed since 2008. The gap between the actual and counterfactual curves represents the rising revenue from PIT-36L at the expense of CIT, further illustrating the taxpayers' shift from corporate to personal taxation.

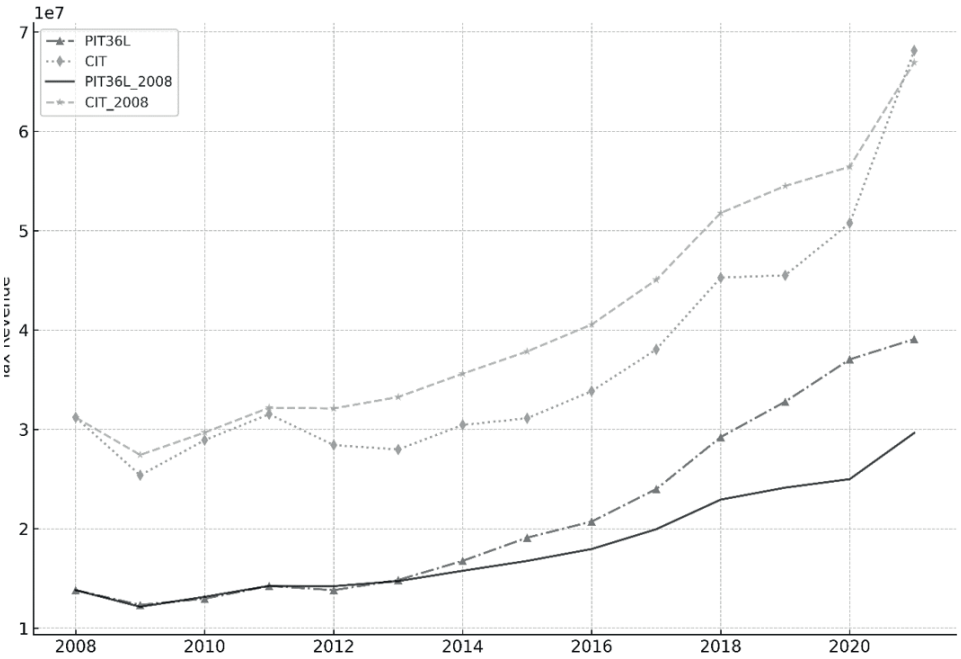
This shift can also be observed by examining the changing numbers of specific company types, such as limited partnerships (*spółka komandytowa*) and joint-stock partnerships (*spółka komandytowo-akcyjna*). Both structures are rather complex and would rarely be chosen for business purposes if not for particular tax incentives. For a time, they were considered tax-transparent and, if properly structured, provided limited liability to the partners. Thus, they became especially popular among medium and large highly profitable enterprises owned by a small group of entrepreneurs. Following the 2004 tax reform, the number of such companies increased. For joint-stock partnerships, the peak was in 2015 (4,566 companies) when they became subject to corporate taxation. Since then, their numbers declined until 2021 (2,850 companies), when they rose again, likely due to a flaw in the 2021 tax reform that exempted partners of joint-stock partnerships from the health care insurance contribution. The number of limited partnerships grew steadily from 17,976 companies in 2015 to 40,577 companies in 2020 but began to decline after 2020 due to their subjection to corporate taxation.

The impact on inequality calculations based on tax data is considerable. Our analysis demonstrates that shifts in the concentration of high-income taxpayers have a pronounced effect on inequality measures. For instance, doubling either the number of DSF-1 (solidarity tax) filers or their average income results in an increase of the Gini coefficient by 0.04 points, while doubling both simultaneously leads to a 0.11-point rise (see Table A6). This experiment demonstrates that the Gini coefficient is highly sensitive to top-income shares, highlighting the need for careful adjustments in inequality studies using tax data.

Suppose two-thirds of the wealthiest taxpayers are entrepreneurs, deriving most of their income from entrepreneurial activities carried out under tax-transparent entities. In that case, all their income, regardless of its disposition, is considered disposable. If, counterfactually, they operated under non-tax-transparent entities (like *spółka z ograniczoną odpowiedzialnością* or *spółka akcyjna* in Poland), only distributed income would be considered disposable. Profits retained in businesses for further investments would increase their wealth but not their annual income. This effect can be estimated using macro-data from national statistical accounts. Among enterprises with more than ten employees in the period 2020–2021, almost 70% of net income remained undistributed. For small enterprises, investments in tangibles and intangibles constituted about a quarter of net profit (Tables A2, A3). This difference is easily explained by the fact that many small entrepreneurs are self-employed, providing personal services with minimal investment. The conclusion is clear: the larger the enterprise, the greater the share of net profits that remains undistributed and should not be counted as disposable income.

Figure 1

Trends in PIT-36L and CIT revenues: actual vs counterfactual values



Note. Actual CIT tax revenues (dotted line with diamonds) are compared to counterfactual revenues based on CIT's share of total taxes in 2008 (dashed line with star). Actual revenues from PIT-36L (dashed line with triangles) are compared to counterfactual revenues based on PIT-36L's share of total taxes in 2008 (solid line). For details of the calculations, see Table A1.

Source: the authors' own calculations based on the Ministry of Finances' statistics on tax revenues.

It is worth noting that the shift from corporate to personal taxation among entrepreneurs is beginning to reverse in response to recent changes in tax law. Since 2017, lawmakers have incentivized corporate taxation over a flat tax rate by reducing CIT for small companies to 15% and to 9% in 2019. In 2021, the so-called Estonian CIT was introduced, allowing the deferral of tax payments until profits are distributed. As illustrated in Figure 2, the actual and counterfactual curves converge after 2020, reflecting this immediate impact.

3. Tax wedge and over-taxation of business income

However, the problem with the disposable income of the wealthiest extends further. A well-known flaw of the Polish tax system is the significant difference in the so-called tax wedge between salaried employees and the self-employed. The tax wedge, which is the difference between the overall employment costs and an employee's net income, is about one-third lower for entre-

preneurs, even after the 2021 tax reform aimed at reducing this disparity (Sawulski et al., 2023, p. 8). The tax wedge also decreases with rising business income due to the lump-sum social security contribution for entrepreneurs and flat tax rate. Consequently, this creates a strong incentive for illegal employment or at least partially compensating workers 'off the books'.

It is often assumed that unregistered compensation comes from the 'grey economy', meaning it is entirely untaxed. However, it is reasonable to argue that a significant portion of that income is not completely untaxed. Due to stringent VAT regulations in the EU, recent efforts to diminish the VAT gap, and the growing institutionalization of the economy (Bednarski, 2016), it is much harder for registered entrepreneurs to hide VAT-taxable proceeds than to misreport their taxable income and illegally compensate employees with already taxed funds. This scheme leads to an intriguing hypothesis. When comparing business income declared in questionnaires and tax returns, most researchers assume a natural tendency to underreport profits and underpay taxes due to evasion, avoidance, or optimization. To our knowledge, no research in Poland has explored the possibility that the income of the wealthiest entrepreneurs might not be undertaxed but rather overtaxed.

To substantiate this claim, we can examine several estimations. For instance, Szulc (2022) observed a noticeable underestimation of low incomes in both household surveys and tax data, as both sets fail to reveal significant discrepancies at the lower end of the income distribution (Kośny, 2019). The underestimation becomes apparent when incomes are compared with declared consumption, or more specifically, the multidimensional household well-being indicator. The curve presented by Szulc (2022, p. 83) showing the dependency of multidimensional poverty on income takes a counterintuitive shape, with poverty rising with income up to PLN400 and falling thereafter.

Szulc then replaced the income of the poorest with predictions based on a large set of well-being correlates using hierarchical correlation reconstruction. The results were notable: the mean equivalent income of the bottom 20% rose by 65% in 2014 and 63% in 2015, up to PLN1189.[1] Based on these estimations, we extrapolated the effect on the entire population, with precise calculations provided in Table A4. We concluded that the total underestimated grossed-up income of the lower quintile amounted to PLN52,526.64 M in 2015, representing 47% of the entire income declared by entrepreneurs in PIT-36L tax returns in 2015 (PLN110,495 M).

A similar calculation was made by Kiełczewska et al. (2021), who estimated the annual amount of under-the-table compensation paid to employees.

[1] It is important to note that Szulc's (2022) correction of low-income data, while insightful, remains somewhat speculative. The correction assumes that the discrepancy between survey income and multidimensional poverty indexes at the lower end reflects shadow economy income. However, shadow economy income likely impacts the entire income distribution, albeit with unknown intensity. Furthermore, it is reasonable to assume that high-income earners may underreport income due to evasion or avoidance practices, which could lead to an underestimation of overall income inequality. These factors underscore the complexity of accurately adjusting income measures in the absence of comprehensive administrative data.

They considered only the extra compensation paid to registered employees, presumably untaxed, excluding compensation to unregistered employees. The methodology involved comparing the consumption expenditures and declared income of employees in private enterprises versus state or state-owned entities, which do not engage in under-the-table payments. The comparison revealed almost no differences in consumption expenditures but significant differences in declared taxable income, estimating a gap of PLN34 billion in 2018. This gap aligns with our previous figures when considering that unregistered employment income was not included.

We lack a reliable instrument to calculate how much of this income was overtaxed. It could be argued that all this income originates from the grey economy and does not reflect overtaxed income. To verify this hypothesis, we compared the profitability of companies using standard double-entry bookkeeping with those using a simplified registry of proceeds and expenditures (*książka przychodów i rozchodów*). Over-taxation and under-the-table payments are difficult to conceal with double-entry bookkeeping but easier with simplified bookkeeping. Simplified bookkeeping users should be more prone to unregistered proceeds. If the grey economy hypothesis is correct, those who use simplified bookkeeping should show lower profitability than standard accountancy users. However, in 2021, simplified bookkeeping users reported nearly twice the pre-tax profit (15.21% vs 6.51%). A similar profitability gap can be observed among PIT-36L taxpayers, with tax profitability at 16.4%, compared to CIT payers at 6.13%. This suggests the over-taxation hypothesis is more plausible.

Part of these extra profits likely reflects the consumption of business owners. Among CIT-payers, this consumed income is often hidden as directors' fees, reducing the tax base, while among PIT-36L payers, it is part of gross profits. To estimate over-taxed income, we consider the consumable income of business owners. Assuming an average business owner consumes the equivalent of the average income of affluent PIT-36 payers (those earning over the second tax threshold of PLN85,528), which for 2021 is PLN137,913 annually, there remains an overtaxed gap of PLN23.54 billion. Detailed computations are presented in the Appendix (Table A5).

IV. CRITIQUE OF POLISH INEQUALITY STUDIES BASED ON TAX DATA

Many researchers investigating income distribution in Poland have turned to tax data, considering it more reliable than survey data. However, some of these studies do not adequately adjust for potential biases. For instance, Kośny (2019) compared survey data and income statistics from the Dolnośląskie province without any corrections. He concluded there was substantial underestimation of income in the upper tail of the distribution. He rightly observed: 'Any division of household income among individuals (based on household

composition) would require estimates on the number of people earning income subject to personal income tax. It also needs assumptions about intra-household income distribution, which usually depends on income level and is very asymmetric in the case of households with high income' (p. 64). Despite this observation, no adjustment was made. He also noted that in the highest areas of income distribution, incomes according to the tax records were, on average, more than twice as high as in the survey. He further concluded that 'the only explanation for these differences is the very high underestimation of the income from the highest areas of income distribution in the HBS data' (p. 70). However, other explanations are possible. The simplest lies in the questionnaire used in the Household Budget Survey (HBS) in Poland. Statistics Poland (Główny Urząd Statystyczny [GUS]) does not ask respondents about their entire business income but only about the part of business income devoted to personal consumption and non-business investments (residential buildings, summer houses, or private vehicles).

The problem of comparing tax units and households also appears in Bukowski and Novokmet (2021). They invested considerable effort in constructing a unified measure of long-run income distribution; however, in our view, they failed to avoid certain tax-related pitfalls. For the most recent period under study, when combining tax and survey data, they assumed that high-income individuals were either single or that spouses reported the same income since, in their case, spouses usually took advantage of joint filing for couples. This may have been true before the 2004 tax reform, which introduced a flat tax rate for entrepreneurs, excluding them from the joint filing tax incentive.

A similar conclusion about the underestimation of survey income was drawn by Brzeziński et al. (2021) and Brzeziński and Sałach (2021). They combined survey and tax data and imputed the highest incomes according to the Pareto model, both in terms of income and the highest earners population. However, they did not make any corrections concerning disposable income.

The most sophisticated model, based on tax income but with many adjustments from surveys and national accounts, was developed by Bukowski et al. (2023). They started with personal income tax (PIT) returns and made several corrections, including adjustments for tax filers versus households, imputation of agricultural income, dividends, interest, and income of entrepreneurs subject to presumptive taxation. For the latter, they assumed high profitability for the taxpayers (75% for services and 90% for rents), which is justified for these categories of income but overly optimistic for small traders and craftsmen, who were disregarded. They also decided to exclude capital gains, income from imputed rents of homeowners, retained business income, and paid corporate income taxes. The latter exclusion would be valid, provided that retained earnings were also excluded from business income taxed as personal income, but they were not. On the other hand, they decided to impute dividend income. Dividends are taxed at the source by the company distributing profits without specifying which stockholders received the dividend. They concluded that 'it has a negligible effect on the level and evolution of income inequality' (p. 10). This is largely accurate, as dividend income is interchange-

able with retained earnings and capital gains from selling company shares. Given Poland's double taxation of distributed corporate profits, this option is rarely chosen due to the high tax burden.

Bukowski & Novokmet (2021) also excluded capital gains, noting that income tax applies only to realized capital gains, which may lead to unintended volatility. However, capital gains earned in pass-through entities are taxed as business profits. Excluding this source of income should have been done consistently, but it was not. In an earlier study, they also disregarded retained earnings, claiming that 'correcting income distribution by imputing corporate retained earnings is less important in Poland because business income is predominantly taxed according to the pass-through concept and hence attributed to individuals' (p. 200). This claim is inaccurate. Comparing the tax units in PIT-36L and CIT, the conclusion holds true, but not in terms of taxed income. Table A2 presents income tax inflows from 2008–2021, distinguishing between PIT-36L (individual entrepreneurs and pass-through entities) and CIT (corporations). Each year, CIT inflows are higher.

Piketty et al. (2018) also addressed retained earnings and capital gains, opting for imputation based on dividend distribution. However, this method is questioned. Kopczuk and Zwick (2020) noted that companies distributing dividends usually have lower retained earnings. Atkinson et al. (2011) advocate including both income categories in inequality measurements, assuming that lifetime differences eventually cancel out.

Including retained earnings reveals another issue: unrealized capital gains. If they are considered, the unrealized capital gains of individuals should also be included. This poses a significant problem in countries where middle-income earners invest heavily in appreciating residential properties, including Poland. None of the cited Polish researchers took this effect into account, which is appropriate when calculating disposable income but problematic when including the retained earnings of pass-through entities.

Historical studies pose additional problems. Bukowski and Novokmet (2021) conducted an insightful study of income inequality in Poland over a century, based on tax data from income tax statistics combined with surveys and national accounts. They made many assumptions and adjustments to construct a unified measure of long-run income distribution measure. However, making data from different periods and tax systems comparable is extremely difficult.

For example, in Poland in the 1920s, only 10% of the professionally active population were income taxpayers (Górski, 2018). It is problematic to draw reliable conclusions about income inequality based on a small fraction of the population, especially considering that about 60% of the total population depended on agriculture for their livelihood. Income from agriculture was largely exempt from income tax due to various exemptions and deductions but was instead subject to a chaotic tax system inherited from the foreign powers that had occupied Poland before 1918. The authors did not incorporate data from agricultural tax records into account. Those dependent on agriculture included both farmers (about 50%), who could be extremely poor or relatively

wealthy, and extremely wealthy large landowners (0.3%). Even if their income was subject to tax, it was widely acknowledged in the literature that the declared income was very loosely related to actual disposable income (Ambroch et al., 2018; Ogrodnik, 1977/78).

A similar objection applies to studies of communist Poland after 1945. Bukowski and Novokmet (2021) used tax data from statistics covering payroll taxes paid by workers in the socialized sector, which they claimed accounted for the majority of the labour force. However, in 1960, about 40% of the working population was still employed in agriculture. Bukowski and Novokmet imputed income from private agriculture to the bottom bracket, assuming low productivity and small earning potential did not contribute to top incomes. This attribution is flawed. Despite low productivity, scarcity of produce on the market led to the high profitability of private farmland, especially when sold directly to consumers. They also excluded private craftsmen, who represented the most affluent part of society after 1956. These craftsmen were covered under 'unearned sources of income' and significantly grew in number between 1950–1988 (Leszczyńska, 2018, p. 109). In 1980, the private sector's share of GDP reached 17.5% (Kaliński, 1995).

V. CONCLUSIONS

To accurately study income inequality, it is more effective to use survey-based income data adjusted with administrative data to ensure representation of all income levels. Unadjusted tax data can lead to significant miscalculations (Hümbelin & Farys, 2016).

Correcting tax data to represent real disposable income is a common pitfall for researchers who are not tax specialists. While many of these challenges have been addressed in the vast literature on the topic and most research applies the necessary solutions, some issues continue to be misinterpreted:

First, there is a crucial distinction between disposable income and total income (including retained earnings and unrealized capital gains). Researchers must treat all possible sources of income consistently. Including the retained earnings of pass-through entities while ignoring those of corporations will significantly distort results, especially in the upper tail of the income distribution. If retained earnings are considered, unrealized capital gains should also be included. Holding assets as an investment with the prospect of appreciation is a source of income regardless of the asset type (stocks, shares, bonds, gold, real estate, cryptocurrencies, etc.).

Second, misreporting does not necessarily reduce the tax base or justify the assumption that higher taxable income compared to survey income is inherently more reliable. Taxpayers often adapt creatively to the tax system, sometimes deliberately overtaxing specific income sources.

Third, using tax data requires not only knowledge of the law but also an understanding of how the law is actually applied by taxpayers. Tax law rarely

describes the behaviour of its subjects. The opposite conclusion is more plausible: taxpayers will strategically adapt to the law and the practices of tax collectors to minimize their tax burden. If the law is not regularly enforced, taxpayers will not comply (as was the case in Poland in the 1920s). If the law is implemented inconsistently and arbitrarily, tax data will not be a reliable source of information (as in communist Poland with private enterprises).

In conclusion, while tax data can offer valuable insights into income inequality, its limitations necessitate careful adjustments and the complementary use of survey data to provide a more accurate picture of disposable income distribution. Neither data source alone fully accounts for the complexities of income distribution, particularly the underreporting issues present at both the lower and upper ends of the spectrum. A balanced approach that integrates both sources, where possible, may provide a more comprehensive and reliable view of income inequality.

Author contributions / Indywidualny wkład autora (CRediT): **Marcin Gorazda** – 50% (Conceptualization / Konceptualizacja; Formal analysis / Formalna analiza; Writing – original draft / Pisanie – pierwszy szkic; Writing – review & editing / Pisanie – recenzja i edycja). **Tomasz Kwarciński** – 50% (Conceptualization / Konceptualizacja; Formal analysis / Formalna analiza; Writing – original draft / Pisanie – pierwszy szkic; Writing – review & editing / Pisanie – recenzja i edycja).

Conflict of interest / Konflikt interesów: The authors declare no conflict of interest. / Autorzy nie zgłosili konfliktu interesów.

Funding / Finansowanie: The article presents the results of the Project no. 027/GPF/2024/POT financed from the subsidy granted to the Krakow University of Economics (Poland). / Artykuł powstał w ramach realizacji projektu nr 027/GPF/2024/POT finansowanego z subsydium grantowego Uniwersytetu Ekonomicznego w Krakowie.

The use of AI tools / Wykorzystanie narzędzi AI: The authors declare no use of AI tools. / Autorzy oświadczyli, że nie korzystali z narzędzi AI.

Data availability / Dostępność danych: Not applicable. / Nie dotyczy.

References / Bibliografia

Angel, S., Heuberger, R., & Lamei, N. (2019). What did you really earn last year? Explaining measurement error in survey income data. *Journal of the Royal Statistical Society*: Series A (Statistics in Society), *182*(4), 1411–1437. https://doi.org/10.1111/rssa.12463

Ambroch, M., Czernak, G., Lisiak, E. & Szydłowska, G. (2018). *100 lat Polski w liczbach. 1918–2018* [100 years of Poland in numbers: 1918–2018]. Główny Urząd Statystyczny.

Atkinson, A. B. (1975). Income distribution and social change revisited. *Journal of Social Policy*, *4*(1), 57–68. https://doi.org/10.1017/S0047279400003998

Atkinson, A. B., & Piketty, T. (Eds.). (2007). *Top incomes over the twentieth century: A contrast between continental European and English-speaking countries*. Oxford University Press.

Atkinson, A. B., Piketty, T., & Saez, E. (2011). Top incomes in the long run of history. *Journal of Economic Literature*, *49*(1), 3–71. https://doi.org/10.1257/jel.49.1.3

Bartels, C., & Metzing, M. (2019). An integrated approach for a top-corrected income distribution. *The Journal of Economic Inequality*, *17*(2), 125–143. https://doi.org/10.1007/s10888-018-9394-x

Bednarski, M. (2016). Wysokość kosztów pracy i struktura zatrudnienia w MŚP a dynamika szarej strefy w Polsce [The labour costs and the structure of employment in Polish SME's and

the grey economy]. *Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*, *292*, 20–31.

Blanchet, T., Chancel, L., & Gethin, A. (2022). Why is Europe more equal than the United States? *American Economic Journal: Applied Economics*, *14*(4), 480–518. https://doi.org/10.1257/app.20200703

Brzeziński, M., Myck, M., & Najsztub, M. (2021). Sharing the gains of transition: Evaluating changes in income inequality and redistribution in Poland using combined survey and tax return data. *European Journal of Political Economy*, September. https://doi.org/10.1016/j.ejpoleco.2021.102121

Brzeziński, M., & Sałach, K. (2021). *Nierówności dochodowe i majątkowe w Polsce: nowe wyniki wykorzystujące dane pozaankietowe* [Income and wealth inequality in Poland: New results using non-survey data]. Center for Open Science. https://doi.org/10.31219/osf.io/s43yr

Bukowski, P., Chrostek, P., Novokmet, F., & Skawiński, M. (2023). *Income inequality in the 21st century Poland*. Ministry of Finance of the Republic of Poland.

Bukowski, P., & Novokmet, F. (2021). Between communism and capitalism: long-term inequality in Poland, 1892–2015. *Journal of Economic Growth*, *26*, 187–239. https://doi.org/10.1007/s10887-021-09190-1

Burdín, G., De Rosa, M., Vigorito, A., & Vilá, J. (2022). Falling inequality and the growing capital income share: Reconciling divergent trends in survey and tax data. *World Development*, *152*, 105783. https://doi.org/10.1016/j.worlddev.2021.105783

Burkhauser, R. V., Hérault, N., Jenkins, S. P., & Wilkins, R. (2018). Top incomes and inequality in the UK: Reconciling estimates from household survey and tax return data. *Oxford Economic Papers*, *70*(2), 301–326. https://doi.org/10.1093/oep/gpx041

Carranza, R., Morgan, M., & Nolan, B. (2022). Top income adjustments and inequality: An investigation of the EU-SILC†. *Review of Income and Wealth*. https://doi.org/10.1111/roiw.12591

Central Statistical Office. (2016a). *Demographic Yearbook of Poland 2016*. Statistical Publishing Establishment.

Central Statistical Office. (2016b). *Household budget survey in 2015*. Statistical Publishing Establishment.

Fernández-Albertos, J., & Kuo, A. (2018). Income perception, information, and progressive taxation: Evidence from a survey experiment. *Political Science Research and Methods*, *6*(1), 83–110. https://doi.org/10.1017/psrm.2015.73

Frenette, M., Green, D. A., & Milligan, K. (2007). The tale of the tails: Canadian income inequality in the 1980s and 1990s. *Canadian Journal of Economics / Revue canadienne d'économique*, *40*(3), 734–764. https://doi.org/10.1111/j.1365-2966.2007.00429.x

Galbraith, J. K. (2018). Sparse, inconsistent and unreliable: Tax records and the World Inequality Report. *Development and Change*, *50*(2), 329–346. https://doi.org/10.1111/dech.12475

Główny Urząd Statystyczny. (2021). Bilansowe wyniki finansowe przedsiębiorstw niefinansowych za 2020 rok [Financial results of non-financial enterprises in 2020 (balance sheet)]. https://stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultaktualnosci/5502/9/20/1/bilansowe_wyniki_finansowe_przedsiebiorstw_niefinansowych_w_2020_r.pdf

Główny Urząd Statystyczny. (2022). Bilansowe wyniki finansowe przedsiębiorstw niefinansowych za 2021 r. [Financial results of non-financial enterprises in 2021 (balance sheet)]. https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5502/9/21/1/bilansowe_wyniki_finansowe_przedsiebiorstw_niefinansowych_w_2021.pdf

Górski, M. (2018). Niepodległość kosztuje. Początki podatku dochodowego w II Rzeczpospolitej [The cost of independence: The beginnings of income tax in Inter-war Poland]. *Miscellanea Historico-Iuridica*, *17*(1), 93–108. https://doi.org/10.15290/mhi.2018.17.01.06

Hlasny, V., & Verme, P. (2018). Top incomes and inequality measurement: A comparative analysis of correction methods using the EU SILC data. *Econometrics*, *6*(2), Article 2. https://doi.org/10.3390/econometrics6020030

Hümbelin, O., & Farys, R. (2016). The suitability of tax data to study trends in inequality – A theoretical and empirical review with tax data from Switzerland. *Research in Social Stratification and Mobility*, *44*, 136–150. https://doi.org/10.1016/j.rssm.2016.04.004

Jordá, V., & Niño-Zarazúa, M. (2019). Global inequality: How large is the effect of top incomes? *World Development*, *123*, 104593. https://doi.org/10.1016/j.worlddev.2019.06.017

Kaliński, J. (1995). *Gospodarka Polski w latach 1944–1989. Przemiany strukturalne* [The economy of Poland 1944–1989: Structural changes]. Państwowe Wydawnictwo Ekonomiczne.

Kiełczewska, A., Kośny, M., & Sawulski, J. (2021). *Skala płacenia pod stołem w Polsce* [The scale of under-the-table payments in Poland]. Polski Instytut Ekonomiczny.

Kopczuk, W. (2012). The Polish business "flat" tax and its effect on reported incomes: A Pareto improving tax reform? Columbia University Working Papers.

Kopczuk, W., & Zwick, E. (2020). Business incomes at the top. *Journal of Economic Perspectives*, *34*(4), 27–51. https://doi.org/10.1257/JEP.34.4.27

Kośny, M. (2019). Upper tail of the income distribution in tax records and survey data. Evidence from Poland. *Argumenta Oeconomica*, *42*(1), 55–80. https://doi.org/10.15611/aoe.2019.1.03

Korinek, A., Mistiaen, J. A., & Ravallio, M. (2006). Survey non-response and the distribution of income. *The Journal of Economic Inequality*, April, 33–55. https://doi.org/10.1007/s10888-005-1089-4

Kuznets, S. (1955). Economic growth and income inequality. *The American Economic Review*, *45*(1), 1–28. https://www.jstor.org/stable/1811581

Leszczyńska, C. (2018). *Historia Polski w liczbach: Tom 5. Polska 1918–2018* [History of Poland in numbers: Volume 5. Poland 1918–2018]. Zakład Wydawnictw Statystycznych.

Lustig, N. (2019). The "missing rich" in household surveys: Causes and correction approaches [Preprint]. SocArXiv. https://doi.org/10.31235/osf.io/j23pn

Medalia, C., Meyer, B. D., O'Hara, A., & Wu, D. (2019). Linking survey and administrative data to measure income, inequality, and mobility. *International Journal of Population Data Science*, *4*(1). https://doi.org/10.23889/ijpds.v4i1.939

Ministerstwo Finansów. (2022). Informacja dot. rozliczenia podatku dochodowego od osób fizycznych za 2020 rok [Information on tax-residence: Who qualifies as tax resident and under what conditions]. Departament Podatków MF.

Ogrodnik, S. (1977/78). Ewolucja opodatkowania rolnictwa w Polsce do 1971 r. [Evolution of agriculture taxation in Poland until 1971]. *Annales Universitatis Mariae Curie-Sklodowska. Sectio H*, 11/12I(15), 277–308.

Okun, A. M. (2015). *Equality and efficiency: The big tradeoff*. Brookings Institution Press.

Ostry, J., Loungani, P., & Berg, A., (2019). *Confronting inequality: How societies can choose inclusive growth*. Columbia University Press.

Pasternak-Malicka, M. (2019). Szara strefa – definicje, przyczyny, szacunki. Polska perspektywa zjawiska [Grey economy: Definitions, causes, and size]. *Studia BAS*, *2*(58), 29–56. http://dx.doi.org/10.31268/StudiaBAS.2019.11

Piketty, T. (2017). *Capital in the twenty-first century*. Belknap Press; Harvard University Press.

Sawulski, J., Szewczyk, N., Rafalska, K., & Smółko, M. (2023). Business power against redistribution: The case of watered-down tax reform in Poland. *Business and Politics*, *25*(2), 133–151. https://doi.org/10.1017/bap.2023.3

Statistics Poland. (2022). *Activity of non-financial enterprises in 2021*. https://stat.gov.pl/download/gfx/portalinformacyjny/en/defaultaktualnosci/3317/1/16/1/activity_of_non-financial_enterprises_in_2021.pdf

Stoetzer, L. F., Giesecke, J., & Klüver, H. (2023). How does income inequality affect the support for populist parties? *Journal of European Public Policy*, *30*(1), 1–20. https://doi.org/10.1080/13501763.2021.1981981

Szulc, A. (2022). Polish inequality statistics reconsidered: Are the poor really that poor? *Statistics in Transition*, September, *23*(3), 79–94. http://dx.doi.org/10.2478/stattrans-2022-0031

Titmuss, R. M. (1962). Income distribution and social change: A critical study in British Statistics. George Allen & Unwin.

Yonzan, N., Milanovic, B., Morelli, S., & Gornick, J. (2022). Drawing a line: Comparing the estimation of top incomes between tax data and household survey data. *Journal of Economic Inequality*, *20*(1), 67–95. https://doi.org/10.1007/s10888-021-09515-5

Zwijnenburg, J. (2022). The use of distributional national accounts in better capturing the top tail of the distribution. *The Journal of Economic Inequality*, *20*(1), 245–254. https://doi.org/10.1007/s10888-022-09534-w

APPENDIX

Table A1

Counterfactual PIT36L and CIT

Year	PIT36 1st tax bracket	PIT36 2nd tax bracket	PIT36L	CIT	PIT36 share	PIT36L share	CIT share	Counterfactual PIT36L 2008	Counterfactual CIT 2008
2008	34,008,599	11,042,290	13,816,000	31,197,708	0.5	0.15	0.35	13,816,000	31,197,708
2009	31,924,744	9,550,407	12,323,000	25,375,825	0.52	0.16	0.32	12,145,367.76	27,425,277.71
2010	33,868,987	9,937,362	12,960,000	28,916,177	0.51	0.15	0.34	13,143,785.89	29,679,791.13
2011	35,942,735	11,179,060	14,239,000	31,524,447	0.51	0.15	0.34	14,248,689.79	32,174,758.5
2012	37,042,284	13,393,708	13,817,017	28,417,389	0.54	0.15	0.31	14,215,732.5	32,100,338.13
2013	40,039,109	13,148,434	14,835,795	27,969,249	0.55	0.15	0.29	14,725,359.64	33,251,119.74
2014	41,368,389	14,198,402	16,765,928	30,444,400	0.54	0.16	0.3	15,766,113.69	35,601,231.28
2015	43,752,456	15,275,493	19,100,624	31,113,309	0.54	0.17	0.28	16,757,814.86	37,840,577.2
2016	46,186,833	16,322,664	20,719,382	33,845,088	0.53	0.18	0.29	17,959,264.59	40,553,553.32
2017	49,169,980	18,846,547	23,989,854	38,039,314	0.52	0.18	0.29	19,949,140.75	45,046,863.63
2018	52,715,009	22,266,794	29,221,000	45,268,463	0.5	0.2	0.3	22,929,042.93	51,775,737.26
2019	53,652,860	25,385,486	32,787,000	45,502,102	0.5	0.21	0.29	24,134,189.17	54,497,060.41
2020	48,774,555	26,332,300	37,042,000	50,756,446	0.46	0.23	0.31	24,989,837.45	56,429,187.29
2021	51,741,777	34,267,644	39,079,000	68,132,099	0.45	0.2	0.35	29,640,222.61	66,930,154.17

Source: the authors' own calculation based on information published by the Polish Ministry of Finance and Statistics Poland, https://stat.gov.pl/obszary-tematyczne/podmioty-gospodarcze-wyniki-finansowe/zmiany-strukturalne-grup-podmiotow/miesieczna-informacja-o-podmiotach-gospodarki-narodowej-w-rejestrze-regon--luty-2023,4,69.html

Table A2

Estimation of undistributed profits of SMEs (Change in Equity) and investments in micro-enterprises: non-financial enterprises (> 10 Employees) (Millions PLN)

Year	Gross income	Net income	Tax	Tax base PIT36L+CIT	Tax PIT36L+ CIT	Equity	Change in equity	Change in equity/net income (%)
2020	163,400	132,500	30,900	662,451	87,798	1,618,700	–	–
2021	298,400	250,200	48,200	662,451	107,211	1,792,400	173,700	69

Note: It should be emphasized that the year-to-year change in equity is only an approximation of undistributed profits. Equity may also change due to incoming contributions. However, as long as these contributions have a national origin, they also represent the non-consumed profits of local taxpayers. Only direct foreign investments might disrupt the results.

Source: the authors' own calculation based on statistics published by Statistics Poland 'Activity of non-financial enterprises' (GUS, 2020, 2021).

Table A3

Estimation of undistributed profits of SMEs (Change in Equity) and investments in micro-enterprises: non-financial enterprises (Millions PLN)

Year	Gross income	Gross income (PIT)	Investments, tangibles and intangibles (PIT)	Tax base PIT-36L	Net income PIT/ investments (%)
2020	517,170	205,742	41,380	212,899	24.83
2021	647,802	215,369	45,308	228,957	25.97

Source: the authors' own calculation based on statistics published by Statistics Poland 'Activity of non-financial enterprises' (GUS, 2020, 2021).

Table A4

Calculation of undertaxed income of the bottom 20% in 2015

Description	Raw survey (PLN)	Corrected (Szulc) (PLN)
Average per capita monthly income of bottom 20%*	764.21 / 728	1,189
Population of Poland on 30.06.2015	38,455,000	–
20% of the population	7,691,000	–
Per capita difference in monthly income estimation ('corrected' less 'raw survey')**	–	461
Underestimated overall monthly net income (= 20% of population x per capita difference)	–	3,545,551,000
Underestimated overall annual net income	–	42,546,612,000
Underestimated overall annual grossed-up income (tax rate 19%)	–	52,526,681,481

Notes: * Szulc (2022) assumes PLN728. However, according to the tables of Statistics Poland (GUS), the value is PLN764,21. We do not know Szulc's sources. ** In this calculation, we took Szulc's value of PLN728.

Source: the authors' own calculation based on Szulc (2022), Central Statistical Office (2016a, 2016b).

Table A5

Profitability of Entrepreneurs (non-financial companies 2021), standard double-entry bookkeeping users vs simplified bookkeeping users, CIT payers vs PIT payers (2021)

Category	No of entities	SME (10-49) employees	Proceeds (bn PLN)	Costs (bn PLN)	Profits/taxable income (bn PLN)	Profitability/tax profitability (profit/proceeds) (%)
Standard double-entry bookkeeping (>10 employees)	48,665	64.20%	4,584.4	4,286	298.4	6.51
Simplified registry of proceeds and expenditures (> 10 employees)	17,038	98.50%	71	60.2	10.8	15.21
CIT payers	606,493		7,787.8	7,310.1	477.7	6.13
PIT36L payers	801,432		1,387.7	1,163.3	224.4	16.17

Source: the authors' own calculation based on statistics published by Statistics Poland 'Activity of non-financial enterprises' (GUS, 2021), and information published by the Polish Ministry of Finance.

Table A6

2020 GINI coefficient for Poland

Variants	GINI 2020
Baseline	0.31
Doubling DSF1 taxpayers	0.35
Doubling DSF1 income	0.35
Doubling taxpayers & income DSF1	0.42

Source: the authors' own calculations based on the number of DSF taxpayers in 2020 published by DGP. The DescTools R-package is used in the calculations.