GOSPODARKA NARODOWA

1(313)2023, 41–65
DOI: 10.33119/GN/159031
gnpje.sgh.waw.pl
SGH

The Polish Journal of Economics

Istnieje od / Published since 1931

Maciej Dudek
University of Michigan, Ann Arbor, MI, United States

Paweł Dudek
Schools Complex in Bychawa, Poland

Konrad Walczyk
(*corresponding author*)
SGH Warsaw School of Economics, Poland

Optimal Labour Income Taxation in Poland: The Case of High-Income Earners*

Optymalne opodatkowanie dochodów z pracy w Polsce – przypadek osób o wysokich dochodach

Abstract

In this paper, we use actual data provided by the Polish tax authority and characterise the properties of income distribution in Poland in the case of high-income earners. By employing a variety of techniques we are able to confirm that the distribution of income in Poland can be approximated with a Pareto distribution in its upper tail. This finding makes the formula for the optimal marginal tax rates of Saez [2001] applicable to the Polish case and allows us to provide estimates of the optimal marginal tax rates for Poland. We show that the current tax policy in Poland is not optimal. Specifically, we show, by relying on empirically viable estimates of the elasticity of the labour supply with respect to the wage, that the optimal marginal tax rate at high income levels exceeds 60%. In other words, we suggest that there is room for a welfare-improving reform in Poland, and we argue that high-income individuals should be expected to contribute substantially more at the margin than they currently do.

Keywords:
social welfare, economic efficiency, Pareto distribution, optimal taxation

JEL classification codes:
H21, H24

Article history:
submitted: March 6, 2022
revised: September 12, 2022
accepted: January 4, 2023

Streszczenie

Na podstawie analizy danych pochodzących z zeznań podatkowych można dowieść, że faktyczny rozkład dochodów osób o wysokich dochodach w Polsce może być dostatecznie dobrze przybliżony przez rozkład Pareto. Oznacza to, że do wyznaczenia optymalnej, krańcowej stopy podatku dochodowego można użyć formuły Saeza [2001]. Z analizy wynika, że obecna taryfa opodatkowania wynagrodzeń w Polsce nie jest optymalna. Jeśli uznać znane szacunki płacowej elastyczności podaży pracy za wiarygodne, optymalna stawka podatku od dochodów z pracy

Słowa kluczowe:
dobrobyt społeczny, efektywność ekonomiczna, rozkład Pareto, podatek optymalny

Kody klasyfikacji JEL:
H21, H24

Historia artykułu:
nadesłany: 6 marca 2022 r.
poprawiony: 12 września 2022 r.
zaakceptowany: 4 stycznia 2023 r.

* The results presented in this paper were obtained as a part of a research project generously funded by Poland's National Science Center: Grant #UMO-2016/21/B/HS4/03055. We would like to express our deepest gratitude to the National Science Center for its funding, which made the completion of our research on optimal labour income taxation possible. In addition, we would like to thank the authorities of the Vistula University for offering us an opportunity to carry out our research at the university and to Marta Woźniak, Agata Papudzińska, Beata Gałecka, Prof. Marta Götz, and Małgorzata Wieteska-Rostek for the professional assistance they offered to us at all stages during this project. Finally, we would like to thank the Polish tax authority for sharing actual Polish data with us.

powinna przekroczyć 60%. Inaczej mówiąc, istnieje pole do poprawy efektywności systemu opodatkowania wynagrodzeń w Polsce poprzez podwyższenie krańcowej stopy podatku dla osób o najwyższych dochodach.

Introduction

Finding the optimal marginal tax rates is not easy as the original problem of **Mirrlees** [**1971**] is technically demanding and its calibration requires rich data sets. Consequently, in practice, governments almost never resort to optimal tax schedules and rely on simplistic and traditional tax functions that are far from welfare maximising. In this paper, we try to bridge the gap between economic theory and practice in the case of Poland. Specifically, we use actual empirical data provided by the Polish tax authority to attack the optimal taxation problem in the range of high-income levels. In particular, we were able to establish that that the formula derived by **Saez** [**2001**] can be applied to the Polish data, and we successfully used Saez's formula to identify the optimal tax rates applicable to high-income earners in Poland. Not surprisingly, our results show that the current tax system in Poland is not optimal and that there is room for a welfare improvement in Poland by allowing the marginal tax rates to be increased at high income levels.

The optimal taxation problem of **Mirrlees** [**1971**], despite its elegance, has not influenced the thinking of policy makers to a large extent. Thus, it appears that, in practice, governments inadequately balance the trade-offs between efficiency and equity. This is probably due to the fact that both the informational and computational demands of the Mirrlees problem are still beyond the means of most policy designers. It is thus imperative to attempt to stimulate progress in the field of optimal taxation as the current tax schedules must lead to welfare losses. Moreover, the situation is likely to worsen in the future as revenue requirements are expected to increase in the face of rising entitlements and adverse demographic shifts. Consequently, one must come up with better tax schedules as the current practice of adjusting tax rates ad hoc can lead to huge welfare losses and can significantly compromise efficiency. Some countries have already recognised the problem possibly given the findings of **Piketty** [**2014**], and have initiated the debate on the need for a tax reform. In fact, given that the situation is expected to shift in such a dramatic manner, calls for at least partial reforms have been put forward as finding the optimal tax schedules as defined by **Mirrlees** [**1971**] can be too demanding. In this paper, we follow a similar approach and try to solve for the optimal tax function applicable on a part of the domain in the case of Poland.

It is widely known that the solution to the optimal taxation problem in general is shaped by four key parameters: the preferences of individuals, the preferences of policy makers, the distribution of skills in the economy, and the revenue requirement. In fact, the solutions happen to be sensitive to the underlying characteristics of the problem, as noted by **Tuomala** [**1990**]. Consequently, from a practical point of view, a more prudent approach may be more appropriate. In fact, as proposed by **Saez** [**2001**], assuming that the distribution of skills is unbounded and focusing on high-income earners can facilitate progress and can be a stepping stone to meaningful tax reforms.

Our approach is formal and allows us to pin down the optimal marginal tax rates in Poland for high-income earners as we utilise the ingenious approach proposed by **Saez** [**2001**], who was able to provide a simple and applicable formula for the optimal marginal tax rates at high income levels. Specifically, we build on **Dudek et al.** [**2021**], and by utilising actual Polish data we establish that the formula of **Saez** [**2001**] is in fact applicable in Poland. In particular, we show that income distribution in Poland can be approximated with a Pareto distribution with parameter m in the range [2.2,3.4] i.e., we verify that income distribution in Poland can be viewed as thick-tailed with relatively many individuals earning high incomes. In such a case, as observed by **Saez** [**2001**], the marginal tax rates at the top are not equal to 0, which happens to be the case in the typical optimal taxation problem of **Mirrlees** [**1971**]. Furthermore, we are able to apply the actual formula of **Saez**

[2001], given the Paretian form of income distribution in Poland, to pin down the optimal marginal tax rates in Poland. We show that the optimal marginal tax rates are bounded between 15% and 70% for income levels exceeding PLN 50,000 per year, the value defining the 83 percentile. The variation in the optimal marginal tax rates is mainly due to the variation in the value of the elasticity of the labour supply with respect to the wage with higher values of the elasticity leading to lower marginal tax rates.

It appears that empirically viable estimates of the elasticity of the labour supply with respect to the wage in Poland (see **Bargain et al.** [**2014**] and **Myck and Trzciński** [**2019**] or **Zawisza** [**2019**]) are, just as in the United States, low. In fact, an estimate of 0.2 can be viewed as a generous one. Under such assumptions, we can conclude that the optimal marginal tax rates at high income levels in Poland should be around 60% i.e., significantly higher than the current rates imposed in practice. Our findings suggest that a tax reform grounded in optimal tax theory would be beneficial in Poland as it would help boost welfare by allowing for a more satisfying trade-off between efficiency and equity.

The results presented in this paper stem directly from a formal application of the optimal tax theory. Still, it may appear to many readers that the numerical values obtained with the optimal tax formula of **Saez** [**2001**] and presented in this paper are extreme and, even if correct, politically infeasible. While we may privately sympathise with such impressions, we want to emphasise that similar results were obtained in the case of the United States, with the optimal marginal tax rates reaching values in the neighbourhood of 75%. Moreover, in the past, marginal tax rates in many economies were much higher than today, sometimes in excess of 90%, and many argue that such high marginal tax rates had a positive impact on welfare with a little negative impact on efficiency. It appears, however, that any reforms should be undertaken – in a synchronised manner, possibly at the EU level – as high marginal tax rates can induce relocation of higher earners, as noted by **Kleven et al.** [**2013**].

Our paper mirrors that of **Doligalski** [**2019**], who also estimates optimal marginal tax rates in the case of Poland using the fact that the optimal marginal tax rate obtained by **Saez** [**2001**] happens to be the revenue maximising rate at high income levels if the marginal social weight attached to high-income earners is negligible. In the paper, we supplement the analysis of **Doligalski** [**2019**] with a sensitivity analysis with regard to the threshold defining the level of high income and with an analysis of the impact of the marginal social weight attached to the welfare of high-income earners on the optimal tax rate.

The paper is organised in five sections. The optimal tax rate formulas are derived in section two. The distribution of income in Poland is characterised in section three. The following section contains the description of the optimal marginal tax rates applicable to the Polish case. Conclusions are included in section five.

Optimal taxation at high-income levels

The issue of taxation has been the subject of numerous studies and generated interest from both researchers and policy makers. Unfortunately, tax studies are normally reduced to partial equilibrium analysis with little thought given to anything more than a discussion of revenue needs or the efficiency properties of different tax schedules. In fact, it was not until 1971 that **Mirrlees** [**1971**] was able to formalise the problem of optimal taxation and devise proper analytical tools to state conditions that, at least implicitly, define the optimal tax schedules. Specifically, **Mirrlees** [**1971**] was able to incorporate the key informational constraint – the lack of an ability to identify individual skill levels – faced by policy makers into an otherwise standard problem involving private decision making. Such an approach allowed **Mirrlees** [**1971**] to attack the key trade-off between efficiency and equity and ultimately state conditions that pin down the optimal tax schedules.

In the Mirrlees framework, it is assumed that the preferences of all individuals are symmetric and can be represented with a utility function. In this paper, we follow the specification of utility later introduced in the work of **Diamond** [**1998**], **Saez** [**2001**], and **Kydland and Prescott** [**1982**]. In particular, we assume that individual preferences can be represented with

$$U(C,L)=C-\frac{1}{1+\frac{1}{\eta}}L^{1+\frac{1}{\eta}},$$

where C denotes consumption and L stands for units of labour supplied. Note that, in particular, the above formulation implies that we abstract from income effects. As noted by **Saez** [**2001**], income effects are likely to be small at high income levels. Moreover, it is straightforward to show that, given the above formulation, the elasticity of the labour supply with respect to the wage is constant and equal to η.

We assume that the form of utility function is known to policy makers. However, following the literature, we assume that the government is not able to observe individual choices, but it has an ability to verify the income actually earned by consumers. Furthermore, following **Mirrlees** [**1971**], we assume that individuals differ with respect to productivity and an individual whose productivity is equal to a and who spends L hours at work earns income $y=aL$. Again, we assume that the government neither observes a or L, but is able to verify income earned y. Consequently, any tax policy can only be based on the level of income earned, y, which, in particular, makes lump-sum taxation infeasible and leaves us by design in a realistic second-best world where equity-efficiency trade-offs can be studied meaningfully.

In the optimal taxation framework, individual skills are not observable. However, it is normally taken as given that policy makers are familiar with the distribution of skills in the entire economy. Here we assume that skills are distributed on interval $[a_L, a_H]$ with the corresponding *pdf* given with $f()$.

The government must take the structure of the economy as given. It is in fact restricted to a simple choice of the tax function, $\tau()$, that is conditioned by income earned y. The tax function is taken as given by consumers and each consumer makes his or her own decisions in private given their private productivity level a. Specifically, a consumer whose productivity is given with a and who earns $y=aL$ solves the following optimisation problem:

$$max_{\{y\}}U()=y-\tau(y)-\frac{1}{1+\frac{1}{\eta}}\left(\frac{y}{a}\right)^{1+\frac{1}{\eta}}.$$

The solution to the above problem is implicitly given with

$$a^{1+\eta}\left(1-\tau'\left(y_a\right)\right)^{\eta}=y_a$$

and leads to the realised level of utility equal to $U_a=y_a-\tau\left(y_a\right)-\frac{1}{1+\frac{1}{\eta}}\left(\frac{y_a}{a}\right)^{1+\frac{1}{\eta}}$.

Observe that the government does not directly control the choices of a given consumer. However, by choosing the tax function, $\tau()$, it can induce proper responses on the part of the consumer. In particular, the tax function can be chosen in such a way as to ensure that the trade-off between equity and efficiency is properly addressed under existing revenue needs. In fact, as noted by **Mirrlees** [**1971**], the optimal taxation problem can be stated as

$$max_{\{\tau()\}}W=\int_{a_L}^{a_H}G\left(U_a\right)f(a)da$$

subject to

$$i)\quad U_a=y_a-\tau\left(y_a\right)-\frac{1}{1+\frac{1}{\eta}}\left(\frac{y_a}{a}\right)^{1+\frac{1}{\eta}}$$

$$ii) \qquad a^{1+\eta}\left(1-\tau'\left(y_a\right)\right)^{\eta}=y_a$$

$$iii) \qquad R=\int_{a_L}^{a_H}\tau\left(y_a\right)f\left(a\right)da,$$

where R denotes the revenue requirement – the amount that the government wants to collect, and $G(\,)$ is an increasing and concave function that represents the degree of aversion of the government to inequality.

Clearly, in the problem of the government, the instrument – the tax function – serves only as an indirect tool that induces behaviour ultimately leading to a given level of social welfare as described with condition (5). The actual solution, the form of the tax function, depends on the four key parameters of the model, specifically the degree of aversion to inequality as captured with $G(\,)$, the revenue requirement R, the *pdf* of the distribution of skills $f(\,)$, and the elasticity of the labour supply with respect to the net wage, η.

Mirrlees [**1971**] is rightfully given credit for formulating the problem and for being able to derive the relevant optimality conditions that characterise the optimal tax function. Specifically, the standard tools of calculus of variations permit for the derivation of a non-linear second order differential equation that identifies the level of realised utility function U_a and the level of income earned as we have $\frac{dU_a}{da}=\frac{y_a^{1+\frac{1}{\eta}}}{a^{2+\frac{1}{\eta}}}$ and finally the actual tax function $\tau\left(y_a\right)$, given condition (4).

Specifically, the relevant condition that identifies the optimum is given with

$$G'\left(U_a\right)f(a)-\lambda f(a)=\frac{\eta}{\eta+1}\frac{d}{da}\left[\lambda\left(a^{\frac{2\eta+1}{\eta+1}}\left(\frac{dU_a}{da}\right)^{\frac{-1}{\eta+1}}-a\right)f(a)\right].$$

It is unfortunately the case that, in general, the above equation must be handled[1] with numerical techniques as there are no analytical methods that would make it possible to approach differential equations of this type. Without imposing any additional restrictions one can only make two general observations. Specifically, as shown by **Sadka** [**1976**] and **Seade** [**1977**], the marginal tax rates at the top and the bottom of the distribution of skills must be equal to 0 unless there is no work provided at the bottom of the distribution of skills. Moreover, as noted originally by **Mirrlees** [**1971**], the optimal marginal tax rates are bounded between 0 and 1.

Indeed, few additional insights can be provided unless one is willing to introduce additional restrictions into the problem of **Mirrlees** [**1971**]. In fact, 40 years after **Mirrlees** [**1971**], **Diamond and Saez** [**2011**] suggested a seemingly minor modification to the problem of **Mirrlees** [**1971**] that allowed for a sharper conclusion with regard to the shape of the optimal tax function at least on a part of the domain. Specifically, **Diamond and Saez** [**2011**] note that stipulating the bounded distribution of skills can be misguided as individual productivity is realised rather than given within a given period and can, in particular, attain values not seen in the past. Accordingly, **Diamond and Saez** [**2011**] argue that modelling skills as being drawn from an unbounded distribution can be more appropriate and can yield more meaningful results. In other words, setting $a_H=\infty$ can be more than reasonable if one is willing to address the issue of optimal taxation in a realistic context.

Again, careful analysis reveals that a simple change to an unbounded distribution of skills by itself does not yield automatically any new insights. In fact, new results emerge only under special circumstances. Specifically, if the distribution of skills has the desired properties then the optimal tax function can be characterised by non-zero and possibly high marginal tax rates at high income levels. In particular, it can be shown that once the distribution of skills has enough mass at the high income level then the optimal marginal tax function is characterised by non-zero marginal tax rates even when income becomes very high. This result is in stark contrast to that obtained nearly automatically when the distribution of skills is bounded. In particular, **Diamond and Saez** [**2011**] show that the distribution of skills in the United States is fat-tailed – assuming

1 We apply proper numerical techniques to solve the above equation and present explicit solutions in a companion paper, **Dudek et al.** [**2020**].

a Pareto form at high income levels – and as a result marginal taxes at high income levels should be positive. Similarly, in the remainder of the paper, we show that the distribution of skills in Poland can also be described with a function of Pareto form and accordingly the optimal marginal taxes in Poland are not zero either.

In particular, if the distribution of income at a high income level takes the Pareto form, i.e., we can state that $P(income > y) = \frac{A}{y^m}$, where A is a constant we can state that the solution of Saez to the optimal taxation problem is particularly elegant and in the case of high income levels is given with $\tau = \frac{1-g}{1-g+\bar{\eta}^u+\bar{\eta}^c(m-1)}$, where m is the parameter characterising the distribution of income, and where $\bar{\eta}^c$ and $\bar{\eta}^u$ denote the compensated and uncompensated elasticities of the labour supply with respect to the wage, and g captures the marginal social welfare weight attached to high-income earners.

The above formula can be simplified even further given our specification (note the absence of income effects) of the utility function given with condition (1) as in our case the two values of elasticities are the same and equal to η. Accordingly, we have $\tau = \frac{1-g}{1-g+\eta m}$.

In the remainder of the paper, we utilise the above formula – originally derived by Saez – in the case of Poland. The formula is elementary, but its application is based on a number of preconditions. First of all, to make the formula useful, one must determine the value of m, a parameter describing the distribution of income in Poland. More precisely, one must verify that indeed the distribution of income, at least over some range, can be described with a Pareto distribution. Only after this is established can one try to apply the above formula to find the optimal marginal tax rates in Poland. In the next section, we verify that indeed the distribution of income in Poland can be approximated with a Pareto distribution and estimate the relevant values of m.

Knowing that m exists and being familiar with its value is not enough to make the formula applicable. One must also know the value of η. In the paper, we provide the optimal marginal tax rates for a range of values of η with a special emphasis on values provided in a number of independently conducted empirical studies.

A direct inspection reveals that there is yet another parameter, g, that must be determined before the above formula becomes viable. Unfortunately, the value of parameter g reflects, in part, the preferences of the policy maker and as such is arbitrary. Accordingly, we must resort to conditional analysis and present a range of estimates of τ for different values of g. In particular, we follow Saez and focus on a special and very extreme case and consider the situation when the welfare of high-income earners weighs very little, i.e., we estimate the optimal marginal tax rates when $g = 0$, i.e., when the above formula takes the form $\tau = \frac{1}{1+\eta m}$.

The above formula is nearly readily applicable. In fact, it suffices to find the actual elasticity and examine the properties of income distribution. **Saez**, in his studies [**2001**], applies the formula to the case of the United States. His findings based on a sizable value of the elasticity and empirically determined value of m equal to about 1.5 led to the conclusion that the optimal marginal tax rates could be as large as 73%. In other words, the findings of **Saez** [**2001**] suggest that the currently implemented rate of 42.5% is somewhat distant from the optimal value. In the remainder of the paper, we follow **Saez's** logic [**2001**] and follow his steps in the context of Polish data. In fact, we reach similar conclusions. First we argue that the distribution of skills in Poland has a shape that makes **Saez's** formula [**2001**] applicable and then we show that the current tax policy in Poland does not coincide with the optimal tax schedule.

High-end income distribution in Poland

The formula for the optimal marginal tax rates derived by **Saez** [**2001**] is simple and readily applicable. However, there are fundamental preconditions that must be satisfied before the formula can in fact be used. In particular, the simple version of the formula – equation (10) – is useless unless the distribution of income in Poland has proper statistical properties. Therefore, it is necessary to first examine whether the distribution of income in Poland has the desired properties. Specifically, we need to determine whether the distribution of income is fat tailed at high income levels or more precisely whether it can be described with a Pareto distribution for a suitably chosen parameter m.

It is worth noting that the observed levels of income earned in Poland are driven by a number of factors. In particular, the current marginal tax in Poland, τ_{PL}, the form of the utility function, relationship (1), and the *pdf* of the distribution of skills, $f(\ )$, drive the observed distribution of income. However, as shown by **Saez** [**2001**], once the observed distribution of income under the existing policy is of the Pareto form, so is the distribution of skills and the distribution of income under the optimal tax schedule. Therefore, it suffices to examine the properties of the currently observable income levels that are driven by the policy in place, which, naturally needs not to be optimal.

We start by reiterating that the actual sample is always finite and bounded. In particular, there exists a value that happens to be the highest. Naturally, given that the highest value of income exists in the sample, there must be the highest level of skills in society. This, in turn, implies, given the findings of **Mirrlees** [**1971**], that the optimal marginal tax rate must be equal to 0 at the highest income level. However, such a conclusion, as again noted by **Saez** [**2001**], can be somewhat premature. In fact, it may be more appropriate to view the observed income levels as simple realisations of a random variable that come from an unbounded distribution of skills. Such a view can be justified on the grounds that the highest income earned fluctuates over time and on the basis of the observation that the income levels at high levels tend to be sparse. Accordingly, in this paper, we follow the lead of **Saez** [**2001**] and try to model the distribution of skill and in turn income as being unbounded and will try to approximate the observed distribution of income as realisations of a random variable described with an unbounded distribution.

In the paper, we utilise actual Polish data. Our data has been obtained from the Tax Collection Department of the Polish Ministry of Finance. Our sample contains data from 2016 and was randomly generated. It contains 54,083 personal income tax records. The majority of the taxpayers in our sample, 35,012, filed individually and the reminder filed either jointly or as single parents. In our analysis, we do not utilise any of the individual characteristics of taxpayers. We understand that this is a shortcoming of our analysis and plan to address those issues in subsequent projects. Similarly, our focus on single earners can be viewed as insufficient as our choice invariably affects percentile income thresholds. Again, we view our approach as a stepping stone to additional research and we believe that our findings – while being only of the first pass type – can be a significant reference and stepping stone to subsequent endeavours. Specifically, in our sample, we observe data on nearly all values for low and moderate income levels. Furthermore, as income levels increase observations become more scarce, and, finally at very high income levels we observe only isolated realisations. We present a summary statistic of our data in Table 1.

Table 1. Percentile income thresholds of actual personal income reported by Polish individual taxpayers in 2016 (thousands of PLN)

1%	2%	3%	4%	5%
[0,0.3)	(0.3,0.7]	(0.7,1.1]	(1.1,1.7]	(1.7,2.3]
6%	7%	8%	9%	10%
(2.3,3.0]	(3.0,3.6]	(3.6,4.4]	(4.4,5.3]	(5.3,6.2]
11%	12%	13%	14%	15%
(6.2,7.2]	(7.2,8.0]	(8.0,8.5]	(8.5,8.8]	(8.8,9.6]
16%	17%	18%	19%	20%
(9.6,10.2]	(10.2,10.5]	(10.5,10.9]	(10.9,11.4]	(11.4,11.9]
21%	22%	23%	24%	25%
(11.9,12.2]	(12.2,12.3]	(12.3,12.7]	(12.7,13.0]	(13.0,13.4]
26%	27%	28%	29%	30%
(13.4,13.8]	(13.8,14.4]	(14.4,14.8]	(14.8,15.3]	(15.3,15.7]
31%	32%	33%	34%	35%

(15.7,16.2]	(16.2,16.6]	(16.6,17.0]	(17.0,17.4]	(17.4,17.5]
36%	37%	38%	39%	40%
(17.5,17.7]	(17.7,17.8]	(17.8,18.2]	(18.2,18.5]	(18.5,18.9]
41%	42%	43%	44%	45%
(18.9,19.3]	(19.3,19.6]	(19.6,20.0]	(20.0,20.5]	(20.5,20.9]
46%	47%	48%	49%	50%
(20.9,21.3]	(21.3,21.7]	(21.7,22.1]	(22.1,22.6]	(22.6,23.1]
51%	52%	53%	54%	55%
(23.1,23.6]	(23.6,24.1]	(24.1,24.6]	(24.6,25.1]	(25.1,25.5]
56%	57%	58%	59%	60%
(25.5,26.1]	(26.1,26.7]	(26.7,27.2]	(27.2,27.8]	(27.8,28.3]
61%	62%	63%	64%	65%
(28.3,28.9]	(28.9,29.6]	(29.6,30.2]	(30.2,30.9]	(30.9,31.6]
66%	67%	68%	69%	70%
(31.6,32.3]	(32.3,33.2]	(33.2,34.0]	(34.0,34.8]	(34.8,35.6]
71%	72%	73%	74%	75%
(35.6,36.5]	(36.5,37.4]	(37.4,38.3]	(38.3,39.3]	(39.3,40.3]
76%	77%	78%	79%	80%
(40.3,41.3]	(41.5,42.5]	(42.5,43.6]	(43.6,44.8]	(44.8,46.2]
81%	82%	83%	84%	85%
(46.2,47.4]	(47.4,48.8]	(48.8,50.4]	(50.4,51.9]	(51.9,53.6]
86%	87%	88%	89%	90%
(53.6,55.3]	(55.3,57.4]	(57.4,59.4]	(59.4,62.0]	(62.0,64.7]
91%	92%	93%	94%	95%
(64.7,67.8]	(67.8,71.1]	(71.1,75.3]	(75.3,80.0]	(80.0,85.9]
96%	97%	98%	99%	100%
(85.9,93.8]	(93.8,105.4]	(105.4,125.4]	(125.4,166.5]	(166.5,∞)

Source: Authors' own elaboration.

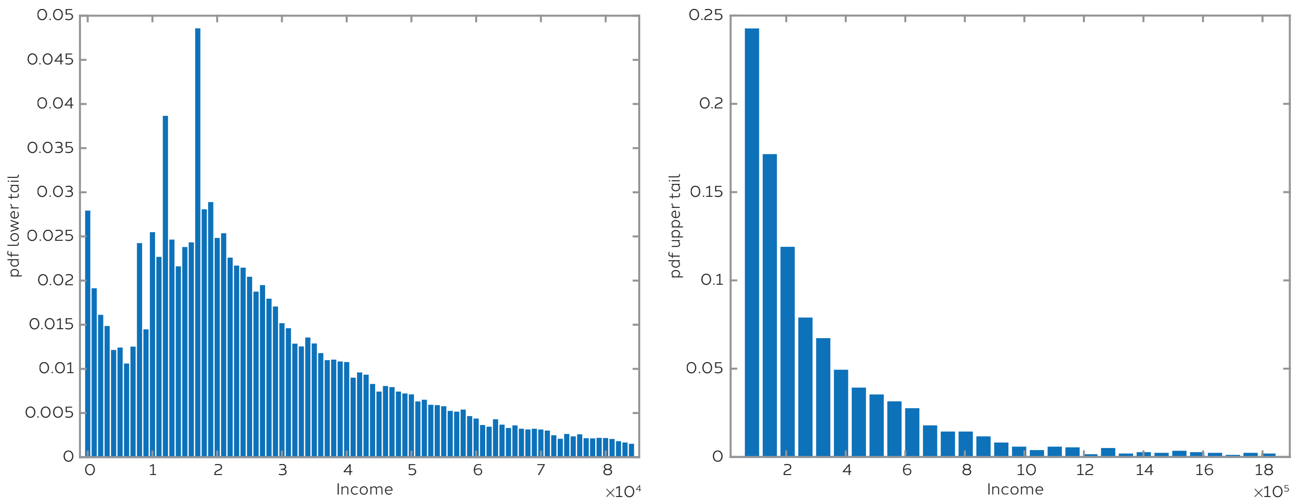
In our analysis, we focus on pre-tax income, which was subject to a 32% marginal tax rate at income levels exceeding PLN 85,000 in 2016. Furthermore, we focus exclusively on income net of any social security contributions. In addition, we limit our attention to single earners and analyse only labour income. Moreover, we do not condition our analysis on gender even though there is evidence that the actual elasticity of the labour supply with respect to net wage could vary drastically across genders. We naturally realise that the approach we take by design can be viewed only as a starting point to any policy discussion as we neither account for potential migration induced by higher tax rates, as noted in **Kleven et al.** [**2013**], nor do we address the duality of the Polish tax system and its potential importance, as outlined in **Kopczuk** [**2012**] and noted in **Doligalski** [**2019**].

Throughout the paper we refer loosely to the data we use in our analysis as current and the policy in place at the time as current even though the data comes from 2016. Naturally, the levels of income, the distribution of income, and the policy itself have evolved since then. In particular, the cut-off level of income, PLN 50,000, routinely used throughout the paper has probably lost its strict meaning. Hence, we repeatedly emphasise in the paper that income of PLN 50,000 corresponds to the 83 percentile. Moreover, recent adjustments in the tax code itself must have influenced the actual distribution of income in Poland in recent years. This potentially could have led to a change in the value of the parameters describing the distribution of income and potentially affect the specific values of the estimates of the optimal tax rates. We deal with this potential additional source of uncertainty by allowing for conditional analysis and by performing a multitude of robustness checks. We believe that it is important to add that the data that we used is likely to reflect the actual behaviour of rational economic agents and as such can be used as a basis for a formal analysis as it is not affected by more recent adverse shocks related to the ongoing crisis.

Again, let us reiterate one more time that the realised levels of income are by design bounded as those pertain to the actual Polish population, which is naturally finite. Nevertheless, individual productivity can and does change over time possibly in a stochastic manner, which naturally implies that the observed income levels can vary over time especially at high income levels where there are few observations. Thus, assuming that the distribution of skills is in fact unbounded can be more than appropriate. It is also worth noting that the relative scarcity of observations of realised income at very high income levels does not necessarily contradict the premise that the distribution of skills is thick tailed. It can be de facto a simple consequence of the fact that our sample is finite and purely mechanically there are few observed realised values at very high income levels, which is consistent with and implied by the underlying likelihood function.

We complement the content of Table 1 by providing the actual *pdfs* of the observed conditional distributions in figures (1a) and (1b).

Figure 1. The conditional *pdfs* of the actual distribution of income



Source: Authors' own elaboration.

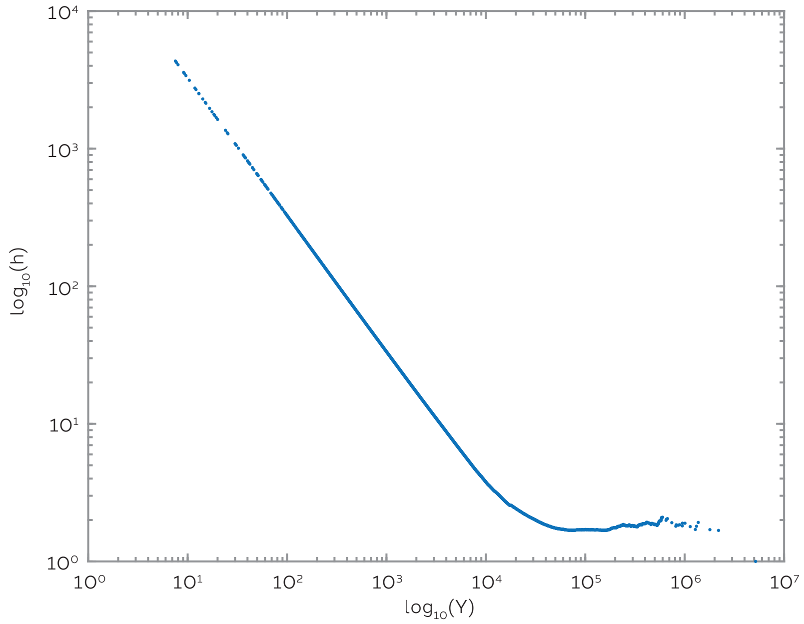
In the two figures, we present the *pdf* of the distribution of income for two distinct ranges. First we depict the *pdf* for income levels smaller than PLN 85,000. and then for income levels higher than PLN 85,000. Naturally, we can conclude, by direct inspection, that income in Poland is not distributed according to a Pareto distribution on the entire domain. However, as it is revealed in panel (1b) income levels that are high enough can in fact be potentially drawn from a Pareto distribution. We attempt to quantify this observation below by relying on a number of characteristics that can be used to identify the relevant parameters of the distribution.

Let us start by looking at a simple metric that plays an important role in the general Saez formula [2001]. Specifically, let us analyse the behaviour of a simple ratio defined with

$$h(z) = \frac{\int_z^{\infty} y f_y(y)\,dy}{z\int_z^{\infty} f_y(y)\,dy}.$$

The above expression can assume different values for different values of z. However, when the distribution of income takes a Pareto form the expression above simplifies neatly to a constant. Specifically, if we continue to assume that the distribution of income is described with $F_y(y) = P(income > y) = \frac{A}{y^m}$ then the corresponding *pdf*, $f_y(y)$, is proportional to $\frac{1}{y^{m+1}}$ and the resulting value of $h(z)$ is constant and given with $\frac{m}{m-1}$. Therefore, it suffices to perform a simple test in order to determine whether the distribution of income, and over which range, has the desired properties needed for the direct application of the formula of Saez [2001].

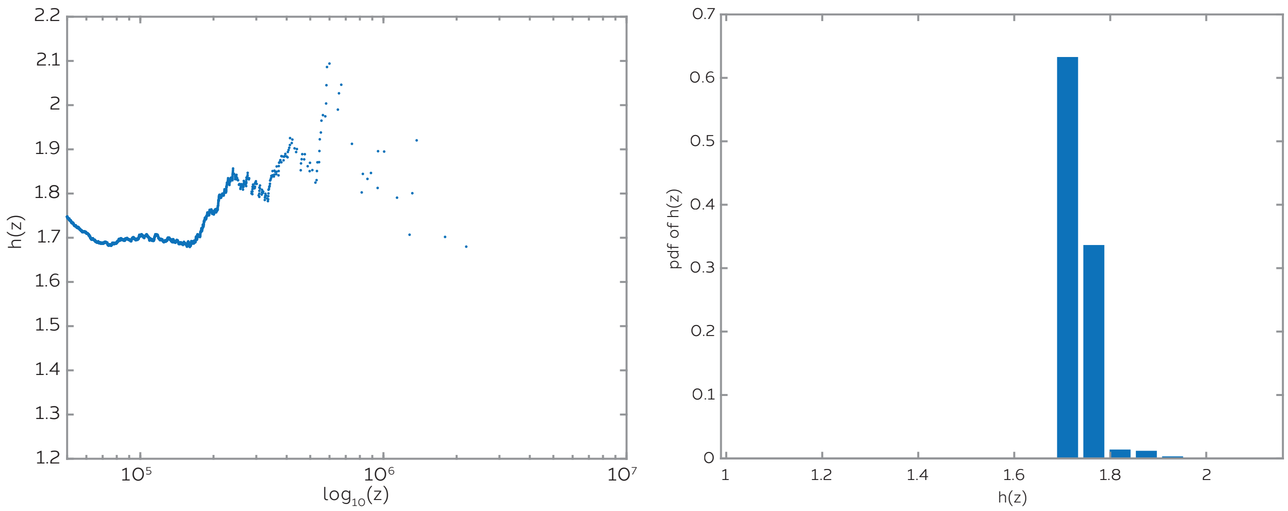
Figure 2. The actual value of *h*( ) in Poland



Source: Authors' own elaboration.

Observe that the provided sample allows us to construct the relevant histograms and to determine the empirically observed shape of $f_y(y)$. Having found the empirical expression describing $f_y(y)$ we can use formula (11) to determine whether it indeed remains constant and if so over which range. The value of the log of the hazard rate, $h(z)$, is presented in Figure (1). Clearly, the value of the hazard is not constant. Nevertheless, it is apparent that the actual variation of the hazard changes dramatically when one traverses the domain. Specifically, for income levels that are high enough the hazard rate visibly stabilises at least in the relative sense potentially giving credence to the claim that indeed the distribution of income at high income levels can be of Pareto form. We naturally explore this possibility next by focusing on the upper ranges of income.

Figure 3. The Behavior of *h*(*z*) at High Income Levels

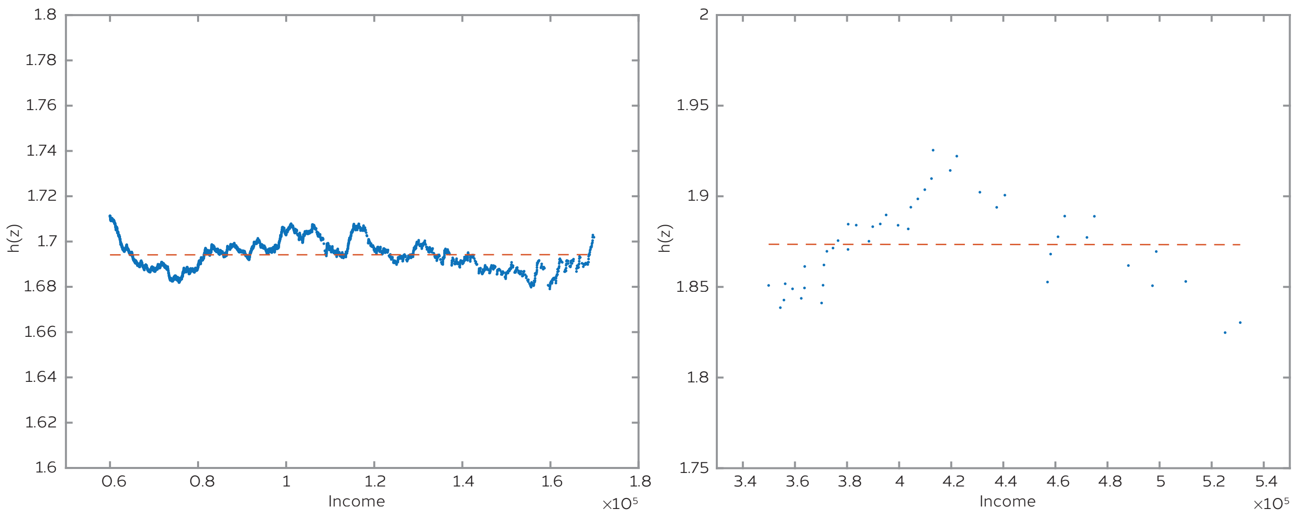


Source: Authors' own elaboration.

To be able to effectively apply **Saez's** formula [**2001**], we must first establish that indeed the hazard rate is roughly constant at high income levels. However, this in itself requires us to first agree what range of income can be viewed as high income. In this paper, we will take a very pragmatic approach and consider a number of ranges to try to identify the range that is most applicable, i.e., the range where the hazard rate indeed exhibits a fair amount of stability. Specifically, if one focuses on income levels in excess of PLN 50,000, which corresponds to the 83 percentile in the income distribution, then the hazard rate can viewed as roughly stable as visualised in Figure (3). The instability reappears again at somewhat higher income levels. However, it is largely driven by the relatively small number of observations at somewhat higher income levels. Note that given that the sample is finite the hazard rate must reach its final value of 1 at the highest value of realised income even if the underlying distribution of skills is of Pareto form on the entire domain. We next examine the stability of the hazard function on finer parts of the domain.

It is naturally the case that the value of $h(z)$ varies with z even at high income levels. Unfortunately, just like in the case of the United States, we cannot simply attest that the distribution of income in Poland is of Pareto form in the strict sense. Nevertheless, it appears that assuming that the distribution of income in Poland can be approximated with a Pareto distribution can be in fact reasonable, especially if one chooses to acknowledge the technical shortcomings of any estimation procedure applied to finite samples. Again, note that the sample at our disposal is finite. Therefore, even if the true distribution of skills, and hence the true distribution of income, is of Pareto form, the hazard rate cannot be constant for all values as it must be by construction converging to 1 as the observed income converges to the highest value. Thus, it is imperative to introduce an upper bound for the income level if one is willing to assess the stability of the hazard rate, $h(z)$. Moreover, as the observed income increases the number of observed realisations of income must, by construction, be decreasing even if the underlying distribution is thick-tailed. Therefore, the observed values of the hazard rate will be more volatile as income increases because even a single observation in the range of high income levels is likely to contribute heavily to the observed value of the hazard rate. Furthermore, let us note that even in the presence of shortcomings stemming from the finiteness of the sample, the actual values of the hazard rate remain bounded. Moreover, the frequency of realisations affected by the under-sampling at high values is relatively low. To further underscore the above observations, we illustrate the observed values of the hazard rate in the case when the extreme observations are ignored – Figures (4a) and (4b).

Figure 4. The Dependence of $h(z)$ when Extreme Observations are Ignored



Source: Authors' own elaboration.

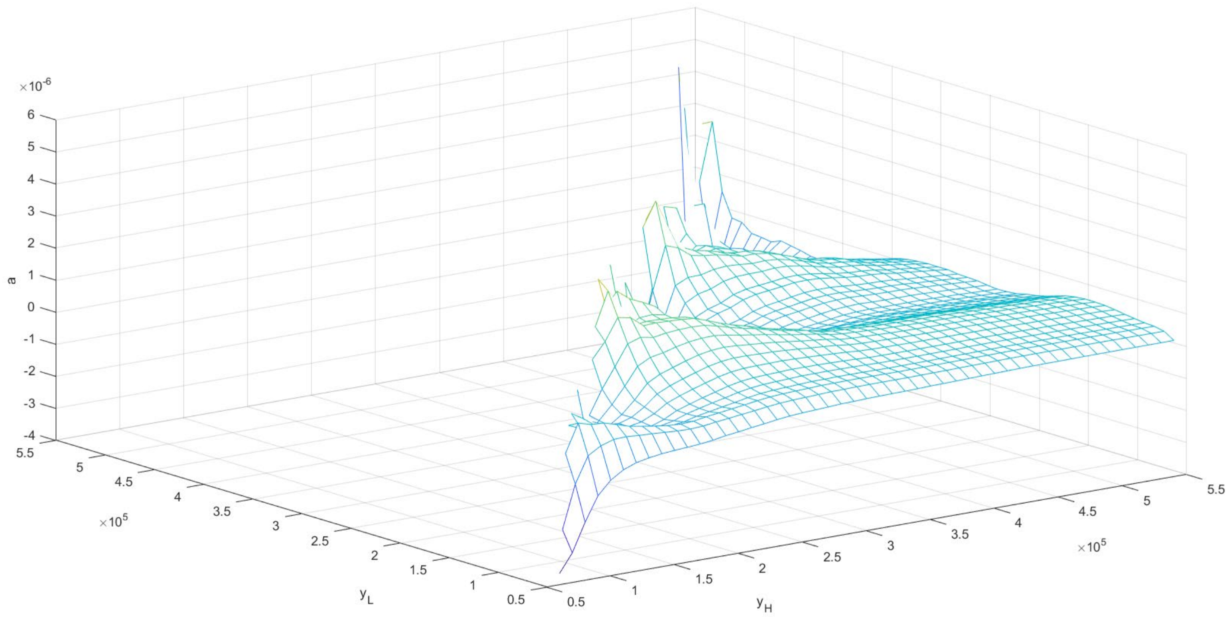
To determine the range of applicability of Saez's formula, let us start by limiting our attention to a range of incomes between y_L and y_H and let us examine how the hazard function varies along such an interval. We already know that for arbitrary values of y_L and y_H the hazard function does not display stability, as visualised in Figure (1). However, if we restrict y_L to be somewhat higher than 50,000 – 83 percentile – at least partial stability seems to appear. Naturally, we are now compelled to identify the range of income levels where the stability of the hazard rate appears to be the highest. We do so by considering a number of potential values for y_L and y_H with increments of PLN 10,000 to identify the range where $h(z)$ is the most stable. Our stability criterion is simple and relies on a simple linear regression of the form

$$h = az + b + \varepsilon,$$

where h is to be equal to the observed hazard rate, $h(z)$.

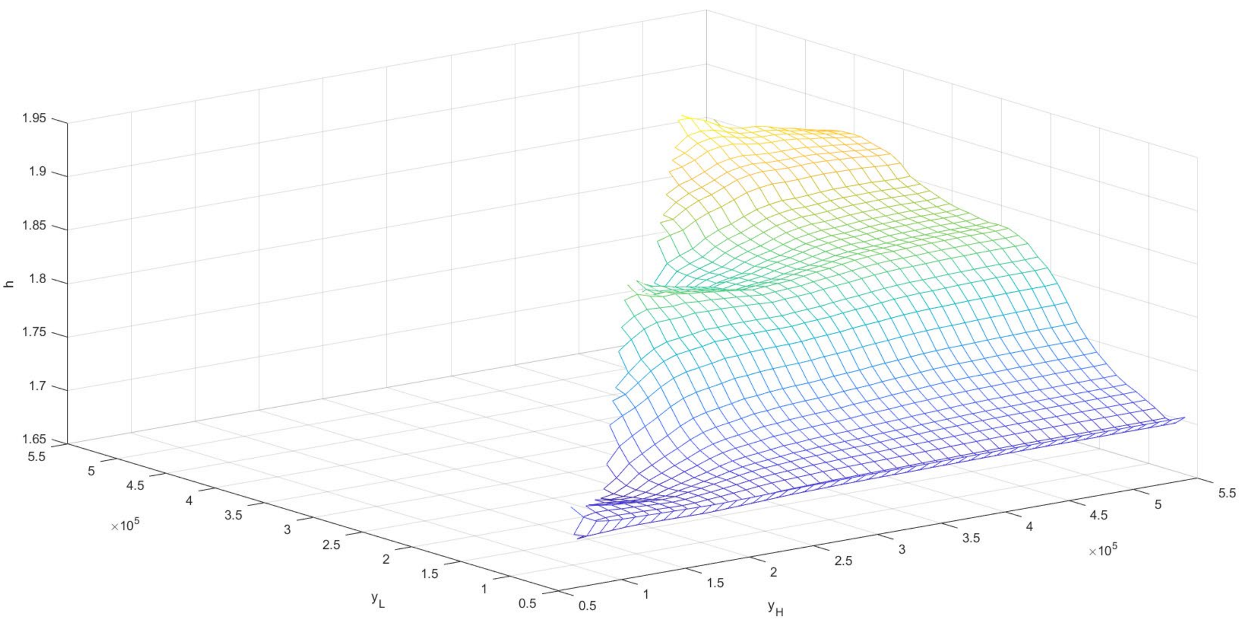
Using a simple OLS approach, we estimate the actual value of a for all the potential values of y_L and y_H and identify the set of values of y_L and y_H for which the estimated value of $a_{\{y_L, y_H\}}$ is the closest to 0 in the statistical sense. In Figure (2), we show the estimated values of a for a range of values of y_L and y_H, where $y_L < y_H$.

Figure 5. Estimates of *a*



Source: Authors' own elaboration.

Figure 6. The average value of *h*(*z*) depending on the range y_L, y_H



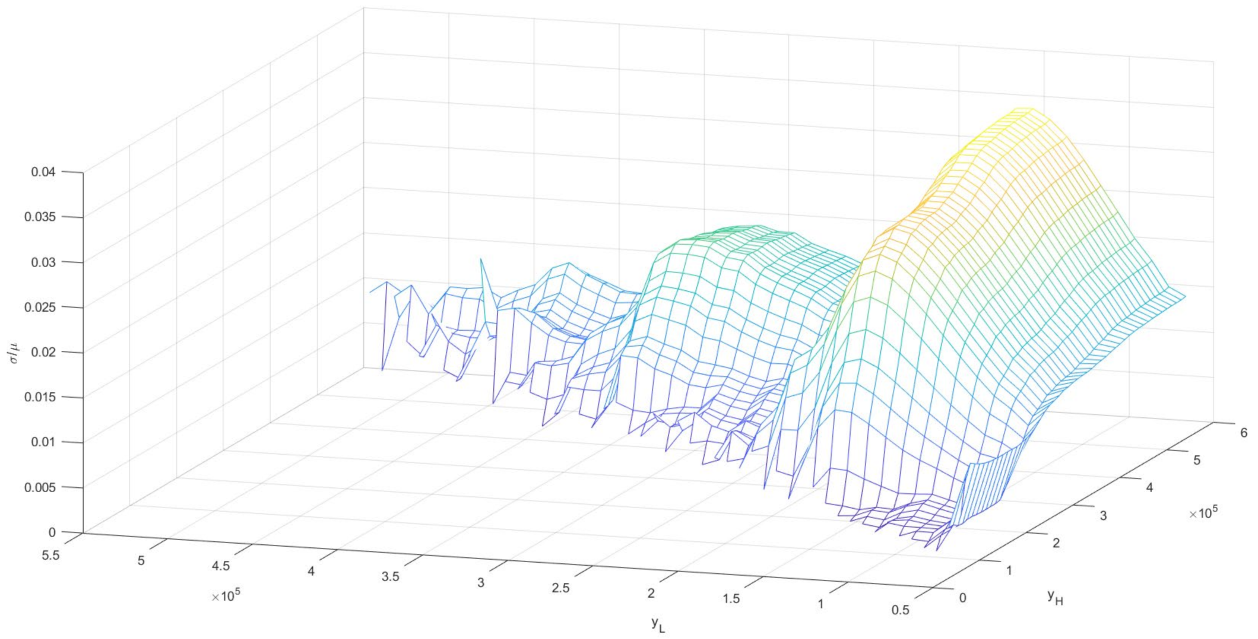
Source: Authors' own elaboration.

Our estimates reveal that there is a wide range of income brackets for which the value of parameter a in our model specification (12) cannot be statistically differentiated from 0. In fact, it turns out that the distribution of income in Poland resembles most closely a Pareto distribution when income is in between 60,000 and 180,000, as visualised in Figure (4a). In this case, the corresponding value of the parameter, as implied by equation (13), is equal to $\hat{m} = 2.4407$.

Again, the actual hazard rate is never constant and always displays some variability. Furthermore, even the choice of the cut-off levels y_L and y_H, as expected, affects the hazard rate and its average value as visualised in Figure (3). Still, one may notice that the average values of $h(z)$, depending on the actual choice y_L and y_H, remain relatively stable. In particular, an increase in y_H for a given value of y_L leads to a relatively modest change in the value of the average value of the hazard.

To further underscore the surprising stability of the hazard rate, we choose to look at the coefficient of variation, $\frac{\sigma}{\mu}$, of the hazard rate for different values of y_L and y_H. The actual values are depicted in Figure (4).

Figure 7. The coefficient of variation as a function of y_L and y_H



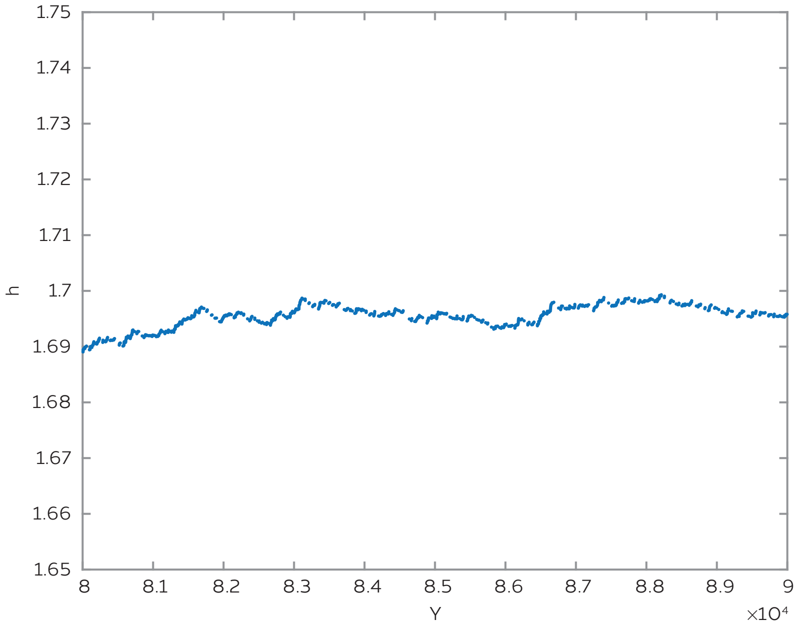
Source: Authors' own elaboration.

The coefficient of variation is not equal to 0, but is fairly low. Furthermore, it normally does increase with y_H as larger and more isolated values of income affect the actual hazard rate. It turns out that the variation coefficient is the smallest when y is between 80,000 and 90,000. The corresponding values of the hazard rate, $h(z)$, are illustrated in Figure (5). In this case, the implied value of m, given condition (13), is equal to 2.4382.

We can also attempt to test whether the observed distribution of income is fat-tailed by comparing the distribution of income that is actually observed to a class of Pareto distributions characterised by parameter m. In particular, one can simply estimate the value of m given the observed income levels for different values of y_L and y_H by utilising the maximum likelihood approach.

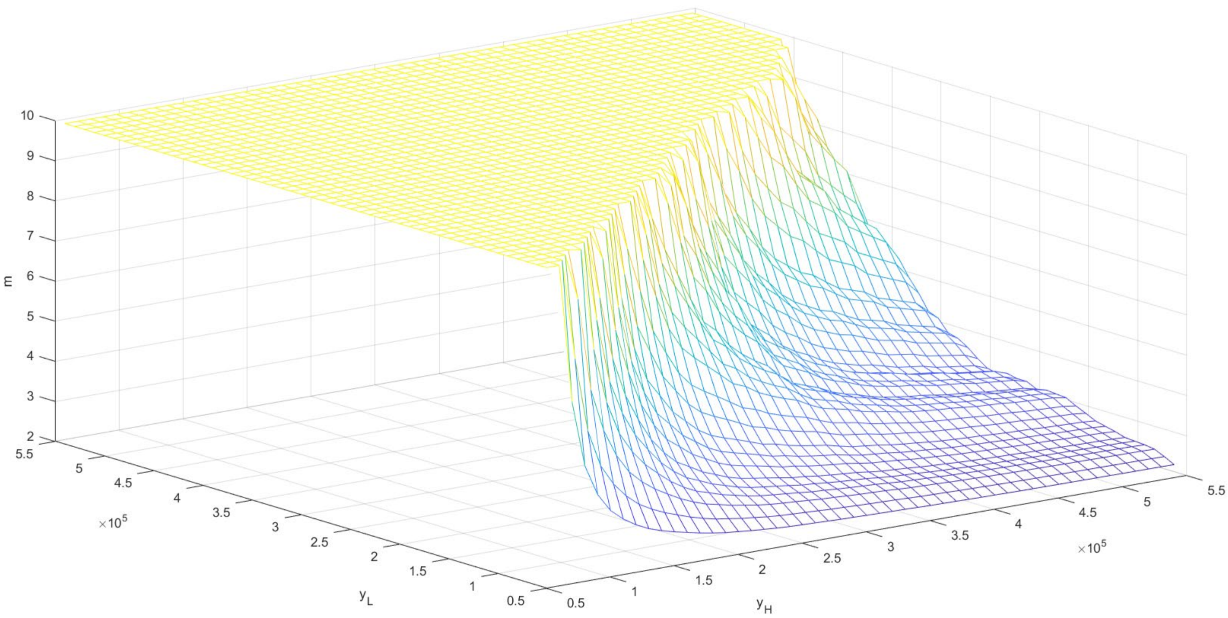
The actual estimates of the value of m vary depending on the range of income chosen. We present the estimated values of m for different choices of y_L and y_H in Figure (6).

Figure 8. Lowest coefficient of variation



Source: Authors' own elaboration.

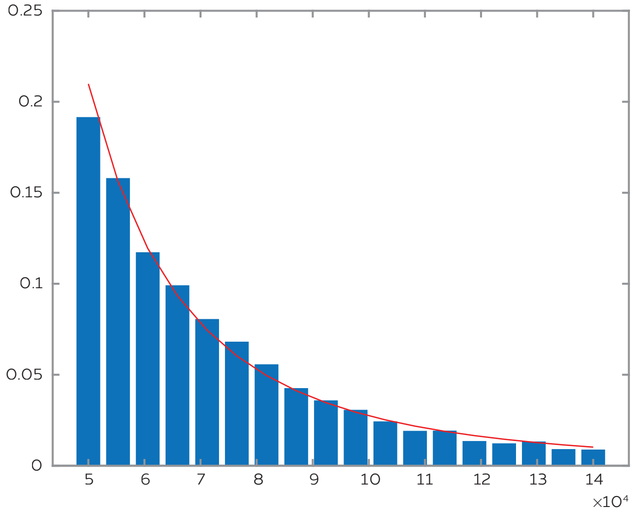
Figure 9. The estimated values (MLE approach) of *m* for different values of y_L and y_H



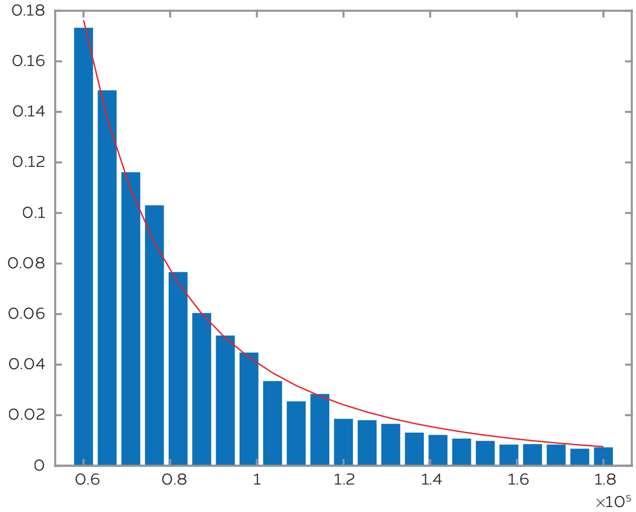
Source: Authors' own elaboration.

The degree of variability of the estimated values of *m* is substantial. Some values are extreme whereas others are much more modest and to a large extent much less volatile. Formally, we must state that the choice of the relevant range of income is critical as an arbitrary choice of y_L and y_H does not lead to a stable value of *m*. This naturally implies that the distribution of income in Poland does not follow the Pareto distribution on an arbitrary domain. Nevertheless, it may still be the case that for a suitably chosen range of income the distribution of income in Poland can in fact be approximated with a Pareto distribution. To underscore this point, we present the actual histograms of the observed income in Poland superimposed on the estimated *pdf* of income under the assumption that the estimated distribution is of Pareto form. We present our findings in Figure (10).

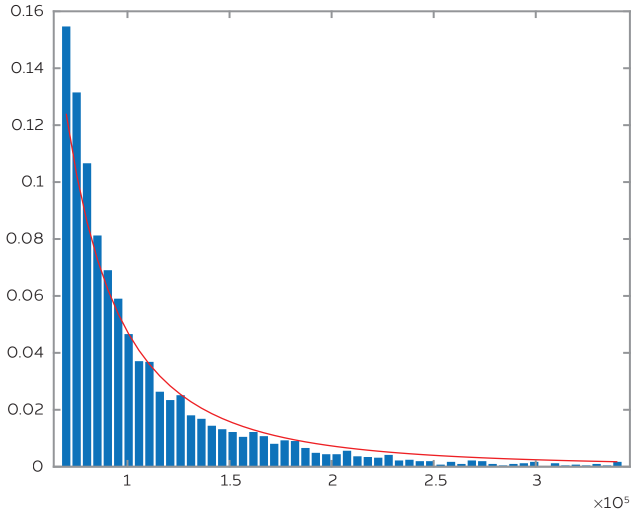
Figure 10. Actual and estimated *pdfs* of the distribution of income in Poland for a selection of choices of *yL* and *yH*



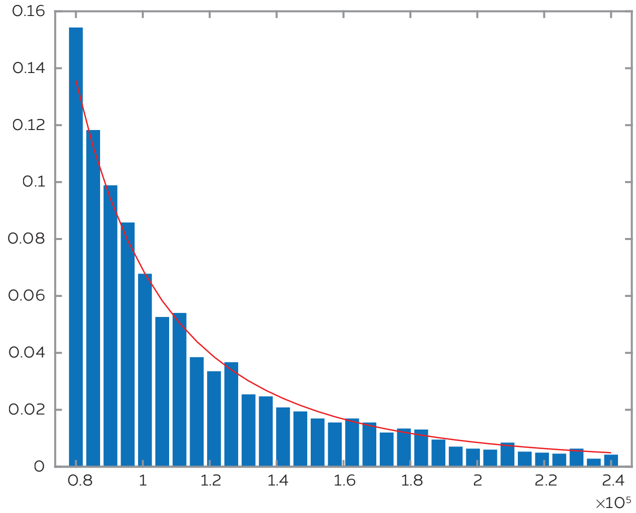
(a) $y_L = 50000$, $y_H = 140000$, $\hat{m} = 1.9344$



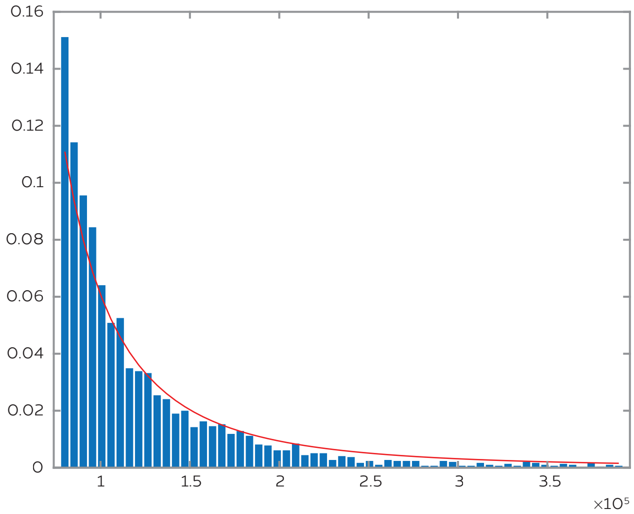
(b) $y_L = 60000$, $y_H = 180000$, $\hat{m} = 1.8695$



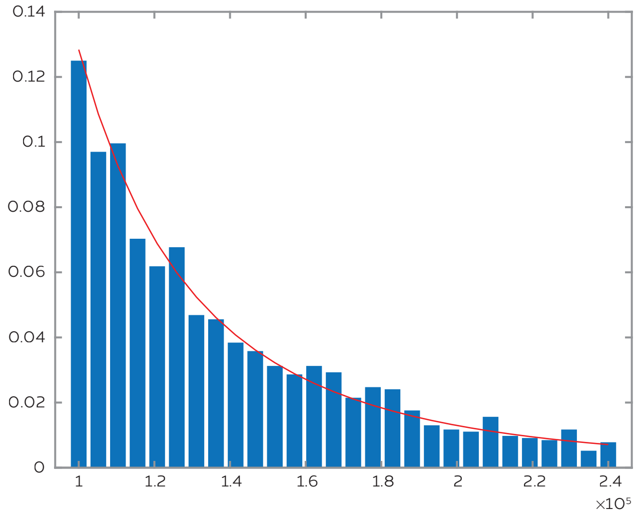
(c) $y_L = 70000$, $y_H = 340000$, $\hat{m} = 1.6977$



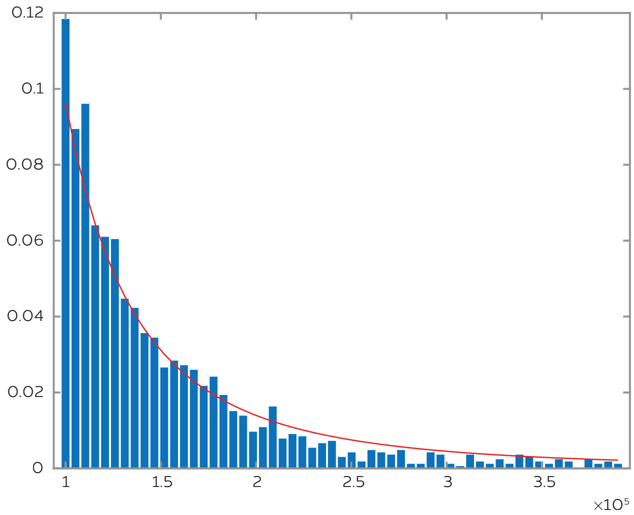
(d) $y_L = 80000$, $y_H = 240000$, $\hat{m} = 2.0213$



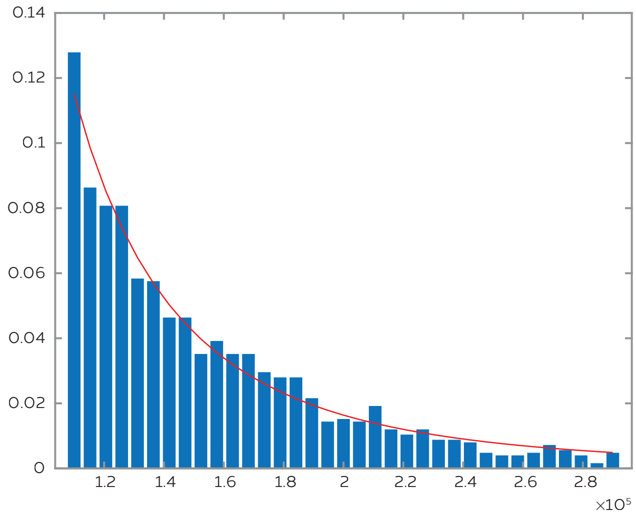
(e) $y_L = 80000$, $y_H = 390000$, $\hat{m} = 1.6994$



(f) $y_L = 100000$, $y_H = 240000$, $\hat{m} = 2.3084$



(g) $y_L = 100000$, $y_H = 390000$, $\hat{m} = 1.7872$



(h) $y_L = 110000$, $y_H = 290000$, $\hat{m} = 2.2612$

Source: Authors' own elaboration.

Even if we were to agree based on a simple visual metric that indeed, given the plots in Figure (10), the distribution of income in Poland, at least over some ranges of income, can be approximated with a Pareto distribution, we must still decide which range to choose and we must determine the corresponding value of the estimated value of m. Such a choice will always be a function of the underlying income bracket. However, we can draw some key basic conclusions. First of all, the Pareto distribution remains a good approximation for income distribution in Poland at high income levels. Moreover, the estimated parameter defining the corresponding distribution is somewhere in between 1.6977 and 2.3084 if one relies on a maximum likelihood estimation for the income ranges considered above. We will show below that such a degree of variability of the value of the parameter will not impact the estimated value of the optimal tax in a significant manner. Naturally, the precise value of the tax can only be found once one decides which income level is relevant and can be viewed as high income in Poland.

Let us reiterate that the actual hazard rate when the income distribution is of Pareto form can be, as noted above, expressed as

$$h(z) = \frac{\int_z^\infty y f_y(y)\,dy}{z\int_z^\infty f_y(y)\,dy} = \frac{\int_z^\infty y \frac{A}{y^{m+1}}\,dy}{z\int_z^\infty \frac{A}{y^{m+1}}\,dy} = \frac{m}{m-1}.$$

Observe that the left-hand side of the above condition can be reconstructed from the Polish data. As noted above, the actual value of $h(\,)$ is never constant. However, its fluctuations are relatively small for sufficiently high income levels. In fact, as argued above, the value of m can be estimated using a number of approaches. In fact, our findings based on the OLS approach suggest that m is close to 2.4407 for a suitably chosen range of income. Similarly, the approach based on the analysis of the coefficient of variation yields 2.4382 as the most reasonable estimate of m. Finally, the maximum likelihood estimation suggests that m is somewhere in between 1.6977 and 2.3084. In general, we can conclude that our findings correspond to those presented in **Chrostek et al. [2019]**. In the next section, we consider all of those values to estimate the actual optimal marginal tax rate in Poland. It turns out that the actual choice of the precise value of m has a limited impact on the value of the marginal tax rate. Most importantly, a fair degree of stability of the value of m for a wide variety of income ranges in Poland makes the formula of Saez applicable and allows us to gauge the actual value of the optimal marginal tax rates in Poland.

Optimal marginal taxes in Poland

The ingenious approach of Saez [2001] to the optimal taxation problem allowed Saez [2001] to derive a simple formula for the asymptotic rate. In particular, Saez [2001] showed that the optimal asymptotic rate is given with expression (9), which we for convenience reiterate below

$$\tau = \frac{1-g}{1-g+m\eta},$$

where η denotes the elasticity of the labour supply with respect to the net wage, g captures the importance of high-income earners in the welfare criterion of the policy maker, and m is a parameter characterising the distribution of income. The formula is simple. However, its applicability hinges on a number of factors. First of all, the distribution of income must be of a specific shape. In the preceding section, we did establish that, perhaps fortunately, the distribution of income in Poland has the desired properties, i.e., it is fat-tailed and can be approximated with a Pareto distribution with a suitably chosen parameter m. This, by itself, makes formula (14) applicable. However, there are additional requirements that must be formally addressed. In particular, since we know that the formula is valid for high-income earners, we must identify the group of taxpayers who can be viewed as high-income earners. In fact, we must decide who should be viewed as a high earner. Is a person earning more than PLN 60,000 a year a high earner? Should we assume that the cut-off is much higher? Is PLN 120,000 per year a more appropriate value? Naturally, there is no definite answer to this question. Therefore, we can, and we do, take a number of conditional approaches and apply the formula to a variety of scenarios and report the relevant optimal tax rate for each of them. We start the process from the assumption that the formula is valid as soon as the distribution of income starts resembling the Pareto distribution. Naturally, in this context, this implies that the range of income to which the formula is applicable is quite wide and it can start at income levels as low as PLN 60,000 – corresponding to 89 percentile. Of course, in our calculations, we choose the value of m that corresponds to the relevant cut-off of income. However, as we noted in the previous section, the value of m remains somewhat stable as the cut-off value increases, which implies that our implied values of the optimal marginal tax rate are also valid for cases when the relevant cut-off of income is higher than the value chosen by us.

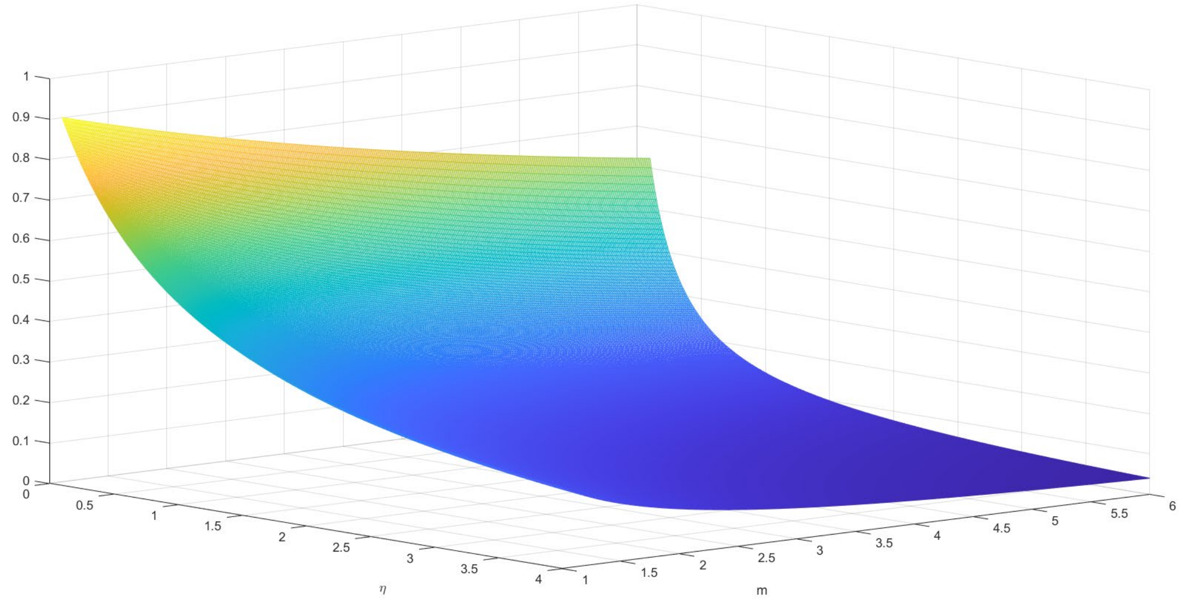
Unfortunately, deciding on the value of the cut-off level of income that makes formula (14) applicable is not sufficient. This is primarily because the formula involves other parameters as well. In particular, the value of g, the parameter that reflects the preferences of the policy maker, must be determined before the formula can be applied. This is not a simple task as g reflects preferences and is to a large extent arbitrary. Moreover, even if we assume that the policy maker favours redistribution, i.e, attaches less and less weight to the welfare of high-income earners, one must describe the evolution of g as the range of relevant income changes. To simplify our analysis, we can, following Saez [2001], assume that g becomes very small and even negligible as income increases. In fact, it may be reasonable to stipulate that g is equal to 0 at very high income levels. In such a case, the asymptotic formula for the optimal marginal tax rates assumes an even simpler form. Specifically, when $g=0$ then formula (14) takes the form given with expression (10), which we restate below

$$\tau = \frac{1}{1+m\eta}.$$

Without a doubt, such a simplification makes the formula even more handy and allows us to determine the optimal marginal tax rates. Still, it needs to be emphasised that the above formula is obtained based on two not necessarily congruent assumptions. Recall that we first chose the range, $y \in [\bar{y}, \infty)$, where in our view the distribution of income takes a Pareto form. Moreover, we assumed that the welfare of high-income earners was unimportant to the policy maker, i.e., we implicitly selected interval $[\tilde{y}, \infty)$ where, again in our view, g could be approximated with 0. Unfortunately, it is not necessarily the case that $\tilde{y} = \bar{y}$ and any policy implications can be only drawn for values of y that exceed the higher value of the two cut-off values. Of course, in the case when $\tilde{y} > \bar{y}$, the formula can be viewed as valid over a range that is fully determined by a subjective – preference based – criterion and as such has only limited applications as a tool for formal policy guidance.

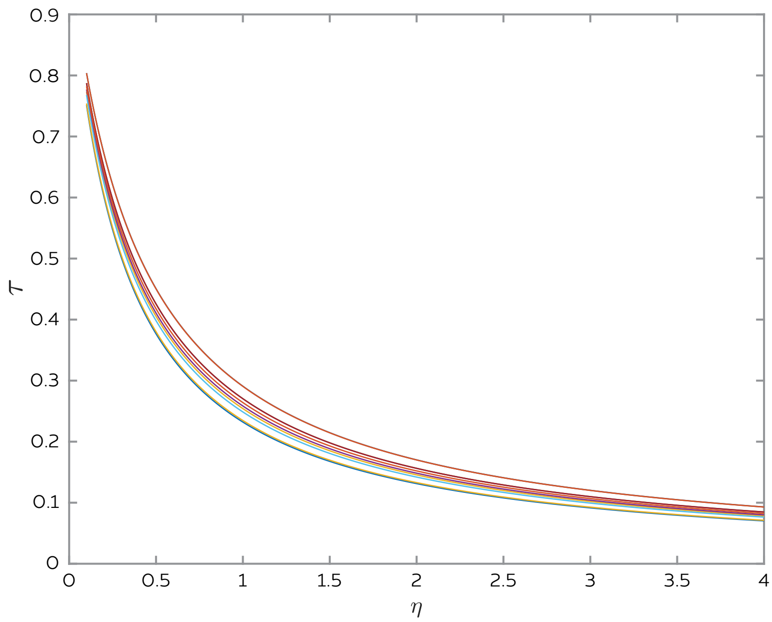
As argued in the previous section, the distribution of income in Poland can be viewed as a type of Pareto distribution over a wide range of high-income levels. Furthermore, if one is willing to take an extreme position and assume that the welfare of high-income earners is completely unimportant, i.e., that $g = 0$, then formula (15) becomes applicable to the Polish data. We start by illustrating, in Figure (7), the implied values of the optimal marginal tax rates for a variety of values of m and η.

Figure 11. The dependence of τ on η and m



Source: Authors' own elaboration.

Figure 12. The dependence of τ on η when $m = 3.44$



Source: Authors' own elaboration.

Clearly, the general formula, even if one restricts attention to the case when $g = 0$, leaves a huge amount of variability in the actual optimal marginal tax rate. In fact, it can be as high as 0.9091 and as low as 0.0400 when $m \in [1,6]$ and $\eta \in [0.1,4]$. However, as discovered in the preceding section, the actual value of m can be estimated using Polish data. In fact, our estimates suggest that the value of m can be confined to range [2.4382,3.308]. In such a case, the values of the implied tax rate exhibit a much smaller degree of variability for a given value of η, as visualised in Figure (8).

Again, as noted earlier, even if we accept some degree of uncertainty with regard to our estimated values of η, the value of the optimal marginal tax rate remains relatively stable for a given value of η However, once we allow for any type of variability in the value of η the value of the optimal tax rate starts exhibiting a huge amount of variability. In fact, the marginal tax rate can be as high as 0.8040 and as low as 0.0703. Naturally, any meaningful progress cannot be made unless one is willing to take a stand on the value of η.

Parameter η represents the elasticity of the labour supply with respect to the net wage. It is not directly observable, but it can be estimated. In his contribution, Saez uses a value of about 0.25, which he finds to be consistent with estimates for the United States. Empirical research done in the case of Poland reveals that the corresponding value for the case of Poland is of a similar order of magnitude. In particular, it has been shown (see **Bargain et al.** [**2014**], and **Myck and Trzciński** [**2019**] or **Zawisza** [**2019**]), that the value of the elasticity in Poland is also low and can be equal to about 0.2. In such a case, the estimated values of the optimal marginal tax rate become even more stable. In fact, our estimated values of m, combined with the value of elasticity of 0.2, lead to the optimal marginal tax being somewhere in the range between 0.6018 and 0.6722, with precise values given in Table 2.

Table 2. The Optimal Asymptotic Marginal Tax Rates when $\eta = 0.2$

m	2.4382	2.4407	2.9344	2.8695	2.6977	3.0213	2.6994	3.3084	2.7872	2.2612
τ_{ES}	67.22%	67.20%	63.02%	63.54%	64.59%	62.33%	64.94%	60.18%	64.21%	60.52%

Source: Authors' own elaboration.

The above values, while formally given with the **Saez** formula [**2001**], are only applicable if one restricts attention to proper income levels. First of all, income must be sufficiently high so that the distribution can indeed be viewed as being of Pareto form. At the same time, it should be sufficiently high so that the welfare of individuals with earnings in that range can be viewed as negligible. In fact, under those assumptions, we can be even more explicit and list – see Table 3 – the values of the optimal marginal tax rate for a wide range of values of η and ranges of income.

Table 3. The Optimal Asymptotic Marginal Tax Rates τ_{ES}

	$\hat{m}$	$\eta = 0.1$	$\eta = 0.2$	$\eta = 0.5$	$\eta = 1$	$\eta = 1.5$	$\eta = 2$
$y > 50000$	2.4103	80.58%	67.47%	45.35%	29.32%	21.67%	17.18%
$y > 100000$	2.3901	80.71%	67.66%	45.56%	29.50%	21.81%	17.30%
$y > 150000$	2.3235	81.15%	68.27%	46.26%	30.09%	22.30%	17.71%
$y > 200000$	2.2228	81.81%	69.22%	47.36%	31.03%	23.07%	18.36%
$y > 250000$	2.1957	82.00%	69.49%	47.67%	31.29%	23.29%	18.55%
$y > 300000$	2.1822	82.09%	69.62%	47.82%	31.42%	23.40%	18.64%
$y > 350000$	2.1457	82.33%	69.97%	48.24%	31.79%	23.71%	18.90%
$y > 400000$	2.1360	82.40%	70.07%	48.36%	31.89%	23.79%	18.97%
$y > 450000$	2.1595	82.24%	69.84%	48.08%	31.65%	23.59%	18.80%

Source: Authors' own elaboration.

The data in Table 3 reveals two distinct patterns. First of all, one can easily spot the remarkable stability of values in the vertical direction and a significant degree of variation in the horizontal direction. The stability along the vertical dimension is particularly valuable as it confirms that the actual distribution of income in Poland can be approximated with a Pareto distribution, which in itself allows us to pursue the approach of Saez to optimal taxation. This perhaps both remarkable and simultaneously fortunate fact is further underscored with the relative stability of the estimated – obtained as the average of the hazard rates in the relevant domain – value of m, which exhibits little variation for a wide range of income. This implies that we can apply the formula of **Saez** [**2001**] broadly and obtain relatively constant marginal tax rates for a given value of η. Specifically, if one is willing to take an extreme value of 0.1 for η then one can see that the optimal marginal tax rate is between 80.58% and 82.40%. At the other extreme, when $\eta = 2$, the optimal marginal tax rate is between 17.18% and 18.97%. Clearly, the values are extremely sensitive to the choice of η. One must therefore take a stand on the value of η. Again, once we reach to the body of empirical work – **Bargain et al.** [**2014**] and see **Myck and Trzciński** [**2019**], and **Zawisza** [**2019**] —we can pin down the value of η to be around 0.2. In such a case, the optimal marginal tax rate lies in between 67.47% and 70.07%, which is in the ballpark of estimates obtained by Saez for the United States. Let us reiterate that the value of the optimal tax rate is substantial and deviates from the currently implemented marginal tax rate in Poland. However, as data in Table 3 reveals, such extreme values are applicable to a wide variety of income ranges. In particular, it can be argued that such high marginal tax rates can be optimal for income levels higher than PLN 50,000 – 83 percentile.

The degree of sensitivity of the optimal marginal tax with respect to η is substantial. This naturally allows us to examine whether and under what conditions the currently observed marginal tax rate of 32% could be viewed as optimal. Answering such a question amounts to a simple reverse engineering of the Saez formula. Naturally, for 32% to be optimal, the values of η and m should be such that $0.32 = \frac{1}{1+\eta m}$, which of course can be true for a wide range of values of η and m. However, once one is willing to adopt – apparently stable – estimates of m from Table 3, it becomes apparent that the currently implemented value of the tax rate in Poland can be optimal only when η is roughly equal to 1. Unfortunately, the value of 1 appears to be somewhat distant from the value that is empirically observed. Thus, we must conclude that the policy in Poland cannot be viewed as optimal, and one may improve welfare in Poland by allowing for a sizable increase in the value of the marginal tax paid by high-income earners.

While formally valid, the above reasoning is conditional. Specifically, so far we have been assuming that the welfare weight attached to high-income earners was negligible, i.e., we were setting g to 0 in the Saez formula. Again, such an assumption can be valid for high income levels as it may be appropriate to assume that the government does not value significantly the welfare of high-income earners at the margin. However, this naturally requires that we define the relevant income range that can be viewed as containing high income levels. The data in Table 3 confirms that, for the applicability of **Saez's** formula [**2001**], we can assume that incomes higher than PLN 50,000 can be viewed as high. However, one really senses that starting from such a threshold could be somewhat inappropriate if one is willing to adhere to the assumption that g is equal to 0 as income levels in the neighbourhood of PLN 50,000 are not viewed as high, and more importantly it is somewhat hard to imagine that the welfare of individuals earning around PLN 50,000 a year can be viewed as negligible at the margin to the government. This naturally opens two possibilities. We can either somewhat arbitrarily limit the range of income and state that our results from Table 3 apply to income levels higher than a given threshold y. Such an approach is again valid, but forces us to define y or simply we could state that the optimal marginal tax rate of roughly 69% would be asymptotically valid. Alternatively, in a fully discretionary manner, we can choose a value for g. In fact, we can follow Saez and argue that his formula is indeed applicable for income levels as low as PLN 50,000, but only when applied with a value of g different than 0. Here, we choose to follow the latter approach. More precisely, we choose to adopt the methodology of **Saez and Stantcheva** [**2013**] and view g as a purely exogenous parameter that reflects the preferences of the policy maker formally unrelated to any part of the model. Let us reiterate one more time that g, while free, can be given an economic interpretation and in fact reflects the importance of high-income earners to policy makers at the margin with high values of g implying that the well-being of high-income earners is important to policy makers.

Observe that a given choice of g does not affect the form of distribution of income. This means that our estimates of m remain valid. Moreover, any choice of g does not affect the value of η either. Therefore we can select g freely in line with our subjective judgement of the importance of the welfare of high-income earners to the policy maker. We consider two values of g. Specifically, we apply the formula of Saez [2001] – equation (14) – when $g \in \{0.1, 0.2\}$. The results are given in Tables (4a) and (4b).

Table 4. The Optimal Asymptotic Marginal Tax Rates τ_{ES}

	$\hat{m}$	$\eta = 0.1$	$\eta = 0.2$	$\eta = 0.5$	$\eta = 1$	$\eta = 1.5$	$\eta = 2$
$y > 50000$	2.932639	78.88%	65.12%	42.75%	27.19%	19.93%	15.73%
$y > 100000$	2.792361	79.02%	65.31%	42.96%	27.35%	20.07%	15.84%
$y > 150000$	2.329861	79.48%	65.95%	43.65%	27.92%	20.52%	16.23%
$y > 200000$	1.630556	80.19%	66.94%	44.74%	28.82%	21.26%	16.84%
$y > 250000$	1.442361	80.39%	67.21%	45.05%	29.07%	21.46%	17.01%
$y > 300000$	1.348611	80.48%	67.34%	45.20%	29.20%	21.57%	17.10%
$y > 350000$	1.095139	80.75%	67.71%	45.62%	29.55%	21.85%	17.34%
$y > 400000$	1.027778	80.82%	67.81%	45.73%	29.64%	21.93%	17.40%
$y > 450000$	1.190972	80.65%	67.57%	45.46%	29.42%	21.74%	17.24%

(a) $g = 0.1$

	$\hat{m}$	$\eta = 0.1$	$\eta = 0.2$	$\eta = 0.5$	$\eta = 1$	$\eta = 1.5$	$\eta = 2$
$y > 50000$	2.932639	76.85%	62.40%	39.90%	24.92%	18.12%	14.23%
$y > 100000$	2.792361	77.00%	62.60%	40.10%	25.08%	18.24%	14.34%
$y > 150000$	2.329861	77.49%	63.26%	40.78%	25.61%	18.67%	14.69%
$y > 200000$	1.630556	78.26%	64.28%	41.85%	26.47%	19.35%	15.25%
$y > 250000$	1.442361	78.46%	64.56%	42.15%	26.70%	19.54%	15.41%
$y > 300000$	1.348611	78.57%	64.70%	42.30%	26.83%	19.64%	15.49%
$y > 350000$	1.095139	78.85%	65.09%	42.72%	27.16%	19.91%	15.71%
$y > 400000$	1.027778	78.93%	65.19%	42.83%	27.25%	19.98%	15.77%
$y > 450000$	1.190972	78.74%	64.94%	42.56%	27.03%	19.81%	15.63%

(b) $g = 0.2$

Source: Authors' own elaboration.

The key results obtained when $g = 0$ carry over to the case when $g \in \{0.1, 0.2\}$. First of all, essentially by construction, the robustness of the Paretian shape of the income distribution in Poland introduces little variation into the estimated optimal marginal tax rates for a given value of the elasticity of the labour supply with respect to the wage. Moreover, the substantial variability with respect to η remains. In fact, depending on the value of η, the optimal marginal taxes can be as high as 80.82% when $g = 0.1$ and $\eta = 0.1$ and as low as 14.23% when $g = 0.2$ and $\eta = 2$. However, once one limits attention to the empirically viable value for η of 0.2, a more familiar picture emerges. In fact, a shift from $g = 0$ to $g = 0.1$ has little impact on the optimal marginal tax rates as those remain high and range between 65.12% and 67.81% depending on income range. Furthermore, increasing g further to 0.2 reduces the optimal marginal tax rates still by a relatively small amount and those remain bounded between 62.40% and 65.19%. Again, as before, overall the optimal marginal tax rates when η assumes a realistic value of 0.2 remain high and exceed the current tax rate in Poland. Therefore, we can conclude that even if one assumes that the welfare of high-income earners is of importance to policy makers, the marginal tax rates at high income levels should be high and probably exceed 60%, which means that there is substantial room for improvement at the policy-making level.

As before, we can try to reverse engineer the problem and identify the value of g for which the current tax rate in Poland is optimal. Specifically, assuming that the value of m is equal to about 2.2, and the value of elasticity is equal to 0.2, the current tax rate of 32% can be viewed as optimal when $0.32 = \frac{1-g}{1-g+2.2*0.2}$. This implies that the corresponding value of g is equal to 0.8, i.e., for the current tax to be optimal the weight attached by policy makers to high-income earners must be substantial. In particular, such a substantial welfare weight would imply that out of USD 1 of revenue the government would be willing to transfer only USD 0.20 to low-income individuals, which, while formally admissible, appears not to reflect the standard preferences of policy makers.

It appears that our basic conclusion that the marginal tax rates in Poland are too low at high income levels withstands a basic robustness check. In fact, making the welfare of high-income payers to be somewhat important does not overturn the previous conclusion that the marginal tax rates should be much higher than they are at high income levels. Still, we must formally concede that in deriving our conclusion we were following the methodology of **Saez and Stantcheva** [**2013**], which formally does not put any restrictions on the value of g and allows g to be treated as a parameter. In fact, formally within that framework, one can assume that g is large and possibly close to 1. If one is willing to make such an assumption, one essentially assumes that the welfare of high-income earners is important to policy makers, which could be true in reality. Naturally, we can conclude that, under such a scenario, the implied value of the marginal tax rate at high income levels is close to 0, which is not surprising as in this case the government will choose not to harm high-income earners with high marginal tax rates, which are highly distortionary, and will instead impose a low marginal tax, which will benefit high-income earners, whose welfare is now important.

The framework of **Saez and Stantcheva** [**2013**], while general and flexible, leaves too much discretion as it does not put, intentionally, any discipline on the choice of g. In this paper, we are interested in identifying optimal marginal tax rates at high income levels. Specifically, we would like to depart from a conditional analysis and focus on a framework that allows for an endogenous determination of g. To achieve that we resort to the original framework of **Saez** [**2001**] and postulate that g be a function of y_L and y_H. Specifically, we set $y_H = \infty$ and assume that g is given with

$$g = \frac{\alpha}{(1-\tau_{ES})y_L},$$

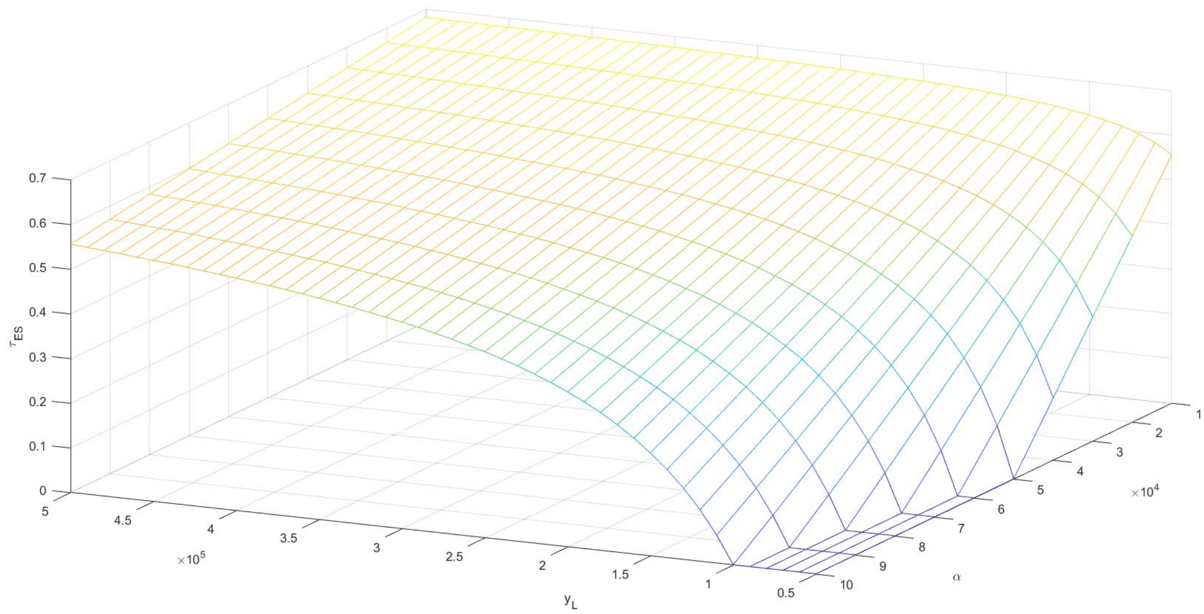
where α is a constant reflecting, in particular, the importance of government revenue to social welfare and y_L denotes the lower bound of income for which, we believe, the formula of **Saez** [**2001**] becomes applicable. Observe that by following **Saez** [**2001**] and defining g with equation (16) we acknowledge that the welfare of an individual is dependent on the net income earned by the individual and more importantly we assume that the importance of high-income earners decreases with actually earned income. Under such assumptions, fully consistent with **Saez's** framework [**2001**], the optimal formula of the marginal tax rate at high income levels takes the form

$$\tau_{ES} = \frac{1-g}{1-g+m\eta} = \frac{1-\frac{\alpha}{(1-\tau_{ES})y_L}}{1-\frac{\alpha}{(1-\tau_{ES})y_L}+m\eta}.$$

We choose to assume that m is roughly equal to 2.2 and that η is equal to 0.2. Under such assumptions, the actual optimal tax rate as a function of y_L is depicted in Figure (9).

Again, it is apparent that even in this scenario the optimal marginal tax rate remains substantial, especially if one is willing to assume that marginal social weights applied to individuals with income levels below PLN 30,000 should be substantial, i.e., for values of α roughly smaller than PLN 30,000. In such a case, the marginal tax rates applicable to high-income earners vary in the range between 27% and 68%, with the majority of values exceeding 50%.

Figure 13. Optimal marginal tax rates depending on α and y_L



Source: Authors' own elaboration.

In summary, we can conclude that the fundamental properties of the income distribution in Poland, i.e., the income distribution being of Pareto shape, make the approach of **Saez [2001]** admissible and his formula applicable to the case of Poland. Unfortunately, our numerical estimates suggest that the current policy put in place in Poland fails to be optimal. In fact, we show that the optimal tax rates at high income levels should be much higher than those currently observed. Specifically, we argue that the marginal tax rates could be as high as 60% for individuals whose income exceeds PLN 50,000. Thus, we conclude that there is room for policy adjustments in Poland.

Conclusions

How much should high-income earners pay in taxes? This question is subject to endless debates. In fact, most individuals would agree that high-income earners should pay more than low-income earners. However, this is where the agreement probably ends. Specifically, there is no universal agreement on the values of the marginal tax rates. Still, many argue that marginal tax rates should be increasing with income, others claim that paying the same fraction of income ensures fairness, while the optimal tax theory in its purest form, when the distribution of skills is bounded from above, implies, at least at the high income level, that the optimal tax schedule is in fact regressive, with the marginal tax rates falling all the way to zero as observed earnings increase. Only in the case of unbounded distribution of skills, as shown by **Diamond and Saez [2011]**, can optimal marginal tax rates remain positive indefinitely. Consequently, only an objective and quantitative analysis can lead us to a meaningful answer to the question of optimal taxation. Unfortunately, finding global solutions to optimal taxation problems can be challenging as informational and computational constraints can be overwhelming. Thus, it may be necessary to limit attention to partial improvements and focus on solutions to optimal taxation problems that are applicable at least locally.

In this paper, we present solutions to the optimal taxation problem applicable to high-income earners in Poland. Our findings are based on the approach of **Saez [2001]**, who came up with a practical formula that can be used to pin down the values of the optimal tax rates. The formula is simple, but it is not universally applicable as it is only valid when the distribution of skills in society is thick-tailed. However, the distribution of income, and hence skill, in Poland, is in fact of Pareto form at the high end. Specifically, by using data obtained from the Polish tax authority, we are able to characterise the distribution of income in the upper tail and show that the distribution can be approximated with a Pareto distribution with parameter *m* in the range [2.2,3.44] which facilitates the applicability of the formula of **Saez [2001]**. In that sense, the distribution of income in Poland resembles that in the United States and, just like in the case of the USA, the formula of **Saez [2001]** implies that the optimal marginal tax rates in Poland can be high. Specifically, we establish that, depending on the value of the elasticity of the labour supply with respect to the wage, the optimal marginal tax rates can be anywhere between 15% and 70% with lower elasticities leading to higher marginal tax rates. Moreover, if one is willing to accept empirically viable estimates of elasticity, one is bound to conclude that the optimal marginal tax rates should be high and equal to about 60% for earners with incomes in excess of PLN 50,000, which corresponds to the 83 percentile. In other words, our findings suggest that the current tax schedule is suboptimal and that there is a need for at least a partial tax reform.

The results presented in this paper are robust and align with the corresponding findings in the case of the United States. However, we must remember about the limited applicability of the results presented in this paper as they pertain only to high-income earners. Moreover, our analysis is confined in scope as we only consider single earners. We omit the labour self-employment trade-off and shun away from potential migration issues. Without a doubt, additional studies are needed to further our understanding of optimal taxation issues in Poland. We have undertaken initial steps in that direction and have searched for optimal tax schedules that are applicable globally and have solved the original problem of **Mirrlees [1971]** in the context of the Polish economy, with the results summarised in a companion paper – see **Dudek et al. [2021]**.

References

Atkinson A., Piketty T., Saez E. [2011], Top Incomes in the Long-Run of History, *Journal of Economic Literature*, 49 (1): 3–71.

Bargain O., Orsini Ch, Peichl A. [2014], Labor Supply Elasticities in Europe and the USA, *Journal of Human Resources*, 49 (3): 723–838.

Berg A., Ostry J. D., Tsangarides Ch. G. [2014], *Redistribution Inequality and Growth*, IMF Discussion Papers.

Chrostek P., Klejdysz J., Korniluk D., Skawiński M. [2019], *Wybrane aspekty systemu podatkowo-składkowego na podstawie danych PIT i ZUS 2016*, MF Opracowania i Analizy, Ministerstwo Finansów, Warszawa.

Diamond P. [1998], Optimal Income Taxation: An Example with a U-Shaped Pattern for the Optimal Marginal Rates, *American Economic Review*, 88: 83–95.

Diamond P., Saez E. [2011], The Case for a Progressive Tax: From Basic Research to Policy Recommendations, *Journal of Economic Perspectives*, 25 (4): 165–190.

Doligalski P. [2019], *Gdzie jest szczyt krzywej Laffera dla najwyższych dochodów w Polsce*, Dobrobyt na Pokolenia, Raport naukowy.

Dudek M. K. [2013], *On Efficient Taxation*, available at SSRN: http://ssrn.com/abstract=2266690 or http://dx.doi.org/10.2139/ssrn.2266690.

Dudek M. K. [2016], *Probabilistic Lump-Sum Taxation*, available at SSRN: https://ssrn.com/abstract=2601764 or http://dx.doi.org/10.2139/ssrn.2601764.

Dudek M. K., Dudek P., Walczyk K. [2020], *Practical Solutions to the Problem of Mirrlees – Optimal Labor Income Taxation in Poland*, available at SSRN: https://ssrn.com/abstract=3621764.

Dudek M. K., Dudek P., Walczyk K. [2021], *Optimal Labor Income Taxation in Poland*, Wydawnictwo AFiB Vistula, Warszawa.

Ebert U. [1992], A Reexamination of the Optimal Nonlinear Income Tax, *Journal of Public Economics*, 49: 47–73.

Guu S.-M., Judd K. [1997], Asymptotic methods for aggregate growth models, *Journal of Economic Dynamics and Control*, 21 (6): 1025–1042.

Kaplow L. [2010] Concavity of Utility, Concavity of Welfare, and Redistribution of Income, *International Tax and Public Finance*, 17: 25–42.

Kleven H., Landais C., Saez E. [2013], Taxation and International Mobility of Superstars: Evidence from the European Football Market, *American Economic Review*, 103 (5): 1892–1924.

Kopczuk W. [2005], Tax Bases, Tax Rates and the Elasticity of Reported Income, *Journal of Public Economics*, 89 (11–12): 2093–2119.

Kopczuk W. [2012], *The Polish business 'flat' tax and its effect on reported incomes: a Pareto improving reform*, Columbia University Working Paper.

Kydland F., Prescott E. [1982], Time to Build and Aggregate Fluctuations, *Econometrica*, 50: 1345–1371.

Lollivier S., Rochet J.-Ch. [1983], Bunching and Second Order Conditions: A Note on Optimal Tax Theory, *Journal of Economic Theory*, 31: 392–400.

Mankiw N. G. [2013], Defending the One Percent, *Journal of Economic Perspectives*, 27 (3): 21–34.

Mirrlees J. A. [1971], An Exploration in the Theory of Optimal Income Taxation, *Review of Economic Studies*, 38: 175–208.

Myck M. [2017], *Modelling Labour Supply in Poland: Elasticity Estimates and Policy Simulations*, IBS Seminar.

Myck M., Trzciński K. [2019], From Partial to Full Universality: The Family 500+ Programme in Poland and its Labor Supply Implications, *ifo DICE Report, ifo Institute – Leibniz Institute for Economic Research at the University of Munich*, 17 (03): 36–44.

OECD [2011a], *The Taxation of Low-Income Workers*, Chapter 2 in OECD Tax Policy Study No. 21: Taxation and Employment, OECD, Paris.

OECD [2011b], *The Taxation of Mobile High-Skilled Workers*, Chapter 4 in OECD Tax Policy Study No. 21: Taxation and Employment, OECD, Paris.

Piketty T. [2014], *Capital in the Twenty-First Century*, Harvard University Press.

Piketty T., Saez E., Stantcheva S. [2014], Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities, *American Economic Journal*: *Economic Policy*, 6 (1): 230–271.

Revesz J. T. [1989], The Optimal Taxation of Labour Income, *Public Finance/Finances Publiques*, 3: 453–475.

Ruiz del Portal X. [2007], The Problem of Optimum Income Taxation: A Remark on Its Monotonicity Constraint, *Journal of Public Economic Theory*, 9: 265–283.

Ruiz del Portal X. [2008], Is the Optimal Income Tax Regressive?, *Economics Letters*, 100: 402–404.

Ruiz del Portal X. [2010], On the Qualitative Properties of the Optimal Income Tax, *Mathematical Social Sciences*, 59: 288–298.

Sadka E. [1976], On Income Distribution, Incentive Effects and Optimal Income Taxation, *Review of Economic Studies*, 43 (1): 261–268.

Saez E. [1999], *A Characterization of the Income Tax Schedule Minimizing Deadweight Burden*, MIT Ph.D. thesis (chapter 3).

Saez E. [2001], Using Elasticities to Derive Optimal Income Tax Rates, *Review of Economic Studies*, 68: 205–229.

Saez E., Stantcheva S. [2013], Generalized Social Marginal Welfare Weights for Optimal Tax Theory, *American Economic Review*, 106 (1): 219–246.

Seade J. K. [1977], On the Shape of Optimal Tax Schedules, *Journal of Public Economics*, 7 (1): 203–236.

Sen A. K. [1973], *On Economic Inequality*, Clarendon Press, Oxford.

Tuomala M. [1990], *Optimal Income Tax and Redistribution*, Clarendon Press, Oxford.

Tuomala M. [2010], On Optimal Non-linear Income Taxation: Numerical Results Revisited, *International Tax and Public Finance*, 17: 259–270.

Zawisza T. [2019], *Optimal Taxation of Employment and Self-Employment: Evidence from Poland and Implications*, IFS Working Paper.