The optimal top marginal tax rate: Application to Hungary

Áron Kiss, Magyar Nemzeti Bank (Central Bank of Hungary)

Abstract

The paper applies recent developments in the theory of optimal income taxation to the Hungarian personal income tax system. The main conclusion is that the optimal top marginal tax rate in Hungary is likely to be higher, perhaps substantially, than the actual rate. It is discussed how this result depends on the parameters describing labour-supply behaviour, the income distribution, and the redistributive preferences of society.

JEL classification

Keywords

Acknowledgements

Introduction

There was hardly a year in the last decade that the Hungarian personal income tax (PIT) remained unchanged. The last three years saw radical changes, including the introduction in 2011 of a flat tax rate. The current wave of reforms is expected to be complete in 2013 when the single PIT rate is 16 percent. The reforms benefited different groups in different years but altogether they brought a radical cut in the marginal (and average) tax rates of high-income earners. At the same time, the elimination of the Employee Tax Credit (adójóváírás) in 2012 brought an increase in the average tax rate of low and middle income earners without children.

Amid such frequent and radical changes it is important to ask what, if anything, economic theory can say about the desired characteristics of the income tax system. This paper builds on recent developments of the theory of optimal income taxation and applies one of its main results to the Hungarian tax system.[1]

The foundational work of the theory of optimal income taxation is by Mirrlees (1971). At the core of the theory is the insight that while society (or a government making policy to implement the preferences of society[2]) would like to redistribute income from high earners to low earners, redistribution dampens the work incentives of both high and low earners. The optimal tax policy thus reflects a balance between redistribution and incentives or, in other words, between the principles of 'equity' and 'efficiency'. While Mirrlees' work was very influential in economic theory, results derived from optimal income tax theory were not seen as being of much practical guidance to tax policy. This changed with the work of Saez (2001). Saez, building on work by Diamond (1998), was able to express the results of the theory as functions of estimable parameters. The relevant parameters include those that describe the shape of the income distribution and those that describe how sensitively people's earnings react to changes in the tax rates.

Even after these developments it should not be expected that the theory gives unequivocal answers to all tax policy questions. One reason is that the optimal income tax schedule depends on the revenue to be achieved, which in turn depends on the desired level of government expenditures and on how distortionary alternative taxes are relative to the PIT. The other important reason is that the optimal tax system depends on the strength of the redistributive preferences of society. An applied optimal taxation model needs those preferences as inputs to derive optimal tax rates. The theory of optimal income taxation can provide useful qualitative guidance for tax policy because some of its results do not depend on these important (but uncertain) factors. It is even possible, under relatively unrestrictive assumptions, to derive even quantitative guidance with respect to the top marginal tax rate.

This paper follows two recent studies that provide examples of how to use the theory of optimal taxation to derive recommendations for tax policy. The first of these studies, by Brewer et al. (2010) was prepared in the framework of the Mirrlees Review, a detailed review of the UK tax system by an international group of researchers coordinated by the Institute for Fiscal Studies. The second study, by

---

[1] An earlier version of the paper appeared as a discussion paper (Kiss 2013). That version had a somewhat broader focus, also surveying theoretical results about the optimal lower tax rates. Some results of the analysis have been discussed in the survey article of Benczúr, Kiss and Mosberger (2013).

[2] In the paper 'society' and 'government' is used interchangeably. The analysis abstracts from mechanisms of policy formation and the political process and adopts the approach of classical public economics in which society (or the government) is searching for the best solution for a given problem. Clearly, the political economy aspect is important in the real world but it is not in the focus of this paper.

Diamond and Saez (2011), describes three broad principles that can be derived from the theory and discusses them in the context of current US tax policy.[3]

This paper applies one result of the literature to the Hungarian context, suggesting that high marginal tax rates for top earners can be optimal. The result is based on the observation, made by Saez (2001), that the optimal top marginal tax rate can be expressed, under general assumptions, by just three parameters: a parameter describing the shape of the top of the income distribution, a behavioural elasticity expressing how sensitively high-income taxpayers react to changes in the marginal tax rate and, finally, a parameter that expresses the social value of an additional dollar kept by a top earner, expressed in terms of public funds. The first two of these parameters can be estimated, while the third parameter is the function of the preferences of society.

Brewer et al. (2010) and Diamond and Saez (2011) argue that the social marginal value of an additional dollar kept by a top earner (an individual belonging to the top 1 percent of earners in their definition) is close to zero. This can be valid on a utilitarian basis (where the welfare of every citizen is equally important to the government) if an extra dollar of income adds to the welfare of a high earner much less than to the welfare of a low earner (or, more precisely, if the marginal utility of consumption declines to zero at very high incomes). In this case, the optimal top marginal tax rate is equal to the top rate that maximises government revenue.

The utilitarian interpretation is appealing because then the question of optimality is viewed from the hypothetical perspective of a not-yet-born individual, 'from behind the veil of ignorance', who expects her earning capacity to be a random draw from the empirical income distribution. This individual prefers a tax system that maximises her expected utility over the possible levels of earning capacity she could be assigned in the 'birth-lottery'.

The present paper derives the formula for the optimal top marginal tax rate and evaluates it for Hungary for alternative values of the social-marginal-value parameter (zero as well as non-zero). The calculations suggest that the revenue-maximising top marginal tax rate is higher than the current top marginal tax rate: the revenue-maximising top marginal PIT rate is estimated to be about 30 percent (or about 40 percent if pension contributions were capped). It is discussed how the optimal top marginal tax rate may depend on the definition of top incomes and on the social-marginal-value parameter. The optimal top marginal tax rate is shown to be lower, although not dramatically, than the revenue-maximising rate if plausible non-zero values are chosen for the parameter expressing the social marginal value of a dollar kept by a top earner. Fiscal effects of hypothetical tax reforms that are consistent with these results are simulated.

The rest of the paper is organised as follows. The next section derives the theoretical formulae for the revenue-maximising and optimal top marginal tax rates. Section 3 evaluates the theoretical benchmarks in the Hungarian case. Section 4 compares these benchmarks to recent actual top marginal tax rates in Hungary, derives conclusions for policy, and surveys alternative considerations not taken into account in the baseline analysis. Section 5 concludes.

The optimal top marginal tax rate: Theory

The surprising result that optimal tax theory can be used to give quantitative guidance about the top marginal tax rate was first derived by Saez (2001). The methodology has since been used to analyse whether actual top marginal tax rates in the US (Diamond and Saez, 2011) and the UK (Brewer et al., 2010) are in line

---

[3] In recent papers, similar approaches have been applied to Germany by Bach et al. (2012) and Finland by Riihelä et al. (2013).

with the optimal tax rates derived from theory. This section follows these studies to derive the simple formula for the optimal top marginal tax rate, evaluates the formula for the Hungarian tax system, and compares the resulting tax rate with actual top marginal tax rates of recent years. The presentation of the theoretical background follows Brewer et al. (2010).

Before setting out to derive the formula for the optimal top marginal tax rate, it should be clarified what is meant by 'top' incomes. Diamond and Saez (2011) defines top incomes as the top 1 percent of the income distribution. In this paper it is be explored how conclusions depend on whether top incomes are defined as the top 1 percent (starting at annual income level of HUF 10.6 million in Hungary in 2008, or about EUR 35,000);[4] or the top 5 percent (starting at HUF 5.3 million, or about EUR 18,000); or the top 10 percent (starting at about HUF 3.8 million, or about EUR 13,000).

Consider an economy where individuals have different earning capacities. Individual i earns a gross income z_i. Of this gross income the individual pays $T(z_i)$ in taxes and consumes the rest, $c_i = z_i - T(z_i)$. Individuals value consumption and leisure and can adjust their hours worked (or, more generally, any aspect of their labour effort) as a response to the rate of exchange between the two. How much the individual can consume in exchange for an additional hour worked depends on the marginal tax rate $\tau(z) = T'(z)$.

The government sets tax rates and transfers in a way as to maximise social welfare while reaching an exogenously given level of net revenue required for the provision of public goods. (As a simplification it is often assumed that the sum of tax revenue has to be equal to the sum of transfers paid, that is, the exogenous level of other public spending is zero.) The marginal weight of an individual in the social welfare function is $g(z)$: this expresses the value society attaches to an additional dollar consumed by an individual with gross income z, expressed in terms of public funds. If the state has redistributive preferences, then $g > 1$ for low earners (a one-dollar increase in their consumption is worth more than one dollar for the government) while $g < 1$ for high earners (a one-dollar increase in their consumption is worth less than a dollar for the government).

The trade off facing the government is this: it values redistribution but redistribution dampens individuals' incentives to work. If individuals work less because of the tax system, revenues decrease and less redistribution can be achieved. This trade off between equity and efficiency is at the centre of the theory of optimal income taxation.

To see this trade off in specific, consider an increase $d\tau$ in the marginal tax rate facing top income earners, i.e., the N individuals earning an income higher than $\dot{z}$. This will affect social welfare in three ways: (1) tax revenue increases mechanically and that is a social gain; (2) the increased tax burden makes those affected worse off and that is a social loss; (3) those affected will reduce their work effort and with that their taxes payable and that is again a social loss.

The first effect is thus the mechanical effect on tax revenue: all individuals earning more than $\dot{z}$ will pay more taxes than before. In the present example there are N taxpayers earning more than $\dot{z}$. Let their average gross income be z^m. The tax increase affects their income above the threshold $\dot{z}$. Thus, the mechanical effect on tax revenue is:

$dM = N[z^m - \dot{z}]d\tau > 0$.

---

[4] A conversion rate of 1 EUR = 300 HUF is used throughout the paper. This rate prevailed during most of 2012 and 2013. The rate was about 270 to 280 from 2008 to 2011, and about 250 in the years before 2008.

Note that this effect is defined as the change in tax revenue before any behavioural change occurs on the part of the individuals affected.

The second effect is the direct welfare loss of those who have to pay more taxes. In the present example the welfare effect is given by:

$dW = -g \cdot N[z^m - \dot{z}]d\tau < 0.$

where g is the average social marginal value of consumption of individuals earning more than $\dot{z}$.

The last is the behavioural effect: the increase of the marginal tax rate induces high earners to decrease their work effort which results in a fall in tax revenues. The decrease in tax revenues is $dB = dz \cdot \tau \cdot N$ where dz is the average change of income for individuals affected by the tax increase. The empirical studies estimating behavioural responses to tax changes estimate a parameter given by the following expression:

$e = (dz/z) \cdot [(1 - \tau)/d(1 - \tau)].$

As can be seen from this expression, parameter e is an elasticity: it measures the percentage change of reported income as a response to a 1 percent change of the marginal net-of-tax rate $(1 - \tau)$, which is the share of the last unit of gross income that the individual can take home as net income. Expressing dz from this formula and substituting into the definition of dB, we get

$dB = -N \cdot e \cdot z^m \cdot d\tau \cdot \tau/(1 - \tau) < 0.$

At the optimal marginal tax rate τ^* the sum of these effects must be equal to zero. If the welfare effect of a small tax increase were positive (negative), the government would want to increase (cut) the tax rate further; thus the initial tax rate could not have been optimal.

From this argument it follows that the equation $dM + dW + dB = 0$ implicitly determines the optimal top tax rate. Introducing the parameter $a = z^m/(z^m - \dot{z})$ we can solve the equation to reach a simple formula:

$\tau^* = (1 - g)/(1 - g + e \cdot a).$

The optimal top marginal tax rate thus depends on three parameters, two of which can be estimated: parameter e is the elasticity of taxable income with respect to the net-of-tax rate, while parameter a characterises the shape of the income distribution.

Since the third parameter g is a function of society's preferences, it is less straightforward to assess its plausible values. An upper bound to the top marginal tax rate can be obtained by considering the case when $g = 0$. In this case the value society attaches to an additional dollar kept by a top earner is negligible compared to the value society attaches to an additional dollar kept by the average earner (or to an additional dollar of government revenue). Then, the only force keeping the marginal tax rate of top earners from rising is the behavioural effect. In this case the optimal top marginal rate is equal to the revenue-maximising rate, with the formula simplifying to:

$\tau^* = 1/(1 + e \cdot a).$

Brewer et al. (2010) and Diamond and Saez (2011) argue that $g = 0$ is plausible for top earners (the top 1 percent of the income distribution in their definition). For most social welfare functions with redistributive preferences it will be the case that g decreases with income, and the zero-marginal-weight result will hold asymptotically for social welfare functions that satisfy the property $\lim g(z) = 0$ as z goes to infinity. This is the case, for example, in a utilitarian framework (where the welfare of every individual is equally important for the government) where the marginal utility from consumption declines to zero. In all these cases it will be true that the narrower the top income bracket is defined, the higher the optimal top

marginal tax rate is and the closer it is going to be to the revenue-maximising tax rate.

But parameter *g* does not have to converge to zero for the revenue-maximising tax rate to be approximately optimal. Note that *g* enters (with the same sign) both the numerator and the denominator of the general formula. This means that the effect of *g* will be of second order as long as *g* is not too large. A sensitivity analysis to the value of *g* is presented in the next subsection.

Evaluating the theoretical benchmarks for Hungary

The revenue-maximising top marginal tax rate in Hungary

The previous subsection established that the revenue-maximising top marginal tax rate depends on only two parameters: parameter *a* describing how thin the income distribution is at the top and parameter *e* describing how sensitively top earners react to changes in the marginal tax rate. In this subsection, estimates for parameters *a* and *e* are presented. With these parameters it is possible to calculate the revenue-maximising top marginal tax rate. As seen above, the revenue-maximising top marginal tax rate is optimal when the value of parameter *g* is zero. The next subsection contains the more general case, calculating the value of the optimal top marginal tax rate as a function of parameter *g*.

The income-distribution parameter

First, the empirical value of parameter *a* is estimated for Hungary. Recall that parameter *a* is defined as $a = z^m/(z^m - \dot{z})$, that is, for any income limit $\dot{z}$, *a* is equal to the average income of individuals above the income limit divided by the difference of that average income and the income limit. A ten-percent sample of the 2008 PIT returns is used for this exercise, compiled by the Hungarian Tax Authority (the population excludes the full-time self-employed). Figure 1 shows the value of *a* for annual gross income levels between HUF 0.1 million (about EUR 330) and HUF 40 million (about EUR 130,000). For easier interpretation of the figure it can be noted that the average annual gross income in 2008 was about HUF 1.9 million (about EUR 6,300) while an individual belonged to the top 1 percent of tax filers with an annual gross income of about HUF 10.6 million (equivalent to about EUR 35,000).

Figure 1. The value of parameter a as a function of the income limit chosen
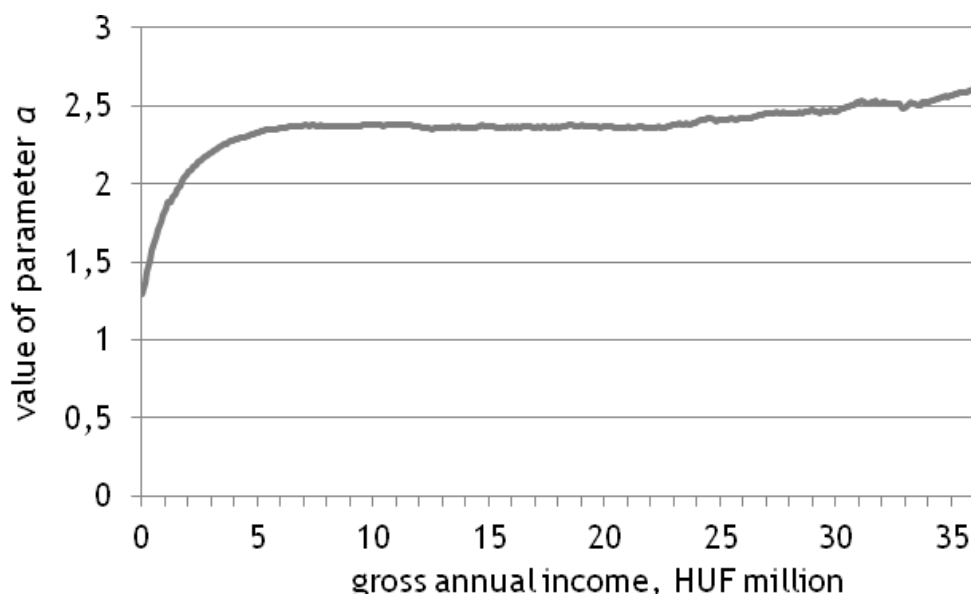
Note: Author's calculation based on 2008 tax return data provided by the Hungarian Tax Authority.

Figure 1 shows that the value of parameter *a* is very stable for income limits above HUF 5 million (about EUR 17,000). It is around 2.35 for incomes between HUF 6 and 23 million and is very close to 2.5 for income levels above that. (Note that individuals above income of HUF 5 million represent about the top 5 percent of the income distribution, while the top 1 percent starts at about HUF 10.6 million.)

That this parameter is stable for the upper part of the income distribution is a general result first noted by the Italian economist and statistician Vilfredo Pareto and confirmed for many countries and time periods ever since. In more intuitive terms it is equivalent with the statement that the average income of individuals earning an income higher than $\dot{z}$ is a constant multiple of $\dot{z}$. In the Hungarian tax data the average income of individuals earning more than $\dot{z}$ is about $1.7 \cdot \dot{z}$.

Based on Figure 1, a central estimate of *a* = 2.5 is adopted but the optimal tax rate formula is also evaluated for somewhat lower and higher values of parameter *a* (also called the Pareto parameter).

The taxable-income elasticity

The second parameter needed to calculate the optimal top marginal tax rate is *e*, the elasticity of taxable income with respect to the marginal rate, net-of-tax. This parameter first came into the focus of economic research with the estimations of Feldstein (1995). Before his work, labour economists estimated the effect of tax changes on hours worked and found low elasticities. It is since Feldstein's work that economists estimate the effects of tax changes on reported taxable income. Clearly, this measure encompasses more than the change in hours worked: it can reflect changes in work intensity, change of jobs, moonlighting, but also tax avoidance and evasion.

This list of possible factors behind the elasticity makes clear that it does not necessarily reflect changes of real economic activity. Nevertheless, it is the welfare-relevant measure of the behavioural effect. This is because, unless there are significant externalities between various tax bases, the elasticity is a direct measure of how government revenue changes in response to a change in the tax rate.

Since the work of Feldstein (1995) panel data with more years and individual observations became available which allowed more robust statistical methods to be applied. The elasticity of e = 0.4 estimated by Gruber and Saez (2002) is considered as representative of the newer literature for the US (see also recent surveys of the literature by Giertz [2004] and Saez et al. [2012]).

There are two studies that estimated the taxable income elasticity for Hungary. In the first such study, Bakos et al. (2008, henceforth BBB), used the elimination of the middle tax bracket in 2005 as the policy experiment to identify the elasticity. Since most of the tax changes in their data affected middle-income earners, most results of BBB reflect the taxable-income elasticity of middle-income individuals. Updated and re-estimated results of BBB have been surveyed, for a non-technical audience, by Benczúr et al. (2013). The re-estimation resulted, in general, in lower estimated elasticities. One income range in which the results were robust to the re-estimation is the range between HUF 1.5 to 1.95 million (about EUR 5,000 to 6,500 at 2013 conversion rates, which was slightly above the average wage in 2004; see Benczúr et al., 2013, Table 2.4). In this income range BBB find an elasticity of about e = 0.12, together with a significant, although moderate income effect. (How optimal tax rates depend on the income effect is discussed in Subsection 2.4.)

The second study of the taxable-income elasticity for Hungary was conducted by Kiss and Mosberger (2011). They use the introduction of an extraordinary tax on high-income individuals in 2007 as their policy episode, thus they focus on the group of individuals most relevant for the present purposes (their main specification includes individuals earning HUF 5 to 8 million in 2005 or about EUR 17 to 27 thousand at 2013 conversion rates). The main estimates for the elasticity fall between 0.15 and 0.2. The present calculation of the optimal top marginal tax rate uses this estimate (e = 0.2) as a benchmark but the formula is evaluated at somewhat lower and higher values as well. The income effect estimated by Kiss and Mosberger (2011) is relatively large in magnitude, but not statistically significant in the main specification.

This central estimate is in line with most estimates for countries outside the US.[5] Some observers attributed the higher estimated US elasticities at least partly to tax optimisation strategies (timing and form of compensation, among others) that are available to executives in the US (Goolsbee 2000). This consideration shows that the elasticity of taxable income is not an immutable parameter but can be influenced by tax policy.

The revenue-maximising top marginal tax rate

Based on these parameter values one can evaluate the formula for the revenue-maximising top marginal tax rate. Table 1 shows the value of the tax rate as a function of parameters a and e, with g = 0. The table shows that the revenue-maximising top marginal tax rate of high earners is 67 percent at the central parameter estimates (a = 2.5 and e = 0.2). Of course, this tax rate cannot be directly compared to the actual PIT rates, since taxpayers also pay social security contributions (SSC) and consumption taxes. Comparable actual top marginal tax rates for the last few years in Hungary are calculated in the next subsection.

---

[5] For references to elasticity estimates in other countries, see Benczúr et al. (2013).

Table 1: The revenue-maximising top marginal tax rate as a function of a and e

values of a	values of e			
	0.1	0.2	0.3	0.4
2	83%	71%	63%	56%
2.5	80%	67%	57%	50%
3	77%	63%	53%	45%

Note: Author's calculations.

Table 1 also shows how, within the plausible range of parameter estimates, the revenue-maximising top marginal tax rate depends on the parameters. If the taxable-income elasticity were 0.1 or 0.3 rather than 0.2, this would change the optimal top marginal tax rate by about 10 percentage points. At the same time, if the value of the Pareto parameter were equal to 2 or 3 instead of 2.5, this would change the optimal top marginal tax rate by 4 percentage points. The intuition behind the effect of this parameter is the following. A smaller *a* implies a fatter tail of the income distribution, which in turn implies that the behavioural effect will be more significant relative to the behavioural effect (the revenue effect is proportional to $z^m - \dot{z}$, while the behavioural effect is proportional to z^m).

The optimal top marginal tax rate

But how sensitive is the optimal top marginal tax rate to the value of parameter *g*? To be able to answer this question, first let us take a simple theoretical benchmark for guidance as to what are sensible values of *g*.

A popular benchmark, and one that is also used by Diamond and Saez (2011), supposes that the social marginal value of consumption is inversely proportional to income, i.e., $g(z) = 1/z$. Such a social marginal welfare function could be a result of social preferences that favour redistribution per se, or it could be a result of utilitarian preferences if the welfare of the individual is a logarithmic function of income. In the utilitarian framework the social marginal welfare weight of high-income individuals is low not because their welfare is less important to the government (or to society) but because an additional dollar is worth less for them than for lower-income individuals.

Tax return data from 2008 are used to calculate the average income of the top 1 percent, the top 5 percent, and the top 10 percent of the income distribution before calculating the value of parameter *g* for these groups according to the formula $g(z) = 1/z$. Table 2 shows the results of this exercise.

Table 2: Average income of top earners and calculated values of parameter g.

Income range	Income limit (HUF million)	Average income (HUF million)	Value of g in example
top 1%	10.6	18.3	0.11
top 5%	5.3	9.3	0.21
top 10%	3.8	6.8	0.28

Note: Author's calculations. In the last column, g is calculated as the inverse of the ratio of the average income of the respective top income group to the overall average income (HUF 1.91 million) in 2008.

Table 2 shows that in 2008 an individual belonged to the top 1 percent of tax filers if her income exceeded HUF 10.6 million (about EUR 35,000). The average income of the top 1 percent was about HUF 18.3 million (about EUR 61,000). In the same year the average gross annual income was about HUF 1.9 million (about EUR 6,300). If the average social marginal value of income is inversely proportional to

income and the value of g is about 1 for the average earner then the value of g is about 0.1 for the top 1 percent.

Similarly, the lower limit of the top 5 percent was, in 2008, at about HUF 5.3 million (about EUR 18,000) while the calculated g is about 0.2. The lower limit of the top 10 percent was, in the same year, at about 3.8 million (about EUR 15,000) while the calculated g is about 0.3. Using these simple calculations as orientation the optimal top marginal tax rate can be calculated as a function of parameter g. Table 3 shows the results of such an exercise, with the value of parameter a held fixed at the central estimate of 2.5.

Table 3: The optimal top marginal tax rate as a function of g and e, with a = 2.5

values of g	values of e			
	0.1	0.2	0.3	0.4
0	80%	67%	57%	50%
0.1	78%	64%	55%	47%
0.2	76%	62%	52%	44%
0.3	74%	58%	48%	41%
0.4	71%	55%	44%	38%

Note: Author's calculations.

Table 3 shows that small changes of g have little effect on the optimal top rate. To take an example, choosing g = 0.1 instead of zero affects the optimal top marginal tax rate very little: for the central parameter values a = 2.5 and e = 0.2 the optimal top marginal tax rate becomes 64 percent instead of 67 percent. As it was shown above, if g is inversely proportional to income, then the value of g should be about 0.1 for the top 1 percent in Hungary. Thus this result means that it matters little whether, for the top 1 percent of earners, we set g = 0 or use $g(z)$ = 1/z as a reasonable approximation.

It follows from the calculations of Table 3 that the broader the top income bracket is defined, the lower the optimal top rate will be. In the numerical example above the marginal welfare weight of the top income group was calculated to be about 0.1 for the top 1 percent, about 0.2 for the top 5 percent, and about 0.3 for the top 10 percent. In this example the optimal top marginal tax rate is approximately 64 percent if it affects only the top 1 percent (applying to income above HUF 10.6 million at 2008 prices) but it is only 58 percent if it affects the top 10 percent (applying to income above HUF 3.8 million at 2008 prices).

Comparing the theoretical benchmarks to actual tax rates

The actual top marginal rate in Hungary, 2005—2013

The actual tax rate that can be compared to the theoretical benchmark is not simply the top PIT rate. If there is a weak link between social security contributions (SSC) and future benefits received, then SSC can be assumed to have the same disincentive effect as the PIT. Ideally, the tax rate corresponding to the theoretical benchmark answers the following question: By how much can an individual increase her consumption if her total labour cost is increased by one unit? Thus, consumption taxes must also be taken into account. Based on these considerations, the effective marginal tax rate of top earners for Hungary is evaluated for recent years according to the following formula:

$\tau_{top} = 1 - (1 - \tau_{cons}) \cdot [(1 - \tau_{pit} - \tau_{ee})/(1 + \tau_{er})].$

Here, τ_{PIT} is the top PIT rate, τ_{ee} is the rate of employee contributions, τ_{er} is the rate of employer contributions and τ_{cons} is the effective tax rate on consumption. (The last term of the r.h.s. expression shows the ratio of net wage to total wage cost.)

Two issues regarding the tax rates merit further discussion.[6] The first issue is related to the employee contributions. The actual top marginal tax rate is calculated here for a hypothetical person with income above the pension contribution ceiling. Until 2013, individuals with income above the ceiling did not have to pay the employee pension contribution, which reduced their marginal effective tax rate by about 10 percent. The ceiling has in recent years been between the 95th and the 99th percentile of the income distribution, which means that individuals at the 99th percentile faced a lower marginal tax rate than individuals at the 95th percentile. In 2013 the pension contribution ceiling was abolished. For our purposes it complicates matters somewhat that the abolition of the contribution ceiling will mean higher future pension benefits for current high-income individuals. If the relationship between additional contributions and additional future benefits is close, this measure shouldn't be seen as a tax increase at all. The issue of the relationship between contributions and benefits is discussed in subsection 4.3.
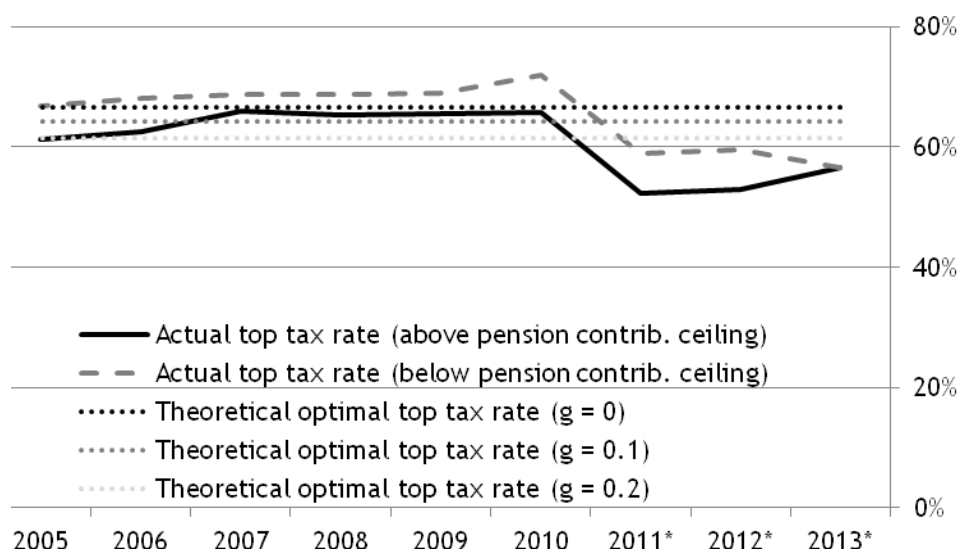
The second issue is the effective consumption tax rate. Ideally, we should be able to measure the tax share of the consumption basket of high-earning households. This is not the case in practice, however. Thus an approximation of the effective rate of consumption taxes has to be used. Similarly to Brewer et al. (2010, see online appendix, p. 3), the effective consumption tax rate is calculated here as government revenue from consumption taxes divided by total consumption from the National Accounts. Consumption taxes include excise duties on tobacco, alcohol products and gasoline, and further smaller items, besides VAT. Table A1 of the appendix shows, for selected years, the tax rates taken into account in the calculation. The table shows that the effective consumption tax rate is between about 15 to 16 percent in the period 2005 to 2010 (for the years after that the 2010 value was used for lack of data). Alternative methods to calculate the effective consumption tax rate would most likely result in similar values.

The actual top marginal tax rates that result from these calculations are shown in Figure 2, along with the optimal top marginal tax rate at different values of *g*. Until 2013 the actual top marginal tax rate is depicted both below and above the pension contribution ceiling. In 2013 the difference ceases to exist due to the abolition of the ceiling.

---

[6] Additionally, it should be noted that this approach abstracts from general-equilibrium effects and assumes that labour bears the burden of all labour taxation.

Figure 2. Actual effective top tax rates in Hungary, 2005-2013

It is apparent from Figure 2 that the gradual introduction of a flat tax of 16 percent between 2011 and 2013 has cut the top marginal tax rate significantly: from about 65 percent (above the pension contribution ceiling) and about 72 percent (below the ceiling) to about 56.6 percent in 2013 (see also Table A1 of the Appendix for the exact figures).[7]

Policy conclusions and simulations

Figure 2 suggests two conclusions about top tax rates in recent years. First, for individuals with income above the pension contribution ceiling, the marginal tax rate before 2010 was around the revenue-maximising tax rate, while the 2011 tax cut resulted in marginal tax rates that were much lower. Second, individuals just below the pension contribution ceiling faced marginal tax rates in the period 2005 to 2010 that were above the revenue-maximising top marginal tax rate.

The first conclusion means that, if the marginal social weight of the highest earners is close to zero, the tax revenue foregone by the post-2010 tax cuts for top earners cannot be justified by the behavioural effect: the estimated taxable income elasticity is not large enough to motivate a PIT rate of 16 percent. The calculated optimal top marginal tax rate is consistent, after the abolition of the pension-contribution ceiling, with a top PIT rate between 24 percent (about optimal under *g* = 0.2) and 31 percent (about optimal under *g* = 0). If the pension contribution ceiling had not been abolished, the calculated optimal top marginal tax rate would be consistent with a top PIT rate between 34 percent (about optimal under *g* = 0.2) and 41 percent (about optimal under *g* = 0).

For the top PIT rate of 16 percent to be optimal, g has to be around 0.35 if the taxable income elasticity is 0.2. Alternatively, if *g* = 0, a top PIT rate of 16 percent is optimal if the elasticity parameter *e* = 0.3, much higher than our central estimate

---

[7] Because of the lack of contemporaneous data, the calculations of Figure 2 do not take into account minor increases in consumption taxes that came into effect in 2011 and 2012 (the most significant among these is the increase from 25% to 27% of the base VAT rate in 2012). These would probably not add more than 1 or 2 percentage points to the top marginal rate, however. Thus they would not affect the results qualitatively.

(this PIT rate corresponds to an effective top marginal rate of 57 percent in Table 3, row 1).

The second conclusion (that individuals with income just below the pension contribution ceiling faced a marginal tax rate above the revenue-maximising top marginal tax rate) shows that, at this point of the analysis, the definition of the 'top income range' becomes important. Even though the theoretical analysis suggests that at high incomes the marginal tax rate should be increasing, individuals below the pension contribution ceiling were facing a higher marginal tax rate than those above the ceiling in recent years (the marginal tax rate schedules for incomes above HUF 3 million [or about EUR 10,000] for selected years between 2005 and 2013 are shown in Figure A1 of the Appendix). Very high marginal tax rates are justifiable at income levels where the taxable-income elasticity is thought to be close to zero (as it might be the case around the average wage), but this is not the case, according to the estimations of Kiss and Mosberger (2011), in the income range just below the pension contribution ceiling.

In the following, two policy simulations are presented to show examples of tax policy measures that are consistent with the theoretical results of this paper. The examples were chosen to be as simple as possible and clearly they are not the only conceivable policy measures that are consistent with the theoretical results. Thus it is not implied here that these examples constitute 'the optimum'. It is also not attempted here to design a whole new PIT system: the example measures affect only the top part of the income distribution.

The simulations were conducted in the microsimulation model described by Benczúr et al. (2011, 2012). The model takes into account the behavioural responses to taxation by incorporating the taxable-income elasticity of 0.2 estimated by Kiss and Mosberger (2011), and it also takes into account general-equilibrium effects through a simple neo-classical macroeconomic model. This means that if individuals change their labour supply as a response to changes in the tax system, the general-equilibrium model calculates how this affects wages and the stock of capital. While the labour-supply response of the individuals is thought to occur relatively quickly, the dynamic macroeconomic effects are long-run effects since the adjustment of the capital stock takes time. Generally, full long-run macroeconomic adjustment takes place within a decade.

Table 4 summarises the results of the simulations. For each simulation, a 'static fiscal effect' is shown first: this says by how much government revenue would increase from the PIT and SSC if the given measure were adopted, absent behavioural change. The 'dynamic fiscal effect,' in turn, takes into account both the behavioural response to the tax change by the taxpayers and the long-run macroeconomic adjustments. Additionally, the ratio of the dynamic and static effects is displayed.

The two simple measures analysed are: (1) the introduction of a PIT rate of 31 percent above HUF 10 million (about EUR 33,000 or roughly 1 percent of taxpayers); (2) the introduction of a PIT rate of 24 percent above HUF 5 million (about EUR 17,000 or roughly 5 percent of taxpayers). All tax policy measures are compared to a baseline of a flat PIT of 16 percent, as effective in 2013, after the abolition of the pension contribution ceiling.

Table 4: Hypothetical tax policy scenarios

		Static fiscal effect (HUF billion)	*Dynamic fiscal effect (HUF billion)*	*Ratio of dynamic to static*
(1)	PIT rate of 31% above HUF 10 million	51	23	45%
(2)	PIT rate of 24% above HUF 5 million	72	39	54%

Note: The 2013 tax system was used as benchmark. Revenue is estimated at 2013 prices. Fiscal estimates include effects on the PIT and SSC. The dynamic simulations are based on the central taxable-income-elasticity estimates by Kiss and Mosberger (2011).

Table 4 shows that the introduction of a 31 percent tax bracket for the top 1 percent would increase government revenue by about HUF 51 billion under no behavioural change, and by about HUF 23 billion after behavioural effects and macroeconomic feedback effects are taken into account. The effects of a lower top rate levied on a broader base (24 percent on income exceeding HUF 5 million) are larger by about one-half. While the modelling strategy of the microsimulation model does not conform fully to the assumptions made in the approach followed in this paper, the simulations seem to confirm that the higher the top marginal tax rate is increased, the higher the behavioural effect (and with that the income lost) as a proportion of the static effect. This illustrates the conflict between equity and efficiency that is in the core of the optimal taxation problem.

Alternative considerations

This subsection looks at some considerations that might qualify or modify the conclusions based on the benchmark results.

The shape of the income distribution

The optimal top marginal rates depend on the shape of the income distribution. Mankiw et al. (2009) find that while the optimal marginal tax rate at $150,000 is between 60 and 70 percent if a Pareto distribution is assumed, it is only about 40 percent with a lognormal distribution. Diamond and Saez (2011), in contrast, give detailed arguments about why calculations should be based on a Pareto distribution. The argument empirically turns on the constancy of parameter *a*. In the case of our data it can be confirmed that the empirical value of this parameter remains very stable even at the highest income levels where only dozens of individuals are observed in the 10 percent sample of Hungarian taxpayers. This is a strong empirical argument for the Pareto distribution.

Alternative welfare paradigms

This paper has reviewed arguments for the plausibility of a utilitarian approach which implies that the social marginal weight of high earners is close to zero if the marginal utility of consumption approaches zero at very high consumption levels. Alternatives to this approach exist. A 'Rawlsian' approach would perhaps assign a positive social marginal welfare weight to a group of low-earning (or low-welfare) individuals and a zero weight to the rest of society. Alternatively, the approach dubbed 'charitable conservatism' by Atkinson (1990) values the reduction of poverty, but does not value redistribution above the poverty line.[8] This approach is consistent with assigning a weight greater than 1 to 'the poor' and a positive weight smaller than 1 to the rest of society. The result that the current top marginal rate in

---

[8] Riihelä et al. (2013) show example calculations for this type of social welfare function besides the approach used in this paper.

Hungary is consistent with a social marginal weight of 0.35 for top earners could potentially be interpreted in this light.

How would the income effect change the results?

Throughout the analysis it was assumed that the marginal tax rate is the only aspect of the tax system that affects the behaviour of high-income earners. It is possible, however, that their behaviour is affected by the so-called income effect, too. The income effect is operative when an individual decreases her work effort after receiving a transfer that does not affect her marginal incentives.

If such an effect were indeed operative, then the gain from the behavioural effect would be dampened in response to a cut in the top marginal tax rate. To see this, suppose that the marginal tax rate is cut by $\Delta\tau$ for those with an income above $\dot{z}$. This tax change would affect an individual with income $z' > \dot{z}$ in two separate ways. First, the cut in the marginal tax rate would have a positive incentive effect. But also, the individual would experience a 'quasi lump-sum' increase in her income of the magnitude $\Delta\tau \cdot (z' - \dot{z})$. If the income effect is operative, this is a disincentive to work.

Since the optimal tax rate is calculated as a function of the behavioural effect (the greater the behavioural effect, the lower the optimal tax rate), the presence of the income effect, by dampening the behavioural effect, increases the optimal top marginal tax rate.

Most studies that estimate the taxpayer response to tax changes have not found a significant income effect. The study by Bakos et al. (2008), is an exception both regarding its original results and as re-estimated (the re-estimated BBB results are surveyed by Benczúr et al. [2013]). Kiss and Mosberger (2011) also found a sizeable income effect which, however, was not precisely estimated. In a robustness analysis of microsimulation results, Benczúr et al. (2012, Table 8) report that the point estimate of the income effect found by Kiss and Mosberger wipes out most of the estimated behavioural effect of the tax changes between 2010 and 2013. If an income effect of such magnitude were established, this would imply much higher optimal tax rates than the benchmark results in this paper.

These considerations show that while the income effect is potentially important, its magnitude cannot be precisely assessed based on the available information. For this reason, among others, the measurable effects of the large recent tax reforms will be of great importance. If the behavioural effect appears to be large (or small) based on incoming data in 2014, this may give some guidance as to which end of the results presented in this paper may be more relevant.

Actuarial considerations

The baseline results have been derived under the assumption that SSC have the same effects as taxes. (This is the approach taken by Brewer et al. [2010] and Diamond and Saez [2011].) This assumption is easy to justify if the link between present contributions and future benefits is weak, but not if they are strong. Arguably, the relationship is weak in the case of healthcare benefits but stronger in the case of pension benefits. If individuals view part of their SSC as savings which they receive back in old age (at least in expectation), then these contributions may not have the same distortionary effect on their working decisions than taxes have. In this case actual marginal tax rates should be calculated with the exclusion of one part of contributions, and the optimal PIT rates will be higher.

In the case of Hungary, it is not clear that the abolition of the pension contribution ceiling in 2013 has the same effect as increasing the top marginal PIT rate would have. High-income individuals previously above the pension contribution ceiling now have to pay employee pension contributions, but their promised pension benefits also increase. In this case, the relationship between additional

contributions and additional benefits may be quite salient and hard to undo by future policy-makers.

For this reason, an example has been calculated in which the marginal link between pension contributions and benefits is perceived by taxpayers to be 50 percent. This is the middle way between two extreme theoretical cases. In the first theoretical extreme, a fixed-sum basic pension is financed from income-dependent tax revenue: here the marginal link between 'contributions' and benefits is zero. In the other theoretical extreme, pension contributions are paid to actual private accounts which form the basis of the future pension benefit annuity. In this case, the marginal link between contributions and benefits is 100 percent. The reality of the Hungarian PAYGO pension system is arguably somewhere in-between.

Figure 3 shows the actual top marginal tax rates in recent years under the assumption that taxpayers can consider half of their pension contributions as 'their own money' (while another part of their future pension benefits are assumed to be lump-sum).

Figure 3. Actual effective top tax rates in Hungary under the assumption that the marginal link between pension contributions and future benefits is 50 percent
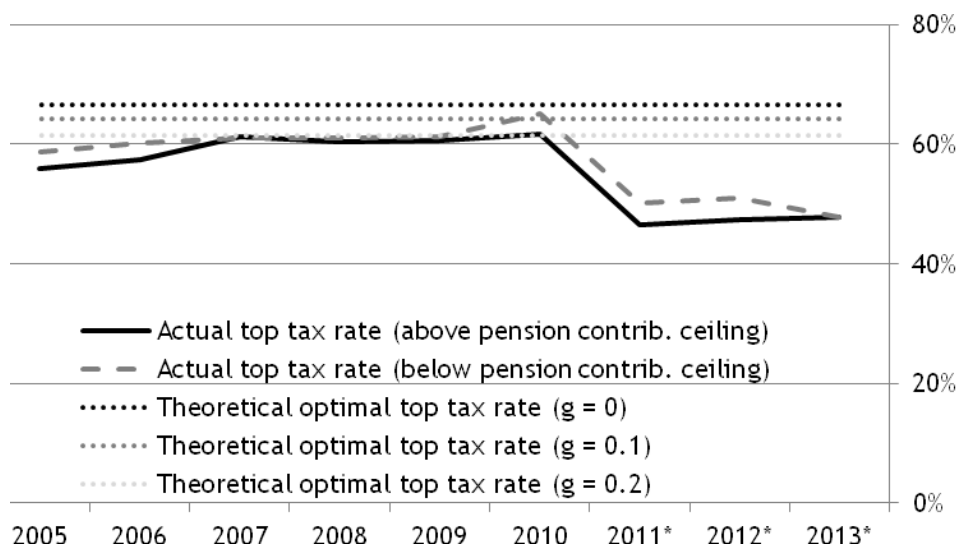
Note: see Notes to Figure 2.

While the example is necessarily somewhat arbitrary, Figure 3 shows that actual top marginal tax rates depend greatly about the perceived status of SSC. In this example, if only half of pension contributions are perceived as a tax, the main conclusions of this analysis are changed significantly. First, in this case top tax rates in Hungary have never reached the revenue-maximising tax rate. Second, the top marginal tax rate in 2013 is below 50 percent, even further from theoretical benchmarks than under the baseline assumptions.

Tax evasion, tax avoidance, long-term elasticities and international tax competition

The logic of the results presented in this paper can accommodate considerations related to tax evasion, tax avoidance, long-term effects of taxes on labour supply, or international tax competition. Tax evasion. If high earners hide some of their income as a response to a tax increase, this will appear to the state as response at the intensive margin and will be taken into account just like a real economic response. Tax avoidance. If taxpayers 'relabel' some of their income after a tax change, then this means that the loss in one tax revenue will be partly made up by another tax revenue. This means that the total tax elasticity might be lower than the

elasticity measured by estimations that focus on a single source of revenue. In Hungary there is no strong evidence for such cross-tax effects although they may exist. Long-term elasticities. Possibly, some types of adjustment by taxpayers take time. Lower top marginal tax rates may imply different career decisions by some individuals or even different educational decisions. There is no strong evidence about these effects either way, but they could lead one to assume that long-term elasticities are greater than short-term ones researchers have estimated so far. International tax competition. If low taxes attract the business of some high-added-value activity (e.g., finance), this will appear to the government as a response of the tax base to changes in the tax rate. In theoretical terms this should be considered as adjustment at the extensive margin. Strong international competition effects would imply an increase in the relevant elasticity to be taken into account in the analysis. For Hungary (or even internationally) evidence for this kind of tax competition is scarce (but see, e.g., recent examples in somewhat special contexts by Kleven et al. [2010 and 2013]).

Conclusion

The theory of optimal income taxation was long considered an esoteric field with little to say about actual tax policy. This view changed with the developments of the last 15 years. Although it is not to be expected that the theory enables us to derive 'the optimal tax system', we can arrive at some robust qualitative results and in some cases even quantitative guidance for tax policy.

This paper applies one recent result of the optimal theory of income taxation to the Hungarian case: it suggests that the optimal top marginal tax rate is likely to be higher than the current actual rate.

The theory requires surprisingly few assumptions, and only two estimable parameters, to derive a value for the revenue-maximising top marginal tax rate. This tax rate is optimal if the value society attaches to an additional dollar kept by a top earner is negligible. If this 'social marginal value of income' for high earners (parameter g) is not negligible, it can be considered as a parameter while the optimal top marginal tax rates can be calculated as a function of this parameter.

According to simulations in this paper, the revenue-maximising top marginal tax rate is about 67 percent in Hungary (including SSC and consumption taxes), consistent with a PIT of about 30 percent (or 40 percent if pension contributions were capped). The actual top marginal PIT rate is 16 percent in 2013. The actual rate is lower than the optimal top marginal tax rate under plausible values of parameter g.

A number of factors are discussed that could qualify or modify the baseline results. Notably, the optimal top PIT rate is higher than in the baseline results if the income effect is significant or if pension contributions have a strong link to future pension benefits. In contrast, the optimal top PIT rate is lower than in the baseline results if the long-run elasticity of taxable income is greater than the short-run elasticity about which evidence is available, or if international competition for talent is a binding constraint for the social planner.

References

Atkinson. Anthony B. (1990) 'Public Economics and the Economic Public', *European Economic Review* 34 (7): 225-248.

Bach, Stefan; Corneo, Giacomo and Viktor Steiner (2012) 'Optimal top marginal tax rates under income splitting for couples', *European Economic Review* 56(6): 1055-1069.

Bakos, Péter; Benczúr, Péter and Dóra Benedek (2008) 'The elasticity of taxable income: Estimates and flat tax predictions using the Hungarian tax changes in 2005', MNB Working Paper 2008 No.7.

Benczúr, Péter; Kátay, Gábor; Kiss, Áron; Reizer, Balázs and Szoboszlai Mihály (2011) 'Analysis of changes in the tax and transfer system with a behavioral microsimulation model', MNB Bulletin, October 2011.

Benczúr, Péter; Kátay, Gábor and Áron Kiss (2012) 'Assessing changes of the Hungarian tax and transfer system: A general equilibrium microsimulation approach', MNB Working Paper 2012 No.7.

Benczúr, Péter; Kiss, Áron and Pálma Mosberger (2013) 'The elasticity of taxable income', in K. Fazekas, P. Benczúr, and Á. Telegdy (eds) *The Hungarian Labour Market 2013*, Budapest: Hungarian Academy of Science.

Brewer, Mike; Saez, Emmanuel and Andrew Shephard (2010) 'Means-testing and Tax Rates on Earnings', in J. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles and J. Poterba (eds) *Dimensions of Tax Design: the Mirrlees Review,* Oxford University Press, pp. 90-173.

Diamond, Peter (1998) 'Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates', *American Economic Review* 61(1): 8-27.

Diamond, Peter and Emmanuel Saez (2011) 'The Case for a Progressive Tax: From Basic Research to Policy Recommendations', *Journal of Economic Perspectives* 25(4): 165-190.

Feldstein, Martin (1995) 'The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act', *Journal of Political Economy* 103(3): 551-572.

Giertz, Seth H. (2004) 'Recent Literature on Taxable-Income Elasticities', Technical Paper Series, Congressional Budget Office, Washington D.C.

Goolsbee, Austan (2000) 'What happens when you tax the rich? Evidence from executive compensation', *Journal of Political Economy* 208(2): 352-378.

Gruber, Jonathan and Emmanuel Saez (2002) 'The Elasticity of Taxable Income: Evidence and Implications', *Journal of Public Economics* 84: 1-32.

Kiss, Áron (2013) 'Developments in the theory of optimal income taxation with applications to the Hungarian tax system', Discussion Papers of the Institute of Economics, Centre for Economic and Regional Studies, Hungarian Academy of Sciences, Budapest.

Kiss, Áron and Pálma Mosberger (2011) 'The elasticity of taxable income of high earners: Evidence from Hungary', MNB Working Paper 2011 No. 11.

Kleven, Henrik J.; Landais, Camille and Emanuel Saez (2010) 'Taxation and International Migration of Superstars: Evidence from the European Football Market', NBER Working Paper No. 16545.

Kleven, Henrik J.; Landais, Camille; Saez, Emanuel and Esben A. Schultz (2013) 'Taxation and International Migration of Top Earners: Evidence from the Foreigner Tax Scheme in Denmark', NBER Working Paper No. 18885.

Mankiw, N. Gregory; Weinzierl, Matthew C. and Danny Yagan (2009) 'Optimal taxation in theory and practice', *Journal of Economic Perspectives* 23(4): 147-74.

Mirrlees, James A. (1971) 'An exploration in the theory of optimal income taxation', *Review of Economic Studies* 38(2): 175-208.

Riihelä, Marja; Sullström, Risto and Matti Tuomala (2013) 'Top incomes and top tax rates: Implications for optimal taxation of top incomes in Finland', Tampere Economic Working Paper Series No. 88.

Saez, Emmanuel (2001) 'Using elasticities to derive optimal income tax rates', *Review of Economic Studies* 68(1): 205-229.

Saez, Emmanuel; Slemrod, Joel B. and Seth H. Giertz (2012) 'The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review', *Journal of Economic Literature* 50(1): 3-50.

Appendix

Table A1. Actual top tax rates in Hungary, 2005-2013

	*2013**	*2012**	*2011**	*2010*	*2008*	*2005*
Top PIT rate	16.0%	20.3%	20.3%	40.6%	40.0%	36.0%
Employee pension contributions	10.0%	10.0%	10.0%	9.5%	9.5%	8.5%
Employee health contributions	8.5%	8.5%	7.5%	7.5%	6.0%	4.0%
Employer contributions	27.0%	27.0%	27.0%	27.0%	32.0%	32.0%
Effective consumption tax rate[a]	15.9%	15.9%	15.9%	15.9%	15.1%	14.8%
Actual top tax rate (above pension contrib. ceiling)[b]	56.6%	52.9%	52.2%	65.7%	65.3%	61.3%
Actual top tax rate (below pension contrib. ceiling)	56.6%	59.5%	58.8%	72.0%	68.8%	66.8%
Theoretical optimal top tax rate (g = 0)	66.7%	66.7%	66.7%	66.7%	66.7%	66.7%

[a] The effective tax rate on consumption was calculated as tax revenue from consumption taxes (VAT, excise taxes and other, minor taxes) divided by total household consumption. For the years 2011 to 2013 the empirical effective consumption tax was not known at the time of writing. Thus for these years the value for 2010 was used.

[b] Until 2013, high-income individuals did not pay employee pension contributions on income exceeding the pension contribution ceiling (employer contributions have not been capped). Thus here only the health-care contributions were accounted on the employee side until 2013. The employee pension contribution ceiling was abolished effective 2013.

Figure A1. Marginal effective tax rates (METR) at high incomes in Hungary, 2005—2013
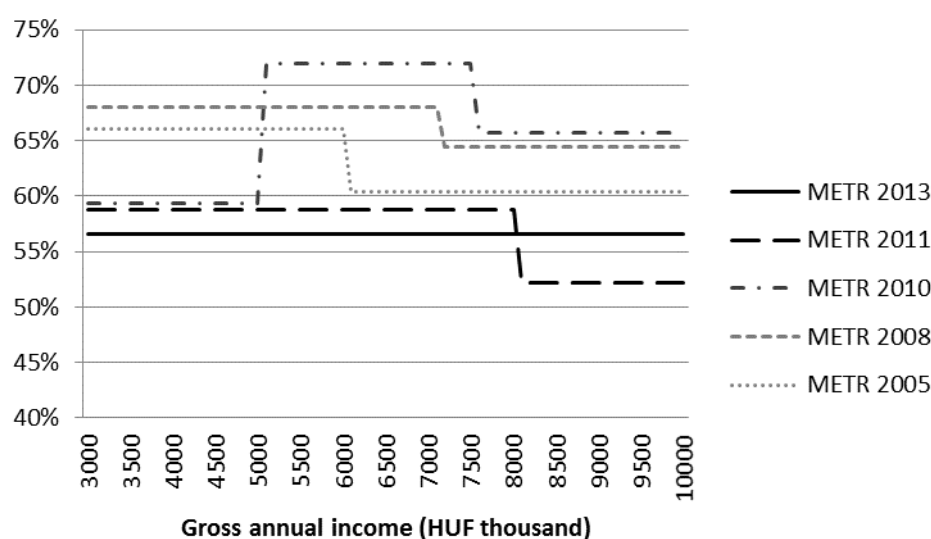
Note: Marginal Effective Tax Rates in this figure are consistent with the calculations of the rest of the paper: They take into account social security contributions as well as the effective consumption tax rate.