MAGYAR TUDOMÁNYOS AKADÉMIA
Közgazdaság- és Regionális Tudományi Kutatóközpont

MTA KRTK KTI

Centre for Economic and Regional Studies
HUNGARIAN ACADEMY OF SCIENCES

MŰHELYTANULMÁNYOK DISCUSSION PAPERS

MT-DP – 2013/17

Developments in the theory of optimal income taxation with applications to the Hungarian tax system

ÁRON KISS

INSTITUTE OF ECONOMICS, CENTRE FOR ECONOMIC AND REGIONAL STUDIES,
HUNGARIAN ACADEMY OF SCIENCES BUDAPEST, 2013

Discussion papers
MT-DP – 2013/17

Institute of Economics, Centre for Economic and Regional Studies,
Hungarian Academy of Sciences

Developments in the theory of optimal income taxation
with applications to the Hungarian tax system

Author:

Áron Kiss
analyst
Central Bank of Hungary
Email: aron.kiss@gmail.com

May 2013

ISBN 978-615-5243-73-8
ISSN 1785 377X

Developments in the theory of optimal income taxation with applications to the Hungarian tax system

Áron Kiss

Abstract

This paper reviews recent developments in the theory of optimal income taxation and draws two broad conclusions with respect to the Hungarian personal income tax system. The first conclusion is that the optimal top marginal tax rate is likely to be higher, perhaps substantially, than the actual rate. The second conclusion is that the optimal tax burden of earnings near the minimum wage is likely to be lower than the actual tax burden. It is discussed how these results depend on the parameters describing labor-supply behavior, the income distribution, and the redistributive preferences of society.

Keywords: optimal income taxation, top income tax rate, Hungary, emerging markets

JEL classification: H21, H24

Acknowledgements:

Views and opinions expressed in this paper are those of the author and do not necessarily represent those of the Central Bank of Hungary (MNB), or any other institution. Calculations in the paper reflect information available as of April 2012. The author would like to thank participants of the 'Simposium on Taxes, Transfers and the Labour Market' organized by Institute of Economics, Centre for Economic and Regional Studies, Hungarian Academy of Sciences (IE-CERSHAS) and MNB in June 2012 and the IE-CERSHAS conference on 'Wages, taxes and transfers', held in Szirák, November 2012. Any remaining error or omission is the responsibility of the author.

Fejlemények az optimális jövedelemadózás elméletében és alkalmazásuk a magyar adórendszerre

Kiss Áron

Összefoglaló

A tanulmány áttekinti az optimális jövedelemadózás elméletének újabb eredményeit, és két következtetést von le a magyar jövedelemadózás rendszerére vonatkozóan. Az első következtetés szerint az optimális felső adókulcs valószínűleg magasabb – talán jelentősen magasabb –, mint a jelenlegi tényleges ráta. A második következtetés szerint a minimálbér körüli jövedelmeket érintő optimális adóteher valószínűleg alacsonyabb, mint a tényleges adóteher. A diszkusszió körüljárja, hogyan függnek ezek a következtetések az adózók munkakínálati viselkedését és a jövedelemeloszlást leíró becsülhető paraméterektől, és hogyan függnek a társadalom újraelosztási preferenciáitól.

Tárgyszavak: optimális jövedelemadózás, felső adókulcs, Magyarország, feltörekvő piacok

JEL kódok: H21, H24

1. 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 completed by 2013 when the single PIT rate is set to be 16%. The reforms benefited different groups in different years but altogether they brought a radical cut in the marginal (and average) tax rate 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 two of its main recommendations to the Hungarian tax system.

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 society1) 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, the exact 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 previous 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 target to be achieved by the tax, 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 a given society. An applied optimal taxation

[1] In the paper 'society' and 'government' is used interchangeably. The analysis abstracts from the 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 approach is important in the real world but it is not the focus of this paper.

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. But in some cases it is even possible to derive quantitative guidance: for example 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, Saez, and Shephard (2010) was prepared in the framework of the so-called Mirrlees Review, a detailed review of the U.K. tax system by an international group of researchers coordinated by the Institute for Fiscal Studies. The second such study, by Diamond and Saez (2011), summarizes three broad principles that can be derived from the theory and discusses it in the context of current U.S. tax policy.

Two main results of this literature are applied for the Hungarian context. The first result suggests that high marginal tax rates for top earners can be optimal. The second result is that low (or perhaps negative) effective tax rates at low earnings levels can be optimal.

The first 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 behavioral 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% of earners in their definition) is close to zero. This can be true on a utilitarian basis (where the welfare of every citizen is equally important to the government) if an extra dollar of income increases the welfare of a high earner by much less than that of a low earner (or, more precisely, if the marginal utility of consumption declines to zero). In this case, the optimal top marginal tax rate is equal to the top rate that maximizes 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 maximizes 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-maximizing top marginal tax rate is higher than the current top marginal tax rate: the revenue-maximizing top marginal PIT rate is estimated to be about 40% as applying to income above the pension contribution ceiling (or about 30% if pension contributions were not 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-maximizing rate if plausible non-zero values are chosen for the parameter expressing the social marginal value of a dollar kept by a top earner. Macro-economic and fiscal effects of hypothetical tax reforms that are consistent with these results are simulated.

The second main result is based on the insight, first derived by Saez (2002), that reducing the tax burden of low income earners has an unequivocally positive effect on social welfare as long as the tax burden is still positive and labor supply adjustment of low earners occurs mainly at the extensive margin (that is, entering and exiting the labor market). The reason for this result is that reducing the tax burden of low income earners makes it more attractive to take up a low-paying job (as opposed to being inactive), which reduces the amount of transfers the government pays to the inactive. Thus this policy must raise social welfare as long as society has any redistributive preference.

This result is weakened if adjustment at the intensive margin is important as well. This may be the case if wage under-reporting behaviour (thought to be wide-spread in Hungary) is sensitive to the tax rates. Still, the extensive margin is likely to be dominant for two reasons: First, this is the conclusion suggested by the research literature on labor supply and, second, Hungary has a very low employment rate in international comparison.

The theory of optimal income taxation has rarely been cited in tax policy debates in Hungary. Studies by Benczúr (2007), Bakos, Bíró, Elek and Scharle (2008) and Scharle et al. (2010) analyzed the tax system from an efficiency point of view, explicitly taking into account behavioral responses to taxation, but they have not drawn on the new optimal-income-tax literature following Saez (2001) to draw conclusions for Hungary. Looking at a different aspect of optimal taxation, Valentinyi (2001) surveyed the literature on the optimal taxation of capital income. Since this question is analyzed by a somewhat separate literature it is outside the focus of this paper. (Optimal income tax theory has mostly focused on the taxation of wage incomes, even though Diamond and Saez (2011), for example, do include a discussion of capital taxation.) The present paper is, to the author's knowledge, the first one that applies the results of the new literature on optimal income taxation to the Hungarian tax system.[2]

The rest of the paper is organized as follows. The next section derives the general formula for the optimal top marginal tax rate and evaluates it for the Hungarian tax system. Section 3 derives the result that the tax rates should be low at low income levels and discusses it in the Hungarian context. Concluding remarks are offered in Section 4.

2. THE OPTIMAL TOP MARGINAL TAX RATE

2.1 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 analyze whether actual top marginal tax rates in the U.S. (Diamond and Saez, 2011) and the U.K. (Brewer et al., 2010) are in line 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% of the income distribution. In this paper it is be explored how conclusions depend on whether top incomes are defined as the top 1% (starting at income level HUF 10—11 million in Hungary); or the top 5% (starting at about HUF 5 million); or the top 10% (starting at about HUF 4 million).

Consider an economy where individuals have different earning capacities. Individual i earns 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 labor 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 $\tau(z) = T'(z)$.

[2] In her diploma thesis, Daragó (2011) conducted similar calculations under the supervision of the author. Some results of the present paper have been presented in simplified form by Benczúr, Kiss and Mosberger (2013).

The government sets tax rates and transfers in a way as to maximize 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 tradeoff 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 tradeoff between equity and efficiency is at the center of the theory of optimal income taxation.

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

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

Note that this effect is defined as the change in tax revenue before any behavioral 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:

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

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

The last is the behavioral 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 behavioral responses to tax changes estimate a parameter $e = \frac{dz}{z} \frac{1-\tau}{d(1-\tau)}$. 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 one-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

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

At the optimal marginal tax rate τ^* the sum of these effects must be equal 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 be optimal.

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

$$\tau^* = \frac{1-g}{1-g+e \cdot a}$$

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 characterizes 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 behavioral effect. In this case the optimal top marginal rate is equal to the revenue-maximizing rate, with the formula simplifying to:

$$\tau^* = \frac{1}{1+e \cdot a}$$

Brewer et al. (2010) and Diamond and Saez (2011) argue that $g = 0$ is plausible for top earners (the top 1% 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 $\lim_{z\to\infty} g(z) = 0$. 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-maximizing tax rate.

But parameter g does not have to converge to zero for the revenue-maximizing 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.

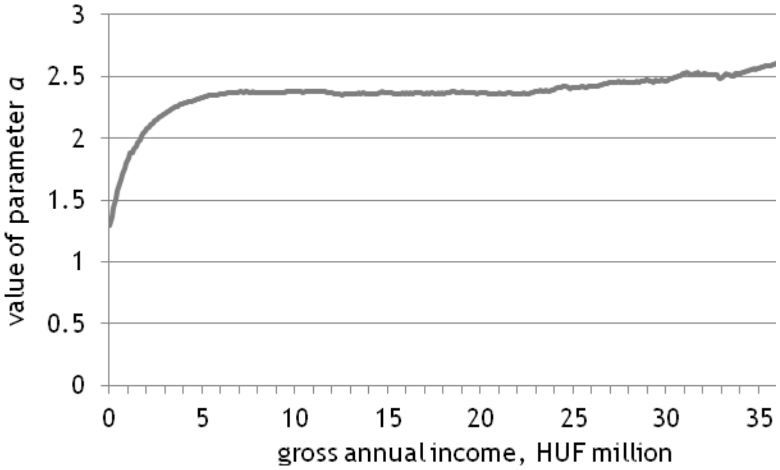
2.2 THE REVENUE-MAXIMIZING TOP MARGINAL TAX RATE IN HUNGARY

The previous subsection established that the revenue-maximizing 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-maximizing top marginal tax rate. As seen above, the revenue-maximizing 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 $a = z^m/(z^m - \bar{z})$, that is, for any income limit $\bar{z}$, 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 personal income tax 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 400) and HUF 40 million (about EUR 160,000). 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 7,500) 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 42,000).

Figure 1.

The value of parameter a as a function of the income limit chosen



Note: The 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. It is around 2.35 for incomes between HUF 6—23 million and is very close to 2.5 for income levels above that.

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 $\bar{z}$ is a constant multiple of $\bar{z}$. In the Hungarian tax data the average income of individuals earning more than $\bar{z}$ is about $1.7 \cdot \bar{z}$.

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 net-of-tax rate. This parameter first came into the focus of economic research with the estimations of Feldstein (1995). Before his work, labor 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 behavioral 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 U.S. (see also recent surveys of the literature by Giertz (2004) and Saez, Slemrod and Giertz (2012)).

There are two studies that estimated the taxable-income elasticity for Hungary. In the first such study, Bakos, Benczúr, and Benedek (2008), henceforth BBB, used the elimination of the middle tax bracket in 2005 as the policy experiment to identify the elasticity. The elimination of the middle tax bracket increased the tax rate of some individuals while reduced the tax rate of others. If the income growth is different for otherwise similar groups who are affected differently by the tax changes, this can be attributed to their behavioral response. 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. Nevertheless they are able to estimate the elasticity for different income groups separately and find an elasticity of $e = 0.34$ for individuals earning more than HUF 2 million (about EUR 8,000). (BBB also find, in many specifications, a significant income effect which dampens the reaction of reported incomes to tax changes. How a significant income effect would affect the optimal tax rates 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—8 million or about EUR 20—32 thousand). The main estimates for the elasticity fall between 0.15—0.2, somewhat lower than the comparable estimate of BBB. 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.

This central estimate is in line with most estimates for countries outside the U.S.[3] Some observers attributed the higher estimated U.S. elasticities at least partly to tax optimization strategies (timing and form of compensation, among others) that are available to executives in the U.S. (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-maximizing top marginal tax rate. Based on these parameter values one can evaluate the formula for the optimal 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-maximizing top marginal tax rate of high earners is 67% at the central parameter estimates ($a = 2.5$ and $e = 0.2$). Of course, this tax rate cannot be directly compared to the rates of the personal income tax, since taxpayers also pay social security contributions and consumption taxes. Comparable actual top marginal tax rates for the last few years in Hungary are calculated in the next subsection.

Table 1

The optimal top marginal tax rate as a function of *a* and *e*, with *g* = 0

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 calculation

Table 1 also shows that, within the plausible range of parameter estimates, the revenue-maximizing top marginal tax rate depends more sensitively from the value of the taxable-income elasticity than from the Pareto parameter. 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. At the same time, 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.

[3] For references to elasticity estimates in other countries, see Kiss and Mosberger (2011).

2.3 THE OPTIMAL TOP MARGINAL TAX RATE AS A FUNCTION OF PARAMETER g.

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 favor 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%, the top 5%, and the top 10% 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%	10.60	18.30	0.11
top 5%	5.27	9.26	0.21
top 10%	3.76	6.83	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% of tax filers if her income exceeded HUF 10.6 million (about EUR 42,000). The average income of the top 1% was about HUF 18.3 million (about EUR 73,200). In the same year the average gross annual income was about HUF 1.9 million (about EUR 7,500). 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%.

Similarly, the lower limit of the top 5% was, in 2008, at about HUF 5.3 million (about EUR 21,000) while the calculated g is about 0.2. The lower limit of the top 10% 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%

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% instead of 67%. 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% in Hungary. Thus this result means that it matters little whether, for the top 1% 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%, about 0.2 for the top 5%, and about 0.3 for the top 10%. In this example the optimal top marginal tax rate is approximately 64% if it affects only the top 1% (applying to income above HUF 10.6 million at 2008 prices) but it is only 58% if it affects the top 10% (applying to income above HUF 3.8 million at 2008 prices).

2.4 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 personal income tax (PIT) rate. Clearly, social security contributions have the same disincentive effect as the PIT, whether they are paid by the employee or the employer. 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 labor 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:

$$\tau_{top} = 1 - (1 - \tau_{cons}) \frac{1 - \tau_{PIT} - \tau_{EE}}{1 + \tau_{ER}}.$$

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.

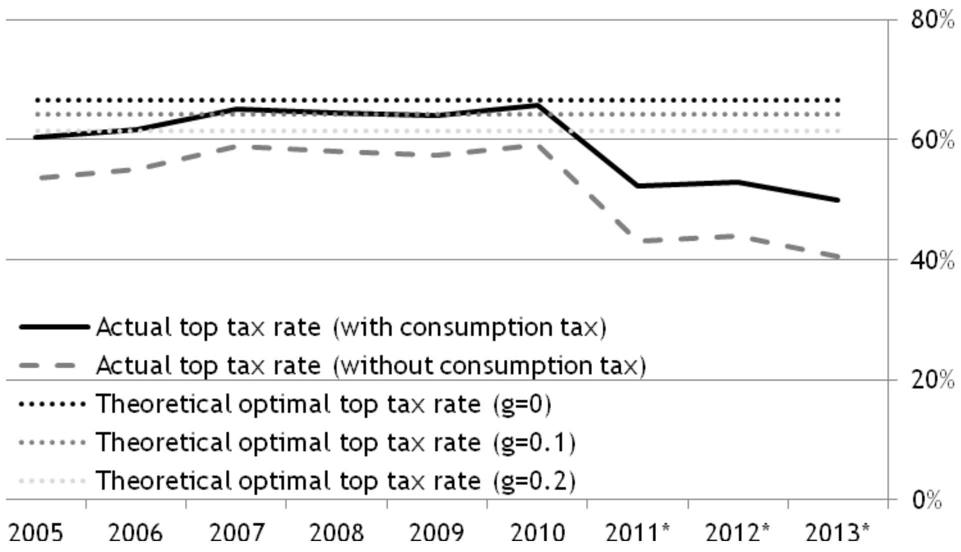
Two issues regarding the tax rates merit further discussion. The first issue is related to the employee contributions. The actual top marginal tax rate is calculated here for a hypothetical person with income above the pension contribution ceiling. Individuals above the pension contribution ceiling do not have to pay the employee pension contribution, which reduces their marginal tax rate by about 10%. (The pension contribution rate varied between 8.5—10% in recent years.) 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 face a lower marginal tax rate than individuals at the 95th percentile.

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 taxes on tobacco, alcohol products and gasoline, and further smaller items, besides the 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—16% in the period 2005—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 (with consumption taxes and without) that result from these calculations are shown in Figure 2, along with the optimal top marginal tax rate at different values of g.

Figure 2.

Actual effective top tax rates in Hungary, 2005-2013



Note: Calculations of the author based on statutory tax rates, budget reports and the National Accounts. Statutory tax rates for 2013 reflect current legislation as of February 2012. The effective consumption tax rate for the years 2011—2013 is, for lack of data, taken to be equal to its 2010 value.

It is apparent from Figure 2 that the flat PIT introduced in 2011 (the rate of which is set to decrease further in 2013 from about 20% to 16%) has cut the top marginal tax rate significantly: from about 65% to slightly above 50%, taking into account consumption taxes.[4]

2.5 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 rates applying before 2010 appear to be around the revenue-maximizing rate, while the 2011 tax cut resulted in marginal tax rates that are clearly below that rate. Second, those individuals in the top 5—10% whose income was below the pension contribution ceiling faced marginal tax rates in the period 2005—2010 that were higher than the revenue-maximizing top income tax rate.

The first conclusion means that the tax revenue foregone by cutting the marginal tax rate of the very top 1-2% of earners cannot be justified by the behavioral effect: the estimated

[4] 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—2 percentage points to the top marginal rate, however. Thus they would not affect the results qualitatively.

income elasticity is not large enough to motivate a PIT rate of 16—20%. The calculated optimal top marginal tax rate is consistent with a top PIT rate between 30% (about optimal under $g = 0.3$) and 40% (about optimal under $g = 0$). Equivalently, if pension contributions were not capped at top incomes, the calculated optimal top marginal tax rate would be consistent with a top PIT rate between 20% (about optimal under $g = 0.3$) and 30% (about optimal under $g = 0$).

Looking at the results from a different perspective, one can calculate, going back to Table 1, under which parameter values the current top marginal rate of about 50% would be optimal. With parameter values of $a = 2.5$ and $g = 0$ (that is, looking at row 2 in Table 1) the taxable-income elasticity (parameter e) should be about 0.4, twice its estimated value, for the top marginal rate of 50% to be optimal.

The second conclusion (that individuals with income just below the pension contribution ceiling were taxed at a rate higher than the optimal top marginal tax rate before 2010) 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 for selected years between 2005—2012 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 available evidence, in the income range just below the pension contribution ceiling.

In the following, some 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 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 next section contains some considerations about the lower part of the tax schedule.

The simulations were conducted in the microsimulation model described by Benczúr et al. (2011, 2012). The model takes into account the behavioral responses to taxation by incorporating the taxable-income elasticities used in the rest of this paper. The microsimulation model is embedded in a small neoclassical general-equilibrium macroeconomic model. This means that if individuals change their labor 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 labor-supply response of 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 summarizes the results of the simulations. Four simple tax policy measures (and a composite one) are considered. For each measure, a 'static fiscal effect' is calculated: this says by how much government revenue would increase from the PIT and the social security contributions if the given measure were adopted, absent behavioral change. Additionally, the 'dynamic fiscal effect' and the 'dynamic GDP effect' are calculated. The dynamic fiscal effect takes into account both the behavioral response to the tax change by the taxpayers and the long-run macroeconomic adjustments. Since the tax measures considered are small in comparison to the economy, most of the dynamic effect comes from individual behavior rather than the general-equilibrium effects. Thus most of the simulated dynamic effects can be thought to take place in the first couple of years after a tax reform. Finally, the dynamic GDP effect says by how much the long-run level of GDP is affected by the tax policy change. If for example a tax increase decreases the long-run level of GDP by 0.5%, this could mean that it slows GDP growth by 0.1% for five years, compared to a baseline scenario with no change in the tax system.

The four simple measures analyzed are: (1) the elimination of the pension contribution ceiling; (2) the introduction of a PIT rate of 30% above HUF 10 million (approximately 1% of taxpayers); (3) the introduction of a PIT rate of 25% above HUF 5 million (approximately 5% of taxpayers); and (4) keeping the upper PIT rate of 20% rate applying to income above HUF 2.4 million in 2012, intended to be eliminated by 2013. All tax policy measures are compared to the baseline of a flat PIT of 16%, which is current legislation for 2013 (as of April 2012).

Measure (1), or an equivalent measure in the PIT increasing the marginal tax rate by 10% on income exceeding HUF 8 million, would prevent the marginal tax rate from decreasing in the top income range. Measure (2), if complemented by measure (1), would approximately implement the revenue-maximizing tax rate for the top 1% of earners. Measure (3), if complemented by measure (1), would approximately implement the optimal tax rate for the top 5% of earners under $g = 0.2$. Finally, measure (4) is just the continuation of 2012 policy.

Table 4

Hypothetical tax policy scenarios

		Static	Dynamic	
		fiscal effect (HUF billion)	fiscal effect (HUF billion)	GDP effect (diff. in level)
(1)	Elimination of the pension contribution ceiling	47	29	-0.3%
(2)	PIT rate of 30% above HUF 10 million	48	29	-0.3%
(3)	PIT rate of 25% above HUF 5 million	82	49	-0.5%
(4)	PIT rate of 20% above HUF 2.4 million	101	69	-0.5%
	A measure composed of (1) + (3) + (4)	191	114	-1.1%

Note: The 2013 tax system (as of 2012 legislation) was used as benchmark. Fiscal estimates include the PIT and social security contributions. The dynamic simulations are based on the taxable-income-elasticity estimates by Kiss and Mosberger (2011).

Table 4 shows that the elimination of the pension contribution ceiling (or an equivalent measure in the PIT) would increase government revenue by about HUF 50 billion under no behavioral change, and by about HUF 30 billion after behavioral effects and macroeconomic feedback effects are taken into account. The measure would decrease the long-run level of GDP by about 0.3%. The effects of increasing the PIT rate from 16% to 30% on income exceeding HUF 10 million are very similar. The effects of a lower top rate levied on a broader base (25% on income exceeding HUF 5 million) are larger by about one-half, while the effects of keeping the upper tax bracket of 2012 are even larger.

A measure composed of measures (1), (3) and (4) is analyzed in the last row of Table 4. This measure implements a moderate ($g > 0$) top marginal rate for about the top 5% of the income distribution while it keeps the top tax bracket of 2012. This composite measure would improve government balance by about HUF 190 billion (static) or HUF 110 billion (dynamic) and decrease the long-run level of GDP by just over 1%. The fiscal effect is equivalent to 0.4—0.7% of GDP, where the lower end of the interval is the dynamic fiscal estimate and the upper end is the static estimate. This total fiscal effect is much smaller in absolute value than that of the 2011 tax reform but it allows a non-negligible adjustment of the government balance.

The simulated macroeconomic effects exemplify the trade-off inherent in optimal tax theory between equity and efficiency. The disincentive effect of higher tax rates means that the economy will grow on a somewhat lower path (but not, after the initial years, at a lower pace) than under the baseline.

As emphasized above, these are not the only conceivable income-tax measures that are consistent with the theoretical results derived above. They should rather serve as an orientation as to the likely range of tax rates, fiscal and macroeconomic effects similar measures would imply.

2.6 DISCUSSION

The results of this section show that high marginal tax rates for high earners can be optimal. This subsection looks at some considerations that might affect this result.

How do these results relate to the zero-top-rate result? In many presentations of optimal income tax theory it is stated that the optimal top marginal tax rate is zero. To see where the zero-top-rate result comes from, consider the usual proof by contradiction (this result was described by Sadka (1976) and Seade (1977)).

Consider an economy with *N* individuals with different earning capacities. Suppose we propose an optimal tax system in which the marginal tax rate of the very top earner is positive. It can be shown that this tax system can be changed in a Pareto-improving way, thus this tax system could not be optimal in the first place. Take the income level z' that the top earner would earn under the proposed tax system. Let us leave the tax schedule unchanged up to z' and set the marginal tax rate equal to zero above z'. What will happen? The very top earner's marginal incentives improve, thus she will choose to exert more effort and reach a higher level of earnings without paying more (or less) taxes than before. Nothing changes for other taxpayers, and total government tax revenue is unchanged, too. But the highest earner is better off: since the old optimum is still attainable, a different choice means that she is better off under the new tax system. Thus the change of the tax schedule is Pareto improving. Thus the originally proposed tax system could not be optimal. *QED*.

What is the relationship between the zero-top-rate result and the high optimal top tax rates derived in this study? The answer is that the zero-top-rate result has little practical value because (1) even in theory it applies only to the single highest income (i.e., practically to one individual) and (2) its implementation would require that the government knows how much the very top earner will be able to earn next year. When planning next year's budget no government knows exactly what the highest income will be. Even though every empirical income distribution is finite, it is reasonable to think of the highest incomes as random draws from a Pareto distribution. For this case it is easy to show that the formula derived in this section still gives the optimal top marginal tax rate (see Saez (2001) and Diamond and Saez (2011, online appendix) for details).

Are there other views on the top marginal tax rate? Two years before Diamond and Saez (2011), Mankiw, Weinzierl and Yagan (2009) published an article in the same journal about the practical consequences of optimal tax theory in which they came to a different conclusion about the top marginal tax rate. The main point of Mankiw et al. is that if the distribution of earning potential (or ability) does not follow a Pareto distribution but rather a lognormal distribution, then optimal marginal tax rates at the top will be decreasing with income. (In the simulations of Mankiw et al. (2009, Figure 3), the optimal marginal tax rate at \$150,000 is between 60—70% with a Pareto distribution but only about 40% with a lognormal distribution.)

The lognormal distribution is 'thinner' at the top than the Pareto distribution which means that parameter $a = z^m/(z^m - \bar{z})$ diverges to infinity, albeit possibly at a slow rate. Mankiw et al. give two arguments against using a Pareto distribution. First, they claim that it takes strong assumptions to infer the true ability distribution and, second, they conduct a test on data from the Current Population Survey (CPS) to show that empirically it is not possible to distinguish the Pareto distribution from the lognormal distribution. Diamond and Saez (2011, p. 188) find both arguments 'invalid'. They argue that the CPS is not the appropriate data set to analyze the top of the income distribution since it is top-coded and thin at top incomes; individual tax return data are much better. Also, Diamond and Saez argue that rather than comparing fitted density functions it is much more appropriate to calculate parameter $a(z)$ directly for the top range of incomes. Based on the empirical function $a(z)$ the evidence for the Pareto-distribution seems strong.

In the case of the Hungarian tax return data it can be confirmed that the empirical value of parameter a remains very stable (and below a value of 3) even at the highest income levels where only dozens of individuals are observed in the data base that contains 10% of Hungarian taxpayers. This is a strong empirical argument for the Pareto distribution.

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 behavior of high-income earners. It is possible, however, that their behavior is affected by the so-called income effect, too. The income effect is operative when an individual decreases her work effort after receiving a lump-sum transfer (having a higher income).

If such an effect were indeed operative, then the gain from the behavioral effect would be dampened in response to a tax cut. To see this, suppose that the marginal tax rate is cut by $\Delta\tau$ for those with an income above $\bar{z}$. This tax change would affect an individual with income $z' > \bar{z}$ 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 $\Delta\tau \cdot (z' - \bar{z})$; if the income effect is operative, this increase in income would be a disincentive to work.

Since the optimal tax rate negatively affected by the magnitude of the behavioral effect, the presence of the income 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. This means that if one adopted the taxable-income elasticity estimated by these authors (about 0.3 for middle to higher-income earners), one should also take into account the income effect they estimated. In the example simulations conducted by the author, the elasticities of Bakos et al. with the income effect imply very similar behavioral effects than the Kiss and Mosberger (2011) elasticities. Thus, the implied optimal tax rates should also be very similar based on both sets of elasticities.

Actuarial considerations. It is sometimes argued in public economics that the effect of social security contributions is different from taxes. The reason is that if social insurance is based on actuarial principles, taxpayers may expect to receive the value of their contributions in (pension or health) benefits. This is arguably more the case in the old-age pension system (where in most countries the benefits are in a positive relation with contributions) than in the health-care system (where contributions are in most countries income-dependent while most benefits are not). If individuals view part of their social security contributions as savings which they will receive back in old age, then these contributions will 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 this paper this approach is not chosen for three reasons: (1) such considerations were not taken into account by Brewer et al. (2010) or Diamond and Saez (2011); and (2) in the pay-as-you-go (PAYGO) pension system of Hungary, as in other countries perhaps, individuals may have the expectation that their future pension benefits depend less on present promises than on future realities; and finally (3) the frequent changes in the rules of the Hungarian pension system may obscure the relationship between contributions and benefits more than in other countries with similar PAYGO systems.

If there is a part of the Hungarian social insurance system where actuarial considerations may have affected the expectations of taxpayers, it is the second, private, 'pillar' of the system. The second pillar meant, from its inception in 1997—98, that taxpayers paid part of their pension contributions to an actual (as opposed to notional) private pension account. Individuals could naturally expect to be able to rely on their private account upon retirement. These private accounts were, however, nationalized in 2010. This episode offers two lessons for the present analysis. First, taxpayers in the decade after 2000 may have considered the approximately eight-percentage-point private contribution as their own money. If this is the case, the marginal tax rate in the higher tax bracket, but below the contribution ceiling, may not have necessarily been higher than the revenue-maximizing tax rate in the period of 2005—2010. The second lesson of the Hungarian pension-reform episode is that even reforms that seem to be set in stone can be undone. Since most people spend 30 to 40 years in the labor market before starting to receive benefits from the pension system, they may rightly feel that the relation between their contributions and the future benefits is due to change many times before (and perhaps even during) their retirement. The more change people experience in the present, the less they will expect the present arrangements to last. Supporting the approach taken in this paper as to actuarial considerations, the Hungarian employer-side pension contribution was renamed 'social contribution tax' in 2012.

Tax evasion, tax avoidance and international tax competition. The logic of the results presented in this section can accommodate considerations related to tax evasion, tax avoidance, or international tax competition. These can be easily interpreted in the theoretical framework used here. *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 're-label' 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. *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 except for the case of professional football players (Kleven et al., 2010).

3. THE PARTICIPATION RESPONSE AND THE OPTIMAL ENTRY-LEVEL TAX RATES

3.1 THEORY

While it is possible to derive a simple formula for the optimal top marginal tax rate, one should not expect to be able to derive the optimal tax and transfer system from the theory of optimal income taxation. Some qualitative results can nevertheless be derived for another important aspect of the tax and transfer system: the marginal tax rate at the bottom of the income distribution.

The guidance that theory gives depends, naturally, on what one thinks about the main features of individuals' behavior in response to taxation. Do people react very sensitively to taxes in their work effort, hours of work, or tax evasion? Or do people stop working altogether, or take up a job in the informal economy, if the tax burden is too high? How significant is the informal economy and on what margins does it affect tax revenue and policy?

The original model developed by Mirrlees (1971) considered only responses at the intensive margin, that is, people might work more or less depending on the tax burden, but they do not enter or exit the labor force altogether. In this framework, the optimal tax system provides relatively high transfers to those with no earnings. These transfers are then phased out at a relatively high rate which effectively means that there is a high marginal tax rate on low earnings. This 'traditional welfare program', as Diamond and Saez (2011) label it, is optimal because the high phase-out rate makes a good targeting of benefits possible. While it reduces the work effort of low earners, their tax base is relatively small, thus the negative behavioral effect is not large (as long as the phase-out rate is not too high).

The empirical literature indicates, however, that for low income earners the dominant margin of adjustment is the extensive margin: moving in and out of the work force (see, for an overview, Meghir and Phillips, 2010). This changes dramatically the implied optimal tax rate at the bottom of the income distribution: It becomes optimal to give higher transfers to low-earners than those out of work (Diamond, 1980; Saez 2002). Real-world transfer programs were changed in this spirit since at least the 1990s (like the Earned Income Tax Credit in the U.S. and the Working Tax Credit in the U.K.).

In the following, this insight is derived in a simple theoretical framework following the treatment of Diamond and Saez (2011, online appendix). Their treatment in turn builds on a discrete version of the Mirrlees (1971) model, as developed by Piketty (1997) and Saez (2002).

Consider an economy populated by individuals with different earning abilities. Suppose that individuals with a low earning ability can choose between earning a certain income z_1 or not working at all and having zero market earnings. To motivate this setup in the Hungarian context, it seems plausible that many low-skilled workers (especially in disadvantaged regions) have little chance of obtaining a wage offer higher than the minimum wage (equal in 2012 to HUF 93 thousand monthly or EUR 320 at the current exchange rate).

The government pays a transfer equal to c_0 to those not working. Those who work and earn z_1 receive $c_1 = (1 - \tau_1)z_1 + c_0$, where the effective marginal tax rate τ_1 can be interpreted as the phase-out rate of out-of-work benefits.

The fraction of the population that earns z_1 is h_1. Depending on the 'financial gains to work' (which are given by $c_1 - c_0$ but can also be summarized by the phase-out parameter τ_1) some individuals might start or stop working. This behavioral response at the extensive margin can be expressed by the elasticity:

$$e_1 = \frac{\Delta h_1}{\Delta(1-\tau_1)} \frac{1-\tau_1}{h_1}.$$

Suppose that initially τ_1 is positive, that is, an individual loses some government transfers when taking up work, and suppose further that the government contemplates increasing the in-work benefit by Δc_1 (without modifying transfers at higher incomes). This change, similarly to the one analyzed in Section 2, has three effects on social welfare.

The first is the mechanical revenue effect. The government will pay an additional transfer Δc_1 to a fraction h_1 of individuals, and this is clearly a social cost (here, since these effects are expressed in terms of fractions of the population, social costs and benefits are to be understood in *per capita* terms):

$$\Delta M = -h_1 \Delta c_1 < 0.$$

The second is the welfare effect. Those individuals receiving the additional transfers are made better off, and this is a social welfare gain:

$$\Delta W = g_1 h_1 \Delta c_1 > 0.$$

The last effect to be considered is the behavioral effect. Since the transfers for low-income working individuals increase, the 'financial gains to work' increase as well. Thus, some individuals previously not working are going to take up work. Expressing Δh_1 from the definition of the behavioral elasticity e_1, we can write the behavioral effect as:

$$\Delta B = \tau_1 z_1 \Delta h_1 = e_1 \frac{\tau_1}{1-\tau_1} h_1 \Delta c_1.$$

At the optimal phase-out rate τ_1 the sum of the three effects must be equal to zero. If a small increase of in-work benefits unequivocally increased social welfare, then we would like to increase it further, thus the τ_1 we started out with could not be optimal (and *vice versa*).

If the government values redistribution then $g_1 > 1$, thus $\Delta M + \Delta W$ is always positive. But at positive values of τ_1 the behavioral effect is positive, too, because an increase of in-work benefits induces some individuals to take up work, which reduces government spending. This means that a positive phase-out rate τ_1 cannot be optimal: government should subsidize low-income work. The optimal phase-out rate, implicitly determined by equation $\Delta M + \Delta W + \Delta B = 0$, is given as:

$$\frac{\tau_1}{1-\tau_1} = \frac{1-g_1}{e_1}.$$

This is a qualitative result. It is not attempted here to derive an approximate value of the optimal tax rate (or subsidy) by making further assumptions. Some of these assumptions would have to be more stringent than in Section 2. Instead, the rest of this section discusses how robust this result is to alternative assumptions and what it means for practical tax policy.

3.2 OTHER CONSIDERATIONS AND APPLICATION TO THE HUNGARIAN TAX SYSTEM

The result of the negative entry-level marginal tax rate is softened if intensive adjustment of the labor supply is significant. To see this, note that reducing the marginal tax rate between zero income and z_1 but keeping the tax burden unchanged at higher incomes $\{z_2, z_3, \dots\}$ means that the marginal tax rate must increase between z_1 and z_2. This means that individuals otherwise earning a gross income of z_2 might find it tempting to reduce their effort to earn z_1 because their net loss becomes smaller. This adjustment at the intensive margin is clearly a social loss. Therefore, taking into account this effect moves the optimal entry-level marginal tax rate up (relative to the case when only the extensive margin is taken into account).

As noted earlier, the labor economics literature suggests that at low earnings the dominant margin of labor supply adjustment is the extensive, rather than the intensive, margin. But there is a consideration that, in the Hungarian context, might affect policy conclusions in a similar way. This issue is the practice of some employers to officially hire their workers at the minimum wage and pay part of the regular compensation in cash (a practice sometimes called 'grey employment' in Hungary). That this is an important factor in Hungary is shown by Elek et al. (2011) who estimate that in 2006 about 50% of those minimum-wage earners probably had their income under-reported (the estimation is based on data on firms with at least five employees). While some government measures in 2007 aimed at reducing the number of false minimum-wage earners, it is likely that the phenomenon is still relevant today.

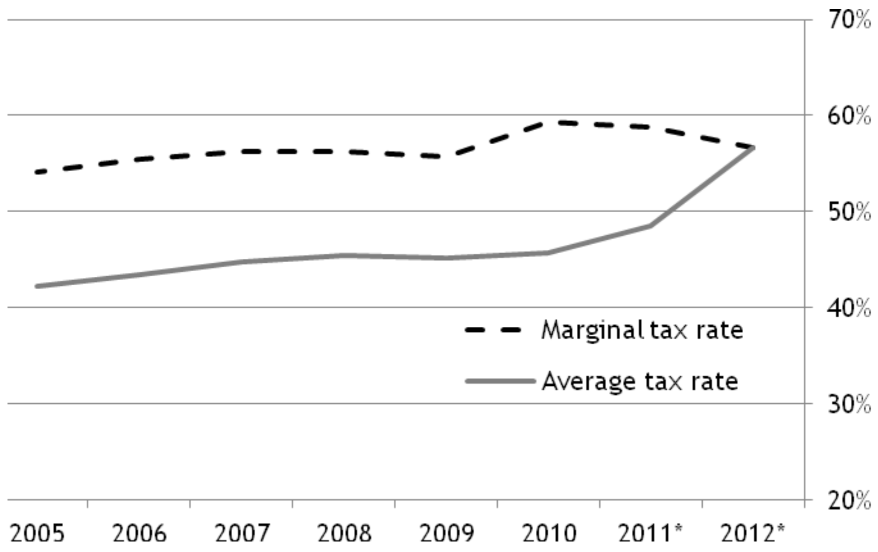
From the point of view of the policy maker, the phenomenon of wage under-reporting has two consequences. First, if the extent of wage-underreporting is responsive to the tax rate, this affects government revenue in the same way as real adjustment of labor supply at the intensive margin: revenues diminish if the marginal tax rate increases because of increasing tax evasion. This constrains the government from heavily subsidizing employment at the entry level. The second consequence of wage underreporting is that the targeting of the wage subsidy is imperfect: there will be individuals who should not receive transfers based on their actual income but they will still receive them based on their reported income. This reduces the desirability of subsidizing employment at the entry level (effectively reducing the average value of parameter ϑ in the recipient group).

In spite of these considerations it seems likely that, since the low employment rate is one of the biggest impediments of the convergence of the Hungarian economy, the entry-level average tax rate should be low (although arguably not negative), to increase the 'financial gains to work' at low incomes.

The elimination in 2012 of the Employee Tax Credit (ETC, *adójóváírás*), lends policy relevance to the issue of the entry-level tax rate. This non-refundable tax credit was designed to reduce (to close to zero) the PIT burden of earnings near the minimum wage. Of course, because of social security contributions and consumption taxes, the effective tax burden on low wages has never been close to zero. Figure 3 shows the effective average and marginal tax rates at the minimum wage in 2005—2012. The parameters of the tax system, including contributions and the Employee Tax Credit, are summarized in Table A2 of the Appendix.

Figure 3.

Actual effective bottom tax rates in Hungary, 2005-2012



Note: Calculations of the author based on statutory tax rates, budget reports and the National Accounts. The effective consumption tax rate for the years 2011—2012 is, for lack of data, taken to be equal to its 2010 value.

As it can be seen in Figure 3, even in the period 2005 to 2010, when the ETC made the minimum wage virtually PIT-free, the average effective tax rate of the minimum wage was above 40% (including, as in Subsection 2.3, employer-side contributions and the effective consumption tax rate) while the marginal effective tax rate was between 50—60%. The effect of the elimination of the ETC is clearly seen in Figure 3 in the substantial increase in the average tax rate; after the elimination of the ETC the average and marginal tax rates applying at the minimum wage are equal.

The consequence of the elimination of the ETC is thus a substantial increase in the average tax rate at low income levels. It has the same effect on the 'financial gains to work', the net gain the individual receives when taking up a job paying the minimum wage (the real-life equivalent of τ_1).

As a consequence, the elimination of the ETC would seem like a move in the right direction if the main issue of the low-wage labor market were adjustment at the intensive margin (or wage under-reporting). But it is probably a move in the wrong direction if adjustment at the extensive margin is the main issue. Based on the government's policy goal of increasing the employment rate to the level of the EU-15 and the U.S., the extensive margin is clearly vital. This, in turn, suggests that in Hungary, like in the U.S. or the U.K., the ETC or a similar measure reducing the average tax burden at low earning levels should is part of the optimal tax system.

4. 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 focuses on two recommendations that can be derived from the theory. First, the optimal top marginal tax rate is likely to be higher than the current actual rate in Hungary. Second, the tax burden of low earnings should be low to encourage employment.

The theory requires surprisingly few assumptions, and only two estimable parameters, to derive a value for the revenue-maximizing 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.

According to simulations in this paper the revenue-maximizing top marginal tax rate is about 67% in Hungary (including, as one should, social security contributions and consumption taxes), consistent with a PIT rate of 40% applying above the pension contribution ceiling (or about 30% if pension contributions were not capped). The top marginal PIT rate is 20% in 2012 and is set to decrease to 16% in 2013. The actual top marginal tax rate is lower than the 'optimal' top marginal tax rate under plausible values of parameter ϑ.

The second conclusion drawn in the paper is that the tax burden on low earnings should be low. The theoretical result is based on the consideration that, absent an intensive-margin response, a reduction in the marginal tax rate on very low incomes has an unequivocally positive welfare effect as long as the tax rate is positive. This conclusion is weakened if an intensive-margin response is present. This may be the case if wage under-reporting behavior (thought to be widespread in Hungary) is sensitive to the tax rates. Nevertheless, even with wage under-reporting the current entry tax burden appears to be too high. The tendencies that can be observed in the tax and transfer system of developed countries (like the Earned Income Tax Credit in the U.S. or the Working Tax Credit in the U.K.) lend support to this conclusion.

REFERENCES

Bakos, Péter, Péter Benczúr and Dóra Benedek (2008): The elasticity of taxable income: Estimates and flat tax predictions using the Hungarian tax changes in 2005. MNB Working Paper 2008/7.

Bakos, Péter, Anikó Bíró, Péter Elek, and Ágota Scharle (2008): The efficiency of the Hungarian tax system. Working Papers in Public Finance 21, Eötvös Loránd University of Sciences, Budapest.

Benczúr, Péter (2007): The effect of tax rates on the behavior of economic agents: A survey of the literature (Title in the original Hungarian: Az adókulcsok hatása a különbözõ gazdasági szereplõk viselkedésére – irodalmi összefoglaló). Közgazdasági Szemle (Hungarian Economic Review) 54 (Feb), 125–141.

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

Benczúr, Péter, Gábor Kátay, and Áron Kiss (2012): Assessing changes of the Hungarian tax and transfer system: A general equilibrium microsimulation approach. MNB Working Paper 2012/7.

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

Brewer, Mike, Emmanuel Saez, 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.

Daragó, Zsófia (2011): Optimal taxation theory and the Hungarian tax system (Title in the original Hungarian: Az optimális adózás elmélete és megvalósulása Magyarországon). Diploma thesis at Corvinus-University, Budapest.

Diamond, Peter (1980): Income taxation with fixed hours of work. *Journal of Public Economics* 13 (1), 101—110.

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.

Elek, Péter, János Köllő, Balázs Reizer, and Péter A. Szabó (2011): Detecting wage under-reporting using a double hurdle model. IZA Discussion Paper No. 6224.

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, pp. 1—32.

Kiss, Áron and Pálma Mosberger (2011): The elasticity of taxable income of high earners: Evidence from Hungary. MNB Working Paper 11/2011.

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

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

Meghir, Costas and David Phillips (2010): Labour Supply and Taxes. 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. 202—274.

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

Piketty, Thomas (1997): La redistribution fiscale face au chomage. *Revue Francaise d'Economie* 12 (1), 157—201.

Sadka, Efraim (1976): On income distribution, incentive effects and optimal income taxation. *Review of Economic Studies* 43 (1), 261—68.

Saez, Emmanuel (2001): Using elasticities to derive optimal income tax rates. *Review of Economic Studies* 68 (1), 205—29

Saez, Emmanuel (2002): Optimal income transfer programs: Intensive versus extensive labour supply responses. *Quarterly Journal of Economics* 117 (2), 1039—73.

Saez, Emmanuel, Joel B. Slemrod 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.

Scharle, Ágota, Péter Benczúr, Gábor Kátay, and Balázs Váradi (2010): How can the efficiency of the tax system be increased? (Title in the original Hungarian: Hogyan növelhető az adórendszer hatékonysága?) MNB Tanulmányok 88, Magyar Nemzeti Bank (Central Bank of Hungary).

Seade, Jesus K. (1977): On the shaps of optimal tax schedules. *Journal of Public Economics* 7, 203—36.

Valentinyi, Ákos (2001): Optimal taxation of capital income (Title in the original Hungarian: A tőkejövedelem optimális adóztatása.) Közgazdasági Szemle (Hungarian Economic Review) 48 (June), 459–479.

APPENDIX — TABLES

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 contributions[a]	8.5%	8.5%	7.5%	7.5%	6.0%	4.0%
Employer contributions	27.0%	27.0%	27.0%	27.0%	29.0%	29.0%
Effective consumption tax rate[b]	15.9%*	15.9%*	15.9%*	15.9%	15.1%	14.8%
Actual top tax rate (without consumption tax)	40.6%	44.0%	43.2%	59.2%	58.1%	53.5%
Actual top tax rate (with consumption tax)	50.0%	52.9%	52.2%	65.7%	64.5%	60.4%
Theoretical revenue-maximizing top tax rate	66.7%	66.7%	66.7%	66.7%	66.7%	66.7%

[a] High-income individuals do not pay pension contributions on income exceeding the pension contribution ceiling. Thus here only the health-care contributions are accounted on the employee side.

[b] 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—2013 the empirical effective consumption tax is not known yet. Thus for these years the value for 2010 was used.

Table A2.

Actual entry-level tax rates in Hungary, 2005-2012

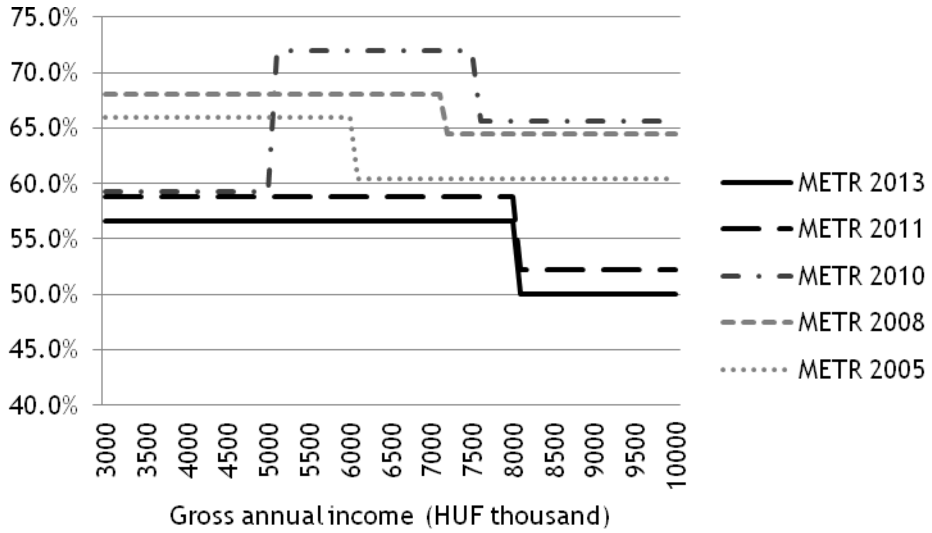
	2012	*2011*	*2010*	*2008*	*2005*
Lower PIT rate	16.0%	20.3%	21.6%	18.0%	18.0%
Monthly minimum wage (HUF)	93,000	78,000	73,500	69,000	57,000
Maximal monthly sum of Employee Tax Credit (HUF)	0	12,100	15,100	11,340	10,260
Monthly taxes due after minimum wage (HUF)	14,880	3,750	769	1,080	0
Average PIT rate at minimum wage	16.0%	4.8%	1.0%	1.6%	0.0%
Employee pension contribution rate	10.0%	10.0%	9.5%	9.5%	8.5%
Employee health care contribution rate	8.5%	7.5%	7.5%	6.0%	4.0%
Employer contribution rate	27.0%	27.0%	27.0%	29.0%	29.0%
Effective consumption tax rate[a]	15.9%	15.9%	15.9%	15.1%	14.8%
Marginal total tax rate at minimum wage	56.6%	58.8%	59.3%	56.2%	54.1%
Average total tax rate at minimum wage	56.6%	48.6%	45.8%	45.4%	42.2%

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—2012 the empirical effective consumption tax rate is not known yet. Thus for these years the value for 2010 was used.

APPENDIX — FIGURES

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.