



University of
St Andrews

School of Economics and Finance Discussion Papers

Lower Tax For Minimum Wage Earners

Jim Jin and Felix FitzRoy

School of Economics and Finance Discussion Paper No. 1703

JEL Classification: H20, D60

Keywords: flat tax, income redistribution, maximin, Pareto improvement

LOWER TAX FOR MINIMUM WAGE EARNERS

Felix FitzRoy*

Jim Jin**

School of Economics and Finance

University of St Andrews

JEL Classification Number: H20, D60

Key Words: flat tax, income redistribution, maximin, Pareto improvement

February 2017

Abstract: We show that minimum wage earners should pay a lower tax than high earners. Though intuitive, this idea is not supported by the existing literature. The optimal maximin tax curve and two-band taxes are usually decreasing. Since decreasing marginal taxes would be unpopular, by continuity a flat tax seems to be superior to increasing marginal taxes and should be a second best solution. However, using a simple utility function and a general income distribution, we find that lowering the marginal tax for minimum wage earners not only dominates the optimal flat tax under maximin, but also make everyone better off.

We thank Tatiana Damjanovic and David Ulph for their helpful comment and discussion. The usual disclaimer applies.

* School of Economics and Finance, University of St Andrews, St Andrews KY16 9AL, Scotland, UK, Tel. (01334)-462425, Fax: (01334)-462444, e-mail: frf@st-andrews.ac.uk.

** Corresponding author: School of Economics and Finance, University of St Andrews, St Andrews KY16 9AL, Scotland, UK, Tel. (01334)-462447, Fax: (01334)-462444, e-mail: jyj@st-andrews.ac.uk.

1. Introduction

In many countries no income tax is levied on the lowest earnings up to a certain level, a politically popular measure to help the poor. As income inequality rises, the threshold for tax exemption has recently been raised to £11,000 in the U.K. Though widely accepted, this measure is not supported by the existing tax literature, although increasing marginal taxes (IMT) are common in developed economies, usually in the form of piece-wise linear tax systems. In this paper we argue that indeed minimum wage earners should be taxed at a lower rate than high earners.

Income redistribution through progressive taxation appears more desirable as inequality has risen to record heights (Stiglitz, 2012; Piketty, 2014; Atkinson, 2015). As the proportion of low wage earners increases, so does the importance of setting appropriate tax rates for them to ensure work incentives and reduce inequality. However the literature provides little justification for uniformly IMT and there is no clear-cut result for the optimal tax structure. With subtle differences in social welfare and income distribution functions, the optimal marginal tax curves can be U-shaped (Diamond (1998), Saez (2001)), or inverse U-shaped (see Sadka (1976), Seade (1977), Tuomala (1984), Hindricks et al, 2006, Kanbur and Tuomala (1994), Boadway et al (2000), Tarkiainen and Tuomala (2007), Hashimzade and Myles (2007), Kaplow (2008) and see a good survey by Jacobs (2013))¹.

Although inverse U-shaped curves suggest lower taxes for the lowest earners, they also imply decreasing tax rates for top earners which are politically infeasible. This political concern and administrative costs may render the continuous tax curves “too far removed from the tax – benefit systems observed in practice to be a useful guide for

¹ One exception is Aaberge and Colombino’s (2013) numerical simulation, based on Norwegian data, which yields the optimal IMT. Their result relies on declining elasticity of labour supply.

policy” (Choné and Laroque 2005, p.396). Diamond and Saez (2011) call for pragmatic research on tax policy. The opposite extreme to the continuous tax curve is a flat tax and the next step is two-band taxes. If we restrict tax rates to be non-decreasing on high income, an inverse U-shaped optimal tax curve might be reduced to increasing two-band taxes. Then one may expect IMT to be justified as a first approximation to the original optimal tax curve.

In fact this is what Sheshinski (1989) initially found when he examined two-band taxes under utilitarian and maximin objectives. However Slemrod et al (1994) later showed that this is not correct and the optimal two-band taxes should be decreasing². Salanie (2003), Hindricks and Myles (2006) obtain similar results in a simple two-class economy. Therefore lower taxes for minimum wage earners are not easily justified. Nonetheless, as we mentioned earlier, decreasing marginal taxes are often viewed as unfair and politically unacceptable. They also imply high marginal tax rates for the lowest income earners and the resulting poverty trap. In contrast a flat tax is politically more acceptable and by continuity it seems to be closer to the optimum than IMT. Hence Mankiw et al (2009) suggest that “A flat tax, with a universal lump-sum transfer, could be close to optimal”. It can save administrative costs and eliminate incentive distortions associated with progressive taxes. Several Eastern European countries have moved in this direction. Atkinson (1995) provides a comprehensive treatment of flat tax. Finally, but not least important, an optimal flat tax is easy to find. Piketty and Saez (2012) obtain the optimal flat tax given a general welfare function and realistic information, though they do not necessarily advocate its implementation³.

²Boadway and Jacquet (2008) also show the optimal marginal tax curve is declining under maximin.

³ They argue for a higher tax for top incomes (approximately the top 5%) based on decreasing marginal utility of income. Jacobs (2013) also argues against a flat tax for various reasons not discussed here. Formal proof that a flat tax is worse than IMT is presented in FitzRoy and Jin (2009), on parts of which this paper is based.

In contrast we argue that any flat tax is Pareto inferior to two-band taxes with a lower rate for minimum wage earners. This result seems to contradict the literature and suggests a non-monotonicity of the optimal tax structure. In a two-group example Slemrod et al (1994) graphically demonstrate that a flat tax can be worse than IMT for both groups. In other words, moving from a flat tax to IMT may be a Pareto improvement. This seems surprising to many economists. In his critique of radical income redistribution through taxation, Mankiw (2013) wrote: “As far as I know, no one has proposed any credible policy intervention to deal with rising inequality that will make everyone, including those at the very top, better off”. The example of Slemrod et al (1994) indicates that such a possibility may exist under a flat tax as advocated by Mankiw et al. (2009). This issue has not been further discussed in the literature.

This paper seeks to fill this gap and show that a Pareto improvement from a flat tax is not only possible, but arises under general conditions, including the most unfavourable assumption of identical elasticity of labour supply and marginal utility of income. We find that given an optimal flat tax under maximin it is always desirable to lower the tax rate for minimum wage earners. Moreover, given any optimal flat tax a tax reduction on minimum wage earners by their proportion in the population must be a Pareto improvement. This result also holds when there is voluntary and involuntary unemployment, or more than two tax bands.

The model is developed in the next section. Section 3 shows that lowering tax for minimum wage earners dominates the optimal flat tax under maximin. In Section 4 we demonstrate that such a simple tax reduction can benefit everyone. Section 5 concludes and proofs are in appendices.

2. The Model

We assume that a population, normalized to unity, consists of a continuum of households, whose wage is denoted by w , and is distributed on $[a, b]$, where $0 < a < b$, and a represents the minimum wage. The wage distribution follows cumulative and density functions $F(w)$ and $f(w)$. A household's pre-tax earnings y are proportional to its wage w and labour supply x , i.e., $y = wx$. The government observes earnings and imposes two tax rates, t_1 and t_2 , when earnings are below or above a threshold \bar{y} . If one's earnings do not exceed \bar{y} , his after-tax earnings are $wx(1 - t_1)$. Otherwise they are $wx(1 - t_2) + \bar{y}(t_2 - t_1)$. The tax revenue is equally distributed as a basic income B to everyone, after a fixed public expenditure P is paid. The unemployed are not included in the unit population and their welfare will be discussed in Section 5.

We assume everyone's utility function is $m - x^{1+1/\varepsilon}/(1 + 1/\varepsilon)$, where m is income and ε is the elasticity of labour supply. Similar utility functions are used in the literature (e.g. Atkinson 1995, Boadway and Jacquet 2008). Our identical elasticity of labour supply for the entire population is obviously less favourable for IMT than empirically plausible declining elasticity, used by Aaberge and Colombino (2013).

Given t_1 , t_2 and \bar{y} , the utility function for different earners can be written as:

$$V_1 = wx(1 - t_1) - \frac{x^{1+1/\varepsilon}}{1+1/\varepsilon} + B \quad \text{for } wx \leq \bar{y} \quad (1)$$

$$V_2 = wx(1 - t_2) + \bar{y}(t_2 - t_1) - \frac{x^{1+1/\varepsilon}}{1+1/\varepsilon} + B \quad \text{for } wx > \bar{y} \quad (2)$$

Every household chooses its labour supply to maximize its utility. When $t_1 < t_2$, we define w_1 and w_2 by $(1 - t_1)^\varepsilon w_1 = \bar{y}$ and $(1 - t_2)^\varepsilon w_2 = \bar{y}$. Then households can be divided into three groups. Those with $w \leq w_1$ will choose labour supply $x = w^\varepsilon(1 - t_1)^\varepsilon$. Those with $w > w_2$ choose $x = w^\varepsilon(1 - t_2)^\varepsilon$. The remaining ones with $w_1 < w \leq w_2$ choose

$x = \bar{y}/w$ and just earn \bar{y} , i.e., bunching. The tax revenues generated from these three groups are $t_1(1 - t_1)^\varepsilon \int_a^{w_1} w^{1+\varepsilon} f(w)dw$, $t_2(1 - t_2)^\varepsilon \int_{w_2}^b w^{1+\varepsilon} f(w)dw - (t_2 - t_1)[1 - F(w_2)] \bar{y}$ and $t_1 \bar{y} \int_{w_1}^{w_2} f(w)dw$ respectively. So the total tax revenue is equal to:

$$R = t_1(1 - t_1)^\varepsilon \int_a^{w_1} w^{1+\varepsilon} f(w)dw + t_2(1 - t_2)^\varepsilon \int_{w_2}^b w^{1+\varepsilon} f(w)dw + \{t_1[1 - F(w_1)] - t_2[1 - F(w_2)]\} \bar{y} \quad (3)$$

We assume the fixed public expenditure P is less than the revenue, so it does not affect our tax decision and the basic income B can be treated as R (minus P). We substitute the optimal labour supply $w^\varepsilon(1 - t_1)^\varepsilon$ into the utility function (1), and let $w = a$, to obtain the utility of minimum wage earners as the maximin objective function.

$$u = \frac{a^{1+\varepsilon}}{1 + \varepsilon} (1 - t_1)^\varepsilon + B \quad (4)$$

Under a maximin objective, we maximize the utility of minimum wage earners. We assume $F(a) > 0$, so minimum wage earners have a positive mass and the maximin is a reasonable social objective. Also part time employees often earn less than full time minimum wage earners. Combining them together we have a significant fraction $F(a)$ in most countries. In UK and USA for instance, this fraction is approximately 20% and 30% respectively⁴. As minimum wages rise in these countries, so will the proportion of minimum wage earners and the importance of the relevant tax policy. In the next section we show the optimal flat tax under maximin is always inferior to two-band IMT.

⁴ In UK the minimum wage was £6.08 per hour in 2011, implying annual earnings of £12,000, exceeding 20% British income (<http://www.hmrc.gov.uk/statistics/tax-statistics/table2-4.pdf>). The US average minimum wage is about \$7.5/hour, or annual earnings \$15,000, more than 30% Americans' 2010 income (http://www.census.gov/hhes/www/cpstable/032011/perinc/new01_001.htm, 2010").

3. Maximin

To justify a lower tax for minimum wage earners, we first find the optimal flat tax under maximin. Piketty and Saez (2012) find an optimal flat tax under a general welfare function, which is a weighted sum of individual welfare. It has a simple form of $(1 - \bar{g})/(1 + \varepsilon - \bar{g})$, where the elasticity of aggregate earnings ε is equivalent to our individual elasticity of labour supply, and \bar{g} is the ratio of the average income weighted by individual social welfare weights to the average income of the population.

When $t_1 = t_2$, the tax revenue in (3) reduces to $t(1 - t)^\varepsilon \int_a^b w^{1+\varepsilon} f(w)dw$. For simplicity, we let $E \equiv \int_a^b w^{1+\varepsilon} f(w)dw$, where the integral includes the mass at a . So E is the total and also average earnings of the population without tax. Then we can write the maximin objective function (4) as $a^{1+\varepsilon}(1 - t)^{\varepsilon+1}/(1 + \varepsilon) + t(1 - t)^\varepsilon E$. Its first-order derivative with respect to t is $(1 - t)^{\varepsilon-1} \{ [1 - (1 + \varepsilon)t]E - a^{1+\varepsilon}(1 - t) \}$, which is positive if and only if t is smaller than the expression t^* given below. Hence we obtain the optimal flat tax under maximin.

$$\textit{Proposition 1: The optimal flat tax under maximin is } t^* = \frac{E - a^{1+\varepsilon}}{(1 + \varepsilon)E - a^{1+\varepsilon}}.$$

The result is a special case of Piketty and Saez (2012), where $a^{1+\varepsilon}/E$ represents their \bar{g} as all the social welfare weight is given to minimum wage earners (they use actual earnings, but the ratio is same under flat tax). The optimal flat tax t^* falls with $a^{1+\varepsilon}/E$ as the gap between the poor and rich is smaller, and falls with ε as a higher elasticity implies more efficiency loss due to taxation.

The next question is whether lowering tax for minimum wage earners can make them better off than the optimal flat tax. The answer seems to be “no” from the existing

literature. With low productivity, minimum wage earners may not gain very much from a lower tax, but lose more in basic income B due to the loss of tax revenue from the rest of the population. However we can show that minimum wage earners can always benefit from a lower tax rate below the optimal tax t^* .

It suffices to show this is true when we let the threshold income $\bar{y} = a^{1+\varepsilon}(1 - t_1)^\varepsilon$. It ensures that minimum wage earners do not face the high tax when choosing optimal labour supply. Given $\bar{y} = a^{1+\varepsilon}(1 - t_1)^\varepsilon$ and our definition of w_1 implies $w_1 = a$. The only people facing the low tax and not bunching are minimum wage earners. If $t_2 > t_1$, we have $w_2 > a$. Starting from t^* , if we lower t_1 , both \bar{y} and w_2 increase. Individuals with w between a and w_2 will be bunching, earning just \bar{y} . We can show that when $t_1 = t_2 = t^*$, $\partial u / \partial t_1 < 0$ (see Appendix A). Hence there always exist some IMT with $t_1 < t_2 = t^*$, that give a higher value of (4), i.e. benefit minimum wage earners.

Proposition 2: Given $\bar{y} = a^{1+\varepsilon}(1 - t_1)^\varepsilon$, lowering t_1 below the optimal flat tax t^ benefit minimum wage earners.*

Thus the optimal flat tax under maximin is dominated by two-band IMT with a lower tax for minimum wage earners. Realistically, however, the government is concerned not only with the wellbeing of minimum wage earners, but also the society as a whole. The maximin optimal flat tax t^* is not the one advocated by Mankiw et al. (2009). To justify lower tax for minimum wage earners, we need to compare it with an optimal flat tax in a more general sense and use a more robust criterion to demonstrate its superiority. In the next section we will show that starting from any optimal flat tax, a lower tax rate for minimum wage earners can make everyone better off.

4. Pareto dominance

The maximin objective is not politically feasible, so the question is whether some tax reduction for minimum wage earners can make everyone better off than under a flat tax. If so, we would have a strong case for such a tax reduction. We will show that reduction of t_1 is indeed a Pareto improvement over any flat tax. The flat tax need not be t^* , but any optimal tax under a general social welfare function.

We follow Piketty and Saez (2012) to construct a general social welfare function. Let $s(w) (\geq 0)$ be the social welfare weight attached to households with wage w , subject to $\int_a^b s(w)f(w)dw = 1$. Following (1) and (2), under a flat tax t , a household's net utility is $u(w) = w^{1+\varepsilon}(1-t)^{\varepsilon+1}/(1+\varepsilon) + B$, where B is equal to $t(1-t)^\varepsilon E - P$. Hence the general social welfare function can be written as

$$\int_a^b s(w)u(w)f(w)dw = \frac{(1-t)^{1+\varepsilon}}{1+\varepsilon} \int_a^b w^{1+\varepsilon} s(w)f(w)dw + t(1-t)^\varepsilon E - P \quad (5)$$

As we mentioned earlier, our elasticity of labour supply plays the same role as the elasticity of aggregate earnings in Piketty and Saez (2012). Differentiating the welfare function (5) with respect to t , we see that its value increases with tax rate t if and only if $t < [E - \int_a^b w^{1+\varepsilon} s(w)f(w)dw] / [(1+\varepsilon)E - \int_a^b w^{1+\varepsilon} s(w)f(w)dw]$, which is the optimal flat tax. Letting $\int_a^b w^{1+\varepsilon} s(w)f(w)dw / E \equiv \bar{g}$, the optimal tax maximizing (5) can be written in the same form as in Piketty and Saez (2012):

$$\textit{Proposition 3: The general optimal flat tax } \hat{t} = \frac{1 - \bar{g}}{1 + \varepsilon - \bar{g}}.$$

The welfare weight $s(w)$ should be non-increasing with w . If $s(w) = 1$, we have the utilitarian welfare function. Thus $\bar{g} = 1$ and the optimal flat tax is zero. If $s(w) = 0$ except for $w = a$, we have $\bar{g} = a^{1+\varepsilon}/E$, \hat{t} becomes the previous optimal flat tax t^* under

maximin. In general the value of \bar{g} lies between 1 and $a^{1+\varepsilon}/E$, and \hat{t} is between zero and t^* . Conversely, any flat tax within this range can be viewed as an optimal tax under a certain social welfare function associated with a non-increasing social welfare weight function $s(w)$. Given a general optimal flat tax \hat{t} , we show that a tax reduction for minimum wage earners can provide a Pareto improvement.

We already know that a reduction in t_1 benefits minimum wage earners. Then we consider the next income group, i.e. individuals with wage between a and w_2 . They only work up to the threshold earnings $\bar{y} = a^{1+\varepsilon}(1 - t_1)^\varepsilon$, i.e. bunching. Their labour supply is $a^{1+\varepsilon}(1 - t_1)^\varepsilon/w$. When t_1 falls, they pay a lower tax rate and gain from increased labour supply due to a higher \bar{y} . Substituting these variables into (1), we get their utility function.

$$U_1 = a^{1+\varepsilon}(1 - t_1)^{\varepsilon+1} - \frac{\varepsilon a^{1+\varepsilon}}{1 + \varepsilon} \left(\frac{a}{w}\right)^{1+1/\varepsilon} (1 - t_1)^{\varepsilon+1} + B \quad (6)$$

Finally we consider high income earners with $w > w_2$. Their labour supply is not affected by t_1 . When t_1 falls and $\bar{y} = a^{1+\varepsilon}(1 - t_1)^\varepsilon$ rises, they will benefit from a lower infra-marginal tax rate and a higher tax threshold \bar{y} . Substituting $\bar{y} = a^{1+\varepsilon}(1 - t_1)^\varepsilon$ into (2), we can write their utility function as:

$$U_2 = \frac{w^{1+\varepsilon}}{1 + \varepsilon} (1 - t_2)^{1+\varepsilon} + a^{1+\varepsilon}(t_2 - t_1)(1 - t_1)^\varepsilon + B \quad (7)$$

Differentiating (6) and (7) with respect to t_1 , we find that both derivatives are negative when $t_1 = t_2 = \hat{t}$ (see Appendix B). Hence (6) and (7) rise when t_1 falls and we have a Pareto improvement over the optimal flat tax \hat{t} .

Proposition 4: Lowering t_1 is a Pareto improvement over the optimal flat tax \hat{t} .

Our result generalize the insight from the example of Slemrod et al (1994) that any flat tax is Pareto inferior to some IMT. This outcome depends on the assumption of $F(a) > 0$, i.e. a positive mass of minimum wage earners. This may seem surprising, but the intuitive explanation is straightforward. When we lower the tax rate for minimum wage earners, the high income group will not change their labour supply. The other two groups, the bunching group and minimum wage earners, will increase their labour supply, which is $a^{1+\varepsilon}(1-t_1)^\varepsilon/w$ and $a^\varepsilon(1-t_1)^\varepsilon$ respectively. Everyone benefits from a lower tax but receives less income transfer, i.e. the basic income B . If minimum wage earners have a positive mass, their increased labour supply will increase the total output, which ensures the total benefit exceeds the reduction in tax revenue. As everyone benefits at least from a tax reduction of $\bar{y} \Delta t_1$, and loses an equal amount of basic income, no one can be worse off. This intuitive argument applies not only in our model with a specific utility function, but also in the general case.

Although everyone benefit from such a tax reduction, they may not benefit equally. One may expect that minimum wage earners benefit the most since they are the main target of tax reduction. This is not necessarily true as higher wage earners also pay less tax even though they may not adjust their labour supply very much. In fact, if we compare the derivatives of the utility respect to t_1 in Appendices A and B, it is easy to see that $\partial U_1/\partial t_1 < \partial u/\partial t_1$ always and $\partial U_2/\partial t_1 < \partial u/\partial t_1$ given $t_1 < t_2$. Hence these two groups will benefit more than minimum wage earners do when t_1 falls.

Proposition 5: Higher income earners benefit more from the tax reduction on minimum wage earners.

This should reduce any political resistance against the tax reduction. This finding can be compared with Sadka (1976), and Seade (1977), who show that the tax rate on the

top earnings should be zero. While the zero tax rate only applies to one individual in their case, our tax reduction should be applied to a positive mass, and thus can bring in a significant improvement.

So far we have only shown the positive impact of a marginal tax reduction for minimum wage earners. Generally this impact may change when the tax reduction becomes significant. For a practical policy recommendation, it is desirable to know if this positive effect can be sustained for a discrete change in the tax rate. Hence we will show that this is guaranteed for a tax reduction by the amount of $F(a)\hat{t}$.

We first consider the impact of a reduction in t_1 by $F(a)\hat{t}$ on minimum wage earners. In Appendix A when we keep $t_2 = \hat{t}$ and let t_1 fall from \hat{t} to $[1 - F(a)]\hat{t}$, (A3) implies that $\partial u / \partial t_1 = -\varepsilon m(1 - t_1)^{\varepsilon-1} \hat{t} [F(w_2) - F(a)] < 0$. So u rises as t_1 falls, and minimum wage earners are better off with the tax reduction. Similarly, the derivatives of U_1 and U_2 respect to t_1 are both negative when t_1 falls from \hat{t} to $[1 - F(a)]\hat{t}$. Hence (6) and (7) must rise when t_1 falls by $F(a)\hat{t}$. Therefore we obtain:

Proposition 6: Reducing the tax rate on minimum wage earners from the optimal flat tax \hat{t} by $F(a)\hat{t}$ is guaranteed to be a Pareto improvement.

Our finding leads to a simple policy recommendation: the tax rate on minimum wage earners should be lower than the next tax rate by a proportion of $F(a)$, which is their proportion in the population. The group, including minimum wage earners and most part-timers, is roughly 20% in the UK (2011) and 30% in the USA (2010). Their full time earnings are roughly £12,000 and \$15,000, with marginal tax rates 20% (on income from £11,000 to £35,000) and 15% (up to \$34,000) respectively. In the UK, national insurance contributions of 12% on incomes over £8,000 raise the effective tax rate to 32% on income over £11,000. Our result suggests that everyone will be better off if the tax rates

up to £12,000 and \$15,000 fall to 16% and 10.5%. Tax reduction of $F(a)\hat{t}$ is not the limit for a Pareto improvement. Even though a further reduction may not benefit everyone, it can be a Pareto improvement over the flat tax. In addition to the positive effect on employment, we may justify zero tax for low income earners, such as the income-tax free allowance of £11,000 in U.K.

We see from the proof that the Pareto improving tax reduction needs not be from the optimal tax \hat{t} , but any flat tax. Furthermore, it may also be from any multi-band taxes. We may assume that beyond \bar{y} , there are other income thresholds, above which other tax rates apply. Then the tax revenue R in (3) remains the same except for the third term, $t_2(1 - t_2)^\varepsilon \int_{w_2}^b w^{1+\varepsilon} f(w)dw$ becomes multiple integrations with different tax bands. However these new terms will not be affected by t_1 . So when we calculate $\partial B/\partial t_1$, the result remains the same. Hence the tax reduction for minimum wage earners is still Pareto improving under a general multi-band tax system.

Finally we argue that the tax reduction for minimum wage earners remains a Pareto improvement when there is unemployment, either voluntary or not. So far we have not explicitly considered unemployment, while the fixed public expenditure may include the unemployment benefit. If the unemployment is involuntary, the tax reduction will not affect it. The voluntarily unemployed may choose to work after the tax reduction. Given the same unemployment benefit and a Pareto improvement for all employed, when the unemployment falls, the extra tax revenue can be used to increase the unemployment benefit. So the tax reduction is still a Pareto improvement.

5. Conclusions

Our results suggest that lowering the tax rate for minimum wage earners always dominates any flat tax under maximin. Moreover a tax reduction for minimum wage

earners can benefit everyone, and hence can be justified under any social welfare function. This result seems intuitive but to the best of our knowledge it has not been reported in the literature. Of course our results depend on a basic income, quite different from existing benefit systems, and income taxes are generally progressive. However, there is growing interest in basic income, often in conjunction with a flat tax. Perhaps of more relevance, total (direct and indirect) average tax paid in the UK is a roughly similar percentage of gross income across the distribution.

The welfare gain from the tax reduction may be small in our model, but is likely to be greater in more realistic cases. For instance we do not consider decreasing marginal utility of income as Piketty and Saez (2012) do, which provides further social justification for progressive taxes. Also, differing from Aaberge and Columbino (2013), we assume an identical elasticity of labour supply for the entire population. In reality its value for full-time and high income earners is much lower than that for low income earners. This means less efficiency loss due to taxation on high earners. In particular, we do not consider the crucial participation decision of low income earners, which generates the ‘poverty trap’ under high implicit marginal tax rates as benefits are phased out. A lower tax for minimum wage earners will reduce the disincentives to work, and generate potentially more wealth for the society. These factors may further justify the tax reduction, but are omitted from our tax model for simplicity and tractability. Tax reduction for low earners and its impact become more significant with a large $F(a)$, which has increased in many developed countries after the recent financial crisis and provides additional support for tax reform.

References

- Aaberge, R. and U. Colombino, 2013, Using a Microeconomic Model of Household Labour Supply to Design Optimal Income Taxes, *Scandinavian Journal of Economics*, vol. 115, pp. 449-475
- Atkinson, A. B., *Public Economics in Action: The Basic Income/Flat Tax Proposal*, Clarendon Press, Oxford, 1995
- Boadway, R, K. Cuff, and M. Marchand, 2000, Optimal Income Taxation with Quasi-Linear Preferences Revisited, *Journal of Public Economic Theory*, Vol. 2, pp. 435-60
- Boadway, R, and L. Jacquet, 2008, Optimal Marginal and Average Income Taxation under Maximin, *Journal of Economic Theory*, vol. 143, pp. 425-441
- Chone, P., and G. Laroque, 2005, Optimal Incentives for Labour Force Participation, *Journal of Public Economics*, vol. 89(2-3), pp. 395–425
- Diamond, P. A., 1998, Optimal Income Taxation: An Example with a U – shaped Pattern of Optimal Marginal Rates, *American Economic Review*, vol. 88(1), pp. 83–95
- Diamond, P. A., and Saez, E., 2011, The Case for a Progressive Tax: From Basic Research to Policy Recommendations, *The Journal of Economic Perspectives*, Vol. 25, No. 4 (Fall 2011), pp. 165-190
- FitzRoy, F. and J. Jin, 2009, Are Two Taxes better than One? IZA dp. 4167
- Hashimzade, N. and G. Myles, 2007, Structure of the Optimal Income Taxation in the Quasi-Linear Model, *International Journal of Economic Theory*, vol. 3, pp. 5-33
- Hindricks, J., E. Lehmann and A. Parmentier, 2006, Optimal Income Taxation and the Shape of Average Tax Rates, *Economic Bulletin*, vol. 8, pp. 1-6
- Hindricks, J. and G. Myles, *Intermediate Public Economics*, MIT Press, London, 2006
- Jacobs, B., 2013, From Optimal Tax Theory to Applied Tax Policy, CES-Info Working Paper, #4151.
- Kanbur, R. and M. Tuomala, 1994, Inherent Inequality and the Optimal Graduation of Marginal Tax Rates, *Scandinavian Journal of Economics*, vol. 96, pp. 275-82
- Kaplow, L., *The Theory of Taxation and Public Economics*, Princeton University Press, Princeton and Oxford, 2008
- Mankiw, N. G., M. Weinzierl & D. Yagan, 2009, “Optimal taxation in Theory and Practice”, NBER Working Paper No. 15071
- Mankiw, N. G., 2013, “Defending the One Percent”, *Journal of Economic Perspective*, 27(3): pp. 21-34.
- Piketty, T. and E. Saez, 2012, Optimal Labour Income Taxation, NBER working paper 18521.
- Sadka, E., 1976, On Income Distribution, Incentive Effects and Optimal Income Taxation, *Review of Economic Studies*, vol. 43, pp. 261-8
- Saez, E., 2001, Using Elasticities to Derive Optimal Income Tax Rates, *Review of Economic Studies*, vol. 68, pp.205–29
- Salanié, B., *The Economics of Taxation*, MIT Press, 2003
- Seade, J., 1977, On the Shape of Optimal Income Schedules, *Journal of Public Economics*, vol. 7, pp. 203–36
- Sheshinski, E., 1989, Note on the Shape of the Optimum Income Tax Schedule, *Journal of Public Economics*, vol. 40, pp. 201–15
- Slemrod, J., S. Yitzhaki, J. Mayshar, and M. Lundholm, 1994, The Optimal Two–Bracket Linear Income Tax, *Journal of Public Economics*, vol. 53, pp. 269–90
- Stiglitz, J., 1982, Self-selection and Pareto Efficient Taxation, *Journal of Public Economics*, vol. 17, pp. 213–40
- Tarkiainen, R. and M. Tuomala, 2007, On Optimal Income Taxation with Heterogeneous Work Preference, *International Journal of Economic Theory*, vol. 3, pp. 35-46
- Tuomala, M., 1984, On the Optimal Income Taxation: Some Further Numerical Results, *Journal of Public Economics*, vol. 23, pp. 351–66

Appendix A, Proof of Proposition 2:

Starting from t^* , we show $\partial u/\partial t_1 < 0$. Given $t_1 \leq t_2 = t^*$ and $\bar{y} = a^{1+\varepsilon}(1-t_1)^\varepsilon$, i.e., $w_1 = a$,

(3) implies the tax revenue can be written as:

$$R = t_1(1-t_1)^\varepsilon a^{1+\varepsilon} + t_2(1-t_2)^\varepsilon \int_{w_2}^b w^{1+\varepsilon} f(w)dw - t_2(1-t_1)^\varepsilon a^{1+\varepsilon}[1-F(w_2)] \quad (\text{A1})$$

As $w_2^{1+\varepsilon}(1-t_2)^\varepsilon = a^{1+\varepsilon}(1-t_1)^\varepsilon$, we have $\partial R/\partial w_2 = -t_2(1-t_2)^\varepsilon w_2^{1+\varepsilon} f(w_2) + t_2(1-t_1)^\varepsilon a^{1+\varepsilon}$

$f(w_2) = 0$. So we ignore indirect impacts via w_2 when differentiate R with respect to t_1 .

$$\frac{\partial B}{\partial t_1} = (1-t_1)^{\varepsilon-1} a^{1+\varepsilon} \{1 - (1+\varepsilon)t_1 + \varepsilon t_2[1-F(w_2)]\} \quad (\text{A2})$$

We substitute these results when differentiating u in (4) and obtain

$$\frac{\partial u}{\partial t_1} = (1-t_1)^{\varepsilon-1} a^{1+\varepsilon} \varepsilon \{t_2[1-F(w_2)] - t_1\} \quad (\text{A3})$$

When $t_2 = t_1 = t^*$, $\partial u/\partial t_1 < 0$ as $t_2 F(w_2) = t^* F(a) > 0$.

Appendix B, Proof of Proposition 4:

We first differentiate (6) with respect to t_1 , and get

$$\frac{\partial U_1}{\partial t_1} = -a^{1+\varepsilon}(1-t_1)^{\varepsilon-1} [1 + \varepsilon - \varepsilon \left(\frac{a}{w}\right)^{1+1/\varepsilon}] + \frac{\partial B}{\partial t_1} \quad (\text{B1})$$

Substituting $\partial B/\partial t_1$ in (A2) into (B1), we obtain

$$\frac{\partial U_1}{\partial t_1} = -\varepsilon a^{1+\varepsilon}(1-t_1)^{\varepsilon-1} \left\{ \left(1 - \left(\frac{a}{w}\right)^{1+1/\varepsilon}\right)(1-t_1) + t_1 - t_2[1-F(w_2)] \right\} \quad (\text{B2})$$

When $t_2 = t_1 = \hat{t}$, (B2) is negative if $[1 - (a/w)^{1+1/\varepsilon}](1 - \hat{t}) + \hat{t} F(a) > 0$. As $a < w$, this inequality holds. So U_1 must rise as t_1 falls. Differentiating (7) with respect to t_1 we get

$$\frac{\partial U_2}{\partial t_1} = -a^{1+\varepsilon}(1-t_1)^{\varepsilon-1} [1 - t_1 + \varepsilon(t_2 - t_1)] + \frac{\partial B}{\partial t_1} \quad (\text{B3})$$

Plugging (A2) into (B3), we obtain $\partial U_2/\partial t_1 = -\varepsilon a^{1+\varepsilon}(1-t_1)^{\varepsilon-1} t_2 F(w_2) < 0$.