A Case for Maximum Wage[♣]

Tomer Blumkin[*] **Efraim Sadka**[♠] **Yotam Shem-Tov**[♦]

April, 2013

Abstract

In this paper we demonstrate that supplementing the optimal non-linear income tax system with a binding maximum wage rule attains a *Pareto* improvement, by serving to mitigate the mimicking incentives of the high-skill individuals without entailing distortions.

JEL Classification: D6, H2, H5

Key Words: redistribution, maximum wage

[♣] We are grateful to an anonymous referee for helpful comments and suggestions. The usual disclaimer applies.

[*] Department of Economics, Ben-Gurion University, Beer-Sheba 84105, Israel, CesIfo, IZA. E-mail: tomerblu@bgumail.bgu.ac.il.

[♠] The Eitan Berglas School of Economics, Tel Aviv University, Tel-Aviv 69978, Israel, CesIfo, IZA. E-mail: sadka@post.tau.ac.il

[♦] The Eitan Berglas School of Economics, Tel Aviv University, Tel-Aviv 69978, Israel, E-mail: yotam.she@gmail.com

1. Introduction

The idea of using maximum wage rules is not a new one. It reaches back to Aristotle, who suggested that no one should have more than five times the wealth of the poorest person. During the second world-war, concerned by war profiteering, Franklin D. Roosevelt proposed a maximum income of 25,000 USD in 1942, accompanied by a 100 percent tax on all income above this level. More recently, the debate about enacting maximum wage rules and salary caps has been revived in earnest in light of the stratospheric rise in chief-executives' pay relative to the median earners over the last decade.[1]

Following Mirrlees (1971) the literature on redistribution usually assumes that there exists some underlying distribution of earning capacities which is the source of inequality and the reason for government intervention on equity grounds. A possible channel via which the government can affect earnings inequality is by designing policy rules that directly influence this underlying distribution supplementing the tax-and-transfer system. One such policy rule which has received much attention is minimum wage legislation [see Lee and Saez (2012) for a broad review of this literature]. Somewhat surprisingly, the literature has by and large overlooked the potential re-distributive role played by a binding maximum wage rule that sets an upper bound on the compensation of the high-skill (wealthiest) individuals. The objective of this paper is to examine the potential welfare-enhancing role of a maximum wage rule as a supplement to the tax-and-transfer system. Embedding a maximum wage rule into a standard optimal non-linear tax framework, we demonstrate that supplementing the labor income tax system with a binding maximum wage would result in a *Pareto* improvement and explain the mechanism at work.

[1] An indirectly related issue is the implementation of salary cap arrangements, which are common in major sports leagues in the world (e.g., the NBA, the NHL and the NFL). The main objective of these arrangements is to attain enhanced competition (promotion of 'equal opportunities') across teams.

2. The model

In order to simplify the exposition we choose the simplest setting possible.[2] Consider an economy which produces a single consumption good employing labor inputs with different skill levels. We assume that the production technology exhibits constant returns to scale and perfect substitution across skill levels. Individuals differ only in their innate earning ability/skill level. We assume that population is equally divided between low- and high-skill individuals and denote by w_1 and w_2, their respective abilities (and corresponding real wage rates in a competitive labor market), where $w_2 > w_1 > 0$. We normalize the population of each skill level to unity with no loss in generality. We follow Mirrlees (1971) by assuming that skill levels are private information unobserved by the government.

Individuals' preferences are represented by the following additively *separable* utility function:

(1) $U(c,l) = g(c) - h(l)$,

where c denotes consumption, l denotes labor, $h(\cdot)$ is strictly increasing and strictly convex, and $g(\cdot)$ is strictly increasing and strictly concave. INADA conditions are assumed to ensure interior solutions throughout.

For later purposes, as is common in the optimal tax literature, we reformulate the utility function and represent it as a function of gross income (y), net income (c) and the individual skill level (w):

[2] Our analysis carries over in a straightforward manner to more general specifications of the production technology, individuals' preferences and the number of skill levels considered.

(2) $V(w,c,y) \equiv g(c) - h(y/w)$.

3. The government problem

We first introduce our benchmark setting. The government is assumed to maximize a weighted *Utilitarian* welfare function given by:

(3) $W \equiv \alpha \cdot V(w_1, c_1, y_1) + (1-\alpha) \cdot V(w_2, c_2, y_2)$,

where α , 0< α <1, denotes the weight assigned to the low skill individual in the social welfare function and is assumed to be sufficiently large,[3] subject to a revenue constraint (assuming with no loss of generality no revenues needs),

(4) $(y_1 - c_1) + (y_2 - c_2) \geq 0$,

and two self-selection constraints, ensuring that each type of individual is as well-off with his bundle as he would be with mimicking the other type:

(5) $V(w_1, c_1, y_1) \geq V(w_1, c_2, y_2)$,

(6) $V(w_2, c_2, y_2) \geq V(w_2, c_1, y_1)$.

The constrained optimization program is fairly standard and technical details are therefore omitted for abbreviation purposes [for details see e.g., Balcer and Sadka (1982) and Stiglitz

[3] Two remarks are in order. First, the set of optimal tax schedules associated with maximizing a weighted *utilitarian* social welfare function for different weights is equivalent to the set of second-best *Pareto* optimal tax schedules [examined in Stiglitz (1982)]. Second, assuming α is large enough ensures that the government is redistributing from the high-skill towards the low skill, in which case the binding incentive constraint in the government optimization program would be that of the high-skill individual. In the case not considered where α is sufficiently small, the redistribution would go in the other direction, hence the binding incentive constraint would be that of the low-skill. In such a case, as shown by Allen (1982), a *minimum* wage would be a desirable supplement to the optimal non-linear income tax schedule.

(1982)]. The standard properties hold: both the revenue constraint (4) and the self-selection constraint of the high-skill individual (6) are binding, the marginal tax rate levied on the high-skill individuals is zero (efficiency at the top); whereas, the marginal tax rate imposed on the low-skill individuals is strictly positive (downward distortion at the bottom). We turn next to introduce our new instrument and prove our main result.

4. A case for maximum wage

Suppose that the government sets an upper bound on the hourly wage rate paid in the labor market. Formally let the upper bound wage rate be denoted by $\overline{w}$, where $w_1 < \overline{w} < w_2$. Several remarks are in order. First, notice that any firm hiring the labor services of high-skill workers would earn strictly positive profits. Clearly, these rents (pure profits) could be taxed away by the government, without entailing any distortions. Second, as noted by Lee and Saez (2012) in the context of minimum wage [see also a related discussion in Blumkin et al. (2007) in the context of anti-discrimination rules], there exists, apparently, some informational inconsistency between the implementation of any policy directly regulating the wage rates and policy focusing on affecting the distribution of earnings (a labor income tax). In reality governments do combine minimum wage policies (based on hours of work) and income taxes (based on earnings). Hence, Lee and Saez (2012) find it useful to consider the constrained optimization problem combining taxes on earnings and minimum wage rates.[4] We follow suit, by allowing the government to combine a maximum wage policy as a supplement to an income tax.

[4] In order to render our analysis tractable we eschew from incorporating misreporting considerations and the entailed enforcement/compliance issues with respect to the maximum wage rule, without discounting their importance. Notice that by doing so we follow the bulk of the literature on optimal taxation and minimum wage legislation, which raises similar concerns about cheating. Clearly and as correctly pointed out by the referee, the incentives to cheat become stronger as the level of taxation of firms' profits increase.

Third, to ensure the effective implementation of the maximum wage rule it is assumed that the government takes into account all forms of remuneration paid by the firm to the high skill worker (measured per hour of work) including salary, bonuses and benefits. Finally, our choice to confine attention to (a pure) intensive margin model (choice of the hours of work rather than the decision on labor market participation) stems from our focus on policy regulating the wage rate of the high skill worker (the upper end of the skill distribution). We turn now to re-formulate the government constrained optimization program in the presence of a maximum wage rule. The government is seeking to maximize the following welfare function:

(3') $$\mathrm{W} \equiv \alpha \cdot \mathrm{V}(\mathrm{w}_1, \mathrm{c}_1, \mathrm{y}_1) + (1-\alpha) \cdot \mathrm{V}(\overline{\mathrm{w}}, \mathrm{c}_2, \mathrm{y}_2),$$

subject to a revenue constraint,

(4') $$(y_1 - c_1) + (y_2 - c_2) + \frac{y_2}{\overline{w}} \cdot (w_2 - \overline{w}) \geq 0,$$

and two self-selection constraints:

(5') $$V(w_1, c_1, y_1) \geq V(w_1, c_2, y_2),$$

(6') $$V(\overline{w}, c_2, y_2) \geq V(\overline{w}, c_1, y_1).$$

Notice that the third term on the left-hand side of the revenue constraint captures the profits earned by the firms (by our constant returns to scale assumption, we may consider a representative firm). Further notice that it is implicitly assumed that the government fully taxes the entire rents accrued by the firms due to the imposition of the maximum wage rule.[5] Notice

[5] Our result will hold even with less than a confiscatory level of profit taxation. In such a case two issues arise. The first is a mechanical (fiscal) effect that derives from the decrease in government revenues due to the decrease in the

further that when $\overline{w} = w_2$; namely, the maximum wage rate is non-binding, the new formulation coincides with the standard one given in section 3 [equations (3)-(6)].

We turn now to state and prove our main result.

Proposition: Supplementing the non-linear income tax with a maximum wage would result in a *Pareto* improvement; making individuals of both skill-levels strictly better off.[6]

Proof: Let the optimal solution for the government optimization program in the absence of a maximum wage rule be denoted by the two bundles $< c_1, y_1 >$ and $< c_2, y_2 >$ self-selected by the low- and high-skill individuals, respectively. To prove our argument we first show that by introducing a binding maximum wage and by slightly modifying the two bundles given above, one can create a fiscal surplus without violating the two self-selection constraints given in equations (5') and (6'). We then show that this surplus could be used to make both types strictly better off.

Let $\overline{w}$ denote the binding maximum wage rate, where $w_2 - \overline{w} = \varepsilon_1$ for some small $\varepsilon_1 > 0$. Now consider a small perturbation around the optimal solution for the government optimization with

level of profit taxation which yields a violation of the revenue constraint. Another one concerns an income effect due to the positive rents accrued by the firms and collected by the owners, which, in the more plausible case where the residual claimant is the high-skill individual, will yield a violation of his incentive compatibility constraint (due to diminishing marginal utility from consumption). However, as shown below (see the proof of the proposition) by introducing a maximum wage rule and modifying the tax schedule the government manages to create a fiscal surplus. Then, by continuity considerations, by only slightly reducing the rate of profit taxation below the confiscatory rate, both the reduction in revenues and the income effect are small in magnitude, so that both could be addressed by resorting to the fiscal surplus. Hence, neither the revenue constraint nor the incentive constraint of the high-skill individual is violated.

[6] The proposition demonstrates the desirability of supplementing the non-linear labor income tax with a maximum wage rule. A different question which can be addressed is regarding the optimal maximum wage rate. It is straightforward to show that as long as the maximum wage rate exceeds the low-skill wage rate, by slightly reducing the maximum wage level and adjusting the tax system in manner similar to that described in the proof of the proposition a fiscal surplus can be created and disbursed in a way that renders individuals of both skill types strictly better off. The optimal solution is therefore to set the maximum wage rate equal to the low-skill wage rate. This will ensure that the utility derived by both types of individuals is equalized.

no maximum wage in place, given by (an illustrative figure accompanied by brief explanations can be found in the appendix):

$$
\begin{aligned}
&c_2' = c_2, \\
&y_2' = \frac{y_2}{w_2} \cdot \bar{w}, \\
(7) \qquad &c_1' = c_1 + \varepsilon_2, \\
&y_1' = y_1 + \varepsilon_3,
\end{aligned}
$$

where $\varepsilon_2, \varepsilon_3 > 0$ are small and satisfy the following two conditions:

$$(8) \qquad \varepsilon_2 \cdot g'(c_1) - \varepsilon_3 \cdot h'(y_1 / w_1) \cdot \frac{1}{w_1} = 0,$$

$$(9) \qquad \varepsilon_2 \cdot g'(c_1) - \varepsilon_3 \cdot h'(y_1 / w_2) \cdot \frac{1}{w_2} - \varepsilon_1 \cdot h'(y_1 / w_2) \cdot \frac{y_1}{w_2^{\;2}} = 0.$$

A few remarks are in order. First, notice that the modified bundle offered to the high-skill individual maintains the same level of utility (derived by the high-skill) as in the original bundle. The consumption level remains unchanged by construction and so is the working time (the gross income is downward adjusted to the binding maximum wage rate to maintain the same number of working hours). Second, notice that the modified bundle offered to the low-skill individual also maintains the same level of utility (derived by the low-skill) as in the original bundle. This derives from condition (8) which is a first-order approximation to the total change in the utility of the low-skill associated with the suggested perturbation. Essentially the condition in (8) captures an upward shift along the indifference curve of the low-skill individual stemming from his original (unperturbed) bundle (see the figure). Finally, notice that the self-selection constraint of the high-skill remains binding by virtue of the condition in (9) which provides a first-order approximation to the total change in the utility of a high-skill individual that is choosing to mimic his low-skill counterpart, associated with the suggested perturbation. This condition takes into account two conflicting effects which offset one another. The first is the effect associated with imposing a maximum wage, captured by the third term on the left-hand side of (9). This term is negative and works in the direction of making mimicking less attractive for the high-skill, as he has to spend more hours at work in order to obtain the same level of income (being rewarded by a lower hourly wage rate). The second effect is captured by the sum of the first and second terms on the left-hand side of (9). This effect is positive and hence works in the direction of relaxing the incentive compatibility constraint faced by the high-skill (making it easier for him to mimic). To see this formally notice that:

(10) $$\varepsilon_2 \cdot g'(c_1) - \varepsilon_3 \cdot h'(y_1 / w_2) \cdot \frac{1}{w_2} > \varepsilon_2 \cdot g'(c_1) - \varepsilon_3 \cdot h'(y_1 / w_1) \cdot \frac{1}{w_1} = 0,$$

where the equality sign follows from condition (8) and the inequality sign follows from the strict convexity of h and the fact that $w_2 > w_1$. In words, the inequality sign follows from the single-crossing property and the fact that the indifference curve of the low-skill individual is steeper than that of his high-skill counterpart (see the figure). Thus the perturbation which amounts to an upward shift along the indifference curve of the low-skill individual makes it more attractive for the high-skill individual to mimic.

We next turn to verify that we create a fiscal surplus. Formally, we need to show that (recalling that the third term captures the profits earned by the representative firm, which are fully taxed by the government):

(11) $$(y'_1 - c'_1) + (y'_2 - c'_2) + \frac{y'_2}{\overline{w}} \cdot (w_2 - \overline{w}) > 0.$$

Substituting from (7) and using the fact that $w_2 - \overline{w} = \varepsilon_1$, one obtains:

(12) $$[(y_1+\varepsilon_3)-(c_1+\varepsilon_2)]+\left[\frac{y_2}{w_2}\cdot(w_2-\varepsilon_1)-c_2\right]+\frac{y_2}{w_2}\cdot\varepsilon_1>0 .$$

Employing the fact that $(y_1-c_1)+(y_2-c_2)=0$; namely, the revenue constraint binds, by virtue of the optimal conditions for the government program in the absence of maximum wage, yields:

(13) $$\varepsilon_3-\varepsilon_2>0.$$

To verify that the condition in (13) holds, notice that the marginal tax rate imposed on the low-skill individuals in the optimal solution for the government program with no maximum wage in place is strictly positive [this is a standard result in the literature; see, e.g., Stiglitz (1982)]. As is standard in the literature, the marginal tax rate is captured by slope of the indifference curve of the low-skill individual going through his chosen bundle. Formally (denoting by MTR_1 the marginal tax rate levied on the low-skill individual)

(14) $$\text{MTR}_1=1-\frac{h'(y_1/w_1)/w_1}{g'(c_1)}>0\leftrightarrow\frac{g'(c_1)}{h'(y_1/w_1)/w_1}>1 .$$

Rewriting (8) and substituting from (14) yields:

(15) $$\frac{g'(c_1)}{h'(y_1/w_1)/w_1}=\varepsilon_3/\varepsilon_2>1,$$

It then follows immediately that $\varepsilon_3-\varepsilon_2>0$, hence (13) is satisfied.

Our final step which will establish our argument is to show that the fiscal surplus can be distributed to both individuals without affecting the two incentive compatibility constraints. To see this consider the following small perturbation to the consumption levels of both types of individuals:

(16) $$\begin{aligned} c_1^{''} &= c_1^{'} + \delta_1 \\ c_2^{''} &= c_2^{'} + \delta_2 \end{aligned},$$

where $c_1^{'}$ and $c_2^{'}$ are defined in (7), $\delta_1, \delta_2 > 0$ are small and satisfy:

(17) $$\delta_1 \cdot g'(c_1^{'}) = \delta_2 \cdot g'(c_2^{'}).$$

Notice that condition (16) defines an increase in the consumption levels of both types of individuals that [by virtue of (17)] induces the same change in the corresponding total levels of utility (using a first-order approximation). By virtue of the *separability* of the utility [see equation (2)] the above perturbation violates none of the incentive compatibility constraints. Setting δ_1 and δ_2 sufficiently small ensures that the revenue constrains is not violated either (in light of the fiscal surplus). This completes the proof. QED

The rationale underlying the proposition is straightforward. In the absence of maximum wage rate, the government sets a strictly positive marginal tax on the low-skill individuals. This downward distortion in the labor-leisure allocation of the low-skill individuals is introduced in order to relax the otherwise binding incentive constraint of the high-skill individuals, thereby attaining enhanced re-distribution (enabling the government to extract additional resources from the high-skill individual). A maximum wage rate makes it harder for the high-skill individual to mimic, as it has to work harder in order to obtain the same level of earnings. Thus, by introducing a binding maximum wage rate, the government creates a slack in the incentive compatibility constraint of the high-skill individual.[7] The government can thus mitigate the downward distortion in the labor-leisure choice of the low-skill individuals without violating the incentive constraint of the high-skill individuals. This would result in a fiscal surplus that can be used to make everybody better-off. Notably, the introduction of the maximum wage rate does not entail any distortion as the high-skill workers obtain the same bundle in the modified allocation as in the original one. Although the maximum wage rate is set at a level lower than the marginal product of the high-skill workers, which, *ceteris paribus*, results in a downward distortion in the labor supply choice of the high-skill workers, the government offsets this distortion by introducing a counter-distortion in the form of a marginal subsidy.

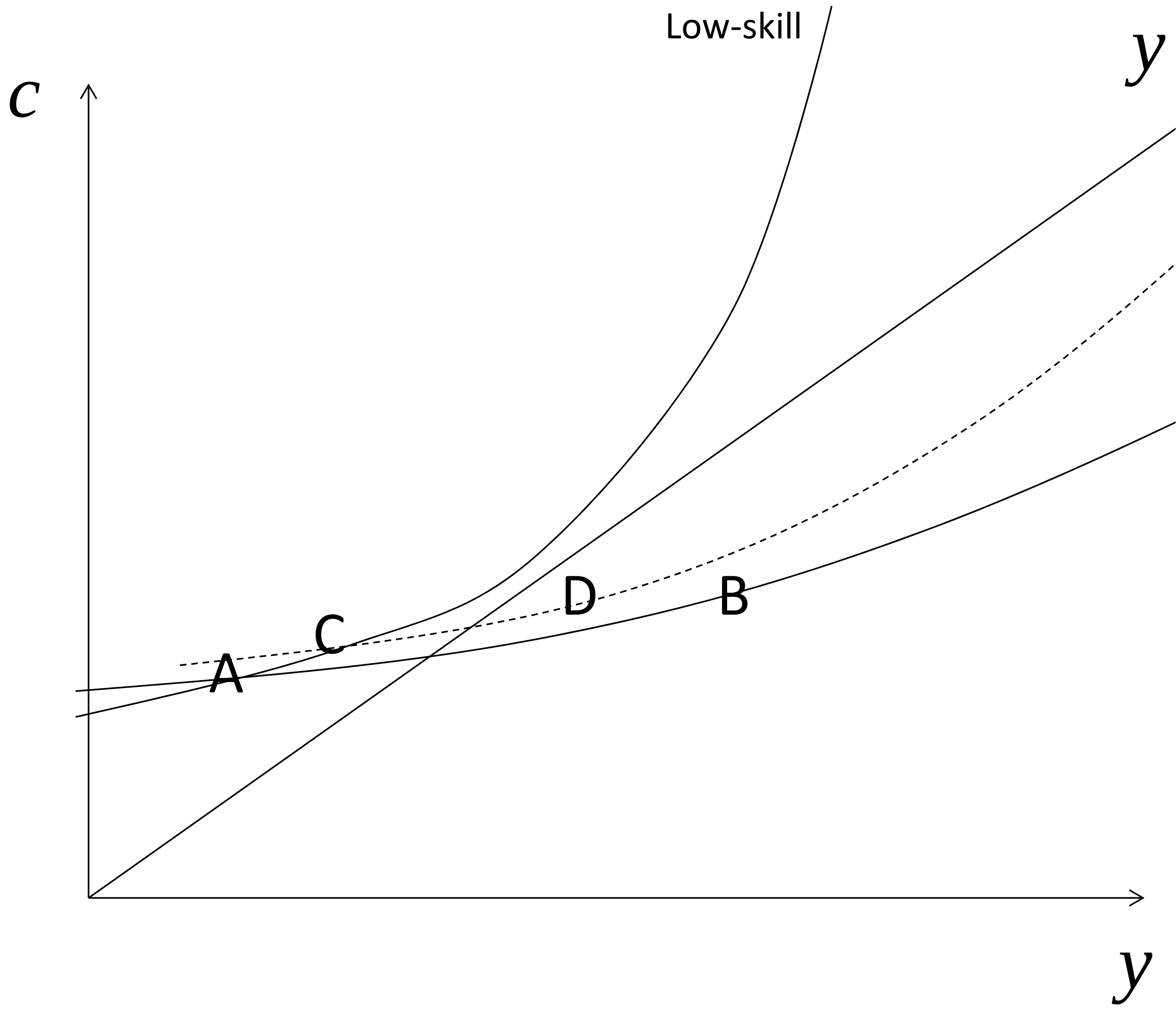
5. Conclusions

In the popular discourse of maximum wage rules, reflecting unease about excessive pay of CEOs, an often raised argument in favor of setting such rules relates to fairness considerations. That is, workers care not only about the absolute remuneration but rather about their relative compensation; hence, goes the line of reasoning, medium- and low paid workers would benefit from knowing that their superiors' take would not exceed their lot substantially. In this paper we abstract from addressing these arguments, without discounting their importance, and rather set focus on the role played by maximum wage rules in shaping the incentive structure of the welfare system and in enhancing its target efficiency. We demonstrate that setting a binding

[7] The same mechanism works in Stiglitz (1982), demonstrating that when the high- and low-skill labor are complements in the production technology, then it is socially desirable to subsidize the high-skill workers at the margin (rather than setting the marginal tax rate to zero) in order to reduce the wage gap across skill levels, thereby making it harder for the high skill to mimic. Subsidizing at the margin induces the high-skill workers to spend more hours in the labor market and results in a reduction in their wage rate (equal to their marginal product which is assumed to be diminishing). As high- and low-skill workers are assumed to be complements in the production technology, this also induces an increase in the wage rate of the low-skill workers (again equal to their marginal product).

maximum wage rule is a socially desirable supplement to the tax-and-transfer system. We show that it serves to mitigate the mimicking incentives of the high-skill workers, thereby allowing the government to reduce the marginal tax rates levied on the low skill workers.

Appendix: Graphical Illustration of the Key Argument



The figure demonstrates how a fiscal surplus is attained by imposing a binding maximum wage rule. The bundles associated with the low-skill and the high-skill individuals under the optimal tax regime prior to introducing the maximum wage rule (the benchmark regime) are depicted in the gross income-net income (y-c) plane and given, respectively, by the points A and B. The (solid) indifference curves are depicted for the high- and low skill individuals for this benchmark case. By the single crossing property the indifference curve of the low-skill agent is steeper than that of the high-skill. Notice that both bundles lie on the (flatter) indifference curve of the high-skill individual, as his incentive constraint is binding.

Imposing a maximum wage rule implies that the high-skill indifference curve shifts in an anticlockwise fashion and is now represented by the dashed line. It intersects with the low-skill indifference curve at a higher point (C). The two new bundles under the perturbed tax regime following the introduction of a maximum wage rule are given by the points C and D (for the low- and high-skill individuals, respectively). Notice that the two bundles still lie on the (dashed) indifference curve of the high-skill (as his incentive constraint still binds). The fiscal surplus is created by the shift along the indifference curve of the low-skill individual (from A to C). To see this, notice that the transfer to the low-skill individual is measured by the vertical difference from his bundle to the 45 degrees line (which represents the *Laissez Faire* regime). As can be observed from the figure, this distance is decreasing as we shift upwards along the indifference curve from A to C. The reason for this is that at point A (the bundle associated with the benchmark tax regime) the slope of the indifference curve is lower than 1 (reflecting a positive marginal tax rate).

References

Allen, S. (1987) "Taxes, Redistribution, and the Minimum Wage: A Theoretical Analysis", Quarterly Journal of Economics, 102, 477-490.

Balcer, Y. and Sadka, E. (1982) "Horizontal Equity, Income Taxation and Self-Selection with an Application to Income Tax Credits," Journal of Public Economics, 19, 291-309.

Blumkin, T., Margalioth, E. and Sadka, E. (2007) "Anti-discrimination Rules versus Income Taxation in the Pursuit of Horizontal Equity", Journal of Public Economics, 91, 1167-1176.

Lee, D. and Saez, E. (2012) "Optimal Minimum Wage Policy in Competitive Labor Markets", Journal of Public Economics, 96, 739-749.

Mirrlees, J. (1971) "An Exploration in the Theory of Optimum Income Taxation", Review of Economic Studies 38, 175–208.

Stiglitz, .J. (1982) "Self-selection and Pareto Efficient Taxation", Journal of Public Economics 17, 213–240.