A Graduated Minimum Wage with Optimal Taxation

Eliav Danziger* and Leif Danziger**

Abstract

In this paper we prove that a graduated minimum wage rate can provide a Pareto improvement of an optimal allocation with nonlinear taxation. The reason is that a graduated minimum wage rate makes it harder for the more productive workers to mimic the income of the less productive workers. We also show that in a utilitarian social welfare optimum, the graduated minimum wage rate increases the consumption of low-productivity workers. However, due to changes in their working hours, they do not necessarily gain in utility. We further establish several comparative-static results for the minimum wage rate in the social optimum.

JEL Classifications: H21, J38
Keywords: Graduated minimum wage, optimal taxation, Pareto improvement

* *Department of Economics, Princeton University, Princeton, NJ 08544, USA*
** *Department of Economics, Ben-Gurion University, Beer-Sheva 84105, Israel*

1 Introduction

In a standard competitive economy where workers have identical preferences but different productivities, Allen (1987) and Guesnerie and Roberts (1987) have shown that a constant minimum wage rate can never raise social welfare above what can be achieved by a nonlinear income tax. In fact, in the presence of optimal taxes, a binding constant minimum wage rate would be counterproductive and decrease social welfare. The reason is that in order to redistribute from the more to the less productive workers, the government must be able to distinguish between different types of workers. Thus, the government generally prefers policies that make it harder for the high-productivity workers to mimic the income of the low-productivity workers. The enactment of a binding constant minimum wage rate, however, has the opposite effect as it reduces the low-productivity workers' income thereby making it easier for high-productivity workers to match the income of low-productivity workers. As a result, with a binding constant minimum wage rate it becomes even harder for the government to distinguish between the different types of workers and social welfare is therefore reduced.[1]

It is therefore surprising that minimum wage policies are so ubiquitous, and subsequent papers have tried to justify the minimum wage rate by introducing unemployment into the competitive model. Thus, if the government's taste for redistribution toward the less productive workers is strong enough, a constant minimum wage rate may be socially beneficial (but not a Pareto improvement) if it forces some of the low-productivity workers to become involuntarily unemployed (Marceau and Boadway, 1994); if it is combined with a welfare policy

[1] In the absence of optimal nonlinear income taxation, a constant minimum wage rate may be socially beneficial in both a competitive environment (Danziger, 2009; Lee and Saez, 2009) and in a monopsonistic environment (Robinson, 1933). Search costs and informational asymmetries may lead to monopsonistic features, and hence rationalize a constant minimum wage rate in other types of labor markets (Mortensen and Pissarides, 1994; Rebitzer and Taylor, 1995; Bhaskar et al., 2002; Flinn, 2006; and Basu et al., 2010). However, with optimal nonlinear income taxation, monopsony cannot justify a constant minimum wage rate (Cahuc and Laroque, 2009).

There exists an extensive empirical literature that examines the consequences of minimum wage rates (for example, Card and Krueger, 1995; Neumark et al. 2004; Neumark and Wascher, 2007; Burkhauser et al., 2000; and Zavodny, 2000).

that obliges unemployed workers to accept job offers even when they prefer unemployment (Boadway and Cuff, 2001); if there are search frictions in the labor market (Hungerbühler and Lehmann, 2009); or if the workers with the highest disutilities of work are laid off first (Lee and Saez, 2009).[2] In this paper we pursue a different path. We suggest keeping the standard competitive model with full employment intact and instead propose a minimum wage design that provides a Pareto improvement of the optimal allocation achievable with nonlinear taxation.

There is an obvious dissonance between insisting on a constant minimum wage rate while at the same time admitting any nonlinear income tax schedule. In the spirit of the optimal nonlinear income tax literature, therefore, this paper analyzes the virtues of a graduated minimum wage rate, that is, a nonlinear minimum wage rate that ties the minimum wage rate a firm must pay to its total minimum-wage labor input. The two-type Stiglitz (1982) style model we develop will show that such a graduated minimum wage rate will break the tight connection between the wage rate and the workers' marginal product, in this way providing the government with a tool for compelling firms to hire low-productivity workers at a minimum wage rate that exceeds their marginal product.[3] By doing so, the government is able to loosen the incentive-compatibility constraint of the more productive workers and thereby increase social welfare.[4]

A major result of the paper is that, starting from an optimal allocation with nonlinear income taxation, the introduction of a graduated minimum wage rate can always enhance

---

[2] A constant minimum wage rate may also be useful by preventing workers from signalling their earning ability (Blumkin and Sadka, 2005).

[3] To make it affordable for firms to pay low-productivity workers above their marginal product, the model assumes that there are decreasing returns to scale and that low-productivity workers have a decreasing marginal product.

[4] We follow the optimal income tax literature in assuming that for legal or technical reasons, taxes can depend only on incomes (earnings), and not directly on wage rates, even though the enactment of a minimum wage policy requires that the enforcement agency can observe workers' wage rates in order to verify that they are paid at least the minimum wage rate (see Guesnerie and Roberts, 1987; Lee and Saez, 2009).

the utility of all individuals, thus constituting a Pareto improvement. The reason is that a graduated minimum wage rate permits the government to raise the low-productivity workers' income from work at the expense of the firms' profits that are distributed to all workers. The high-productivity workers would then have to work more in order to obtain the income of the low-productivity workers, which would make it harder for the high-productivity workers to match the low-productivity workers' income. This, in turn, loosens the incentive-compatibility constraint of the high-productivity workers. Importantly, and in contrast to other justifications of the minimum wage, the logic is independent of the government's taste for redistribution toward the workers with the lowest pay and transcends the specific model assumed in this paper. Therefore, the result that a graduated minimum wage is capable of providing a Pareto improvement even in the presence of an optimal nonlinear income taxation will remain true under very general circumstances.

We show that with a utilitarian social welfare function, an optimally chosen graduated minimum wage rate leads to an increase in the consumption of the low-productivity workers as it decreases the marginal social cost of their consumption. However, the low-productivity workers do not necessarily gain from the graduated minimum wage rate since their working hours could increase so much that their utility decreases. In this case, the gain of the high-productivity workers would more than outweigh the loss of the low-productivity workers.

We also show that if the graduated minimum wage rate is differentiable, then the marginal labor cost is less than the minimum wage rate. This means that, effectively, firms are facing a downward-sloping labor supply of low-productivity workers for which reason they hire more labor than they would at a constant minimum wage rate equal to the marginal product. In other words, by means of the graduated minimum wage rate the government endows the workers with monopoly power and allows them to practice second-degree price discrimination. As a consequence, the firms employ more low-productivity workers than they would in a competitive labor market with the same wage rate.

Finally, we establish several comparative-static results: The minimum wage rate in the social optimum decreases with government expenditure on public goods and with the supply of low-productivity workers, while it increases with the supply of high-productivity workers, with the mass of firms, and with the productivity of the high-productivity workers.

The analysis in this paper emphasizes the normative aspects of a graduated minimum wage rate. Nevertheless, we wish to point out that graduated minimum wage rates are in fact implemented in several countries. For example, Colombia, Honduras, and Panama have multi-bracket minimum wage systems where the minimum wage rate depends on firm size, and within the U.S., thirteen states have two-bracket state minimum wage systems where the minimum wage rate depends on the size of the firm.[5] In addition, in numerous other countries, including Argentina, Greece, Ireland, and Puerto Rico, the minimum wage rate depends on the sector or occupation, both of which could serve as a proxy for the employment of minimum-wage workers.

2 The Economy

We consider an economy with a continuum of low-productivity workers with measure $n_1 > 0$ and a continuum of high-productivity workers with measure $n_2 > 0$. All workers have the same utility function $U(c) - h$, where $c \geq 0$ is the worker's consumption and $h \geq 0$ is his working hours. The utility function satisfies $U'(0) = \infty$, $U' > 0$ and $U'' < 0$.

There is also a continuum of firms with measure $x > 0$. Each firm has a production function $f(\ell_1) + q\ell_2$, where ℓ_1 and ℓ_2 are the total hours of labor input of low- and high-productivity workers, respectively. We assume that $f(0) = 0$, $f' > 0$, and $f'' < 0$. The production function therefore exhibits decreasing returns to scale, with low-productivity workers having a decreasing marginal product and high-productivity workers having a constant mar-

[5] See http://www.dol.gov/whd/minwage/america.htm.

ginal product. The marginal products satisfy $f'(0) < q$, which justifies our classification of workers as having either low or high productivity.

The firms' profits are distributed evenly among the workers, each worker receiving $D \equiv x[f(\ell_1) + q\ell_2 - w_1\ell_1 - w_2\ell_2]/(n_1 + n_2)$, where w_1 and w_2 are the wage rates of low- and high-productivity workers, respectively. Hence, a low-productivity worker's income is $w_1 h_1 + D$, where h_1 is a low-productivity worker's working hours, and a high-productivity worker's income is $w_2 h_2 + D$, where h_2 is a high-productivity worker's working hours.

In addition to the workers and firms, there is a government with a utilitarian social welfare function

$$n_1 \left[U(c_1) - h_1\right] + n_2 \left[U(c_2) - h_2\right], \tag{1}$$

where c_1 and c_2 are the consumption of a low- and a high-productivity worker, respectively. The government has a budget constraint

$$x[f(\ell_1) + q\ell_2] - n_1 c_1 - n_2 c_2 = R, \tag{2}$$

where $R \geq 0$ is an exogenous expenditure on public goods (that do not factor into social welfare). The government observes the workers' incomes, and its objective is to maximize social welfare by choosing an income-tax function for workers and possibly also a graduated minimum wage rate.[6] The income-tax function may be nonlinear, and, as in much of the optimal income tax literature, we assume that income tax cannot be conditioned on the workers' wage rates or their working hours.

For comparison with our later results, we note that the quasi-linear utility function and the lower marginal product of the low-productivity workers imply that if the government could fully control the consumption and working hours of all workers, then c_1 and c_2 should be identical and determined such that the marginal utility of a worker's consumption is equal

[6] Since profits are distributed evenly, a profit tax would be equivalent to a lump-sum tax on workers. Accordingly, allowing the government to also tax profits would not affect the results.

to the lowest possible marginal social cost of producing that consumption, $1/q$. Furthermore, the low-productivity workers should not work. Accordingly, in a first-best allocation $U'(c_1) = U'(c_2) = 1/q$, $h_1 = 0$, and $h_2 = (n_1/n_2 + 1)\,(U')^{-1}\,(1/q)$.

3 Optimal Taxes

Without government intervention in the wage setting, the labor market for low- and high-productivity workers clear,

$$\ell_1 = \frac{n_1 h_1}{x}, \tag{3}$$

$$\ell_2 = \frac{n_2 h_2}{x}, \tag{4}$$

and the wage rates are competitively determined so that the workers are paid their marginal products,

$$w_1 = f'\left(\frac{n_1 h_1}{x}\right), \tag{5}$$

$$w_2 = q. \tag{6}$$

Since $f' < q$, it follows that $w_1 < w_2$.

Any income-tax function is equivalent to the government offering the workers a choice between different consumption-income bundles. However, since low-productivity workers will need to work more than high-productivity workers in order to reach the same income level, a given consumption-income bundle provides the two types of workers with different utilities. Hence, the low- and high-productivity workers will not in general choose the same bundle. This gives the government the ability to differentiate between the two types of workers by designing the income-tax schedule so that each type of worker prefers the consumption-income bundle meant for his own type rather than for the other type. Thus, the income-tax schedule must satisfy the two incentive-compatibility constraints that low-productivity workers cannot make themselves better off by earning the same income (and consequently obtaining the same consumption) as high-productivity workers, and that high-productivity workers cannot make themselves better off by earning the same income (and consequently obtaining the same consumption) as low-productivity workers. Appendix A shows that at a social optimum the incentive-compatibility constraint of the low-productivity workers does not bind while the incentive-compatibility constraint of the high-productivity workers does bind, i.e.,

$$U(c_2) - h_2 = U(c_1) - \frac{w_1 h_1}{w_2}, \tag{7}$$

where $w_1 h_1 / w_2$ is the number of hours that a high-productivity worker would have to work in order for his income to reach that of a low-productivity worker.

By substitution of the market-clearing conditions (3)-(4), the government's budget constraint (2) becomes

$$xf\left(\frac{n_1 h_1}{x}\right) + qn_2 h_2 - n_1 c_1 - n_2 c_2 = R, \tag{8}$$

and by substitution of the competitively determined wage rates (5)-(6), the incentive-compatibility constraint of the high-productivity workers (7) becomes

$$U(c_2) - h_2 = U(c_1) - \frac{f'(n_1 h_1 / x)\, h_1}{q}. \tag{9}$$

The government's problem is then to determine a tax function so as to maximize social welfare (1) subject to its budget constraint (8) and the incentive-compatibility constraint of the high-productivity workers (9). However, we need only be concerned with the two consumption-income bundles intended for low- and high-productivity workers since the government can use the optimal income tax system to make any other bundle less attractive. Furthermore, since choosing the consumption-work bundles for the low- and high-productivity workers is equivalent to choosing their consumption-income bundles, we can restate the government's problem as how to determine c_1, h_1, c_2, and h_2 in order to maximize (1) subject to (8) and (9).

For later reference, we note that a social welfare optimum (assuming an internal solution) satisfies

$$n_1 U'(\hat{c}_1) - \lambda n_1 - \mu U'(\hat{c}_1) = 0,$$

$$-n_1 + \lambda n_1 f'\left(\frac{n_1 \hat{h}_1}{x}\right) + \frac{\mu}{q}\left[f''\left(\frac{n_1 \hat{h}_1}{x}\right)\frac{n_1 \hat{h}_1}{x} + f'\left(\frac{n_1 \hat{h}_1}{x}\right)\right] = 0,$$

$$n_2 U'(\hat{c}_2) - \lambda n_2 + \mu U'(\hat{c}_2) = 0,$$

$$-n_2 + \lambda q n_2 - \mu = 0,$$

where a circumflex over a variable denotes the optimal value of the variable, $\lambda > 0$ is the multiplier associated with the budget constraint, and $\mu > 0$ is the multiplier associated with the incentive-compatibility constraint of the high-productivity workers. By eliminating the multipliers, we obtain that the marginal utilities of the low- and high-productivity workers' consumption satisfy

$$U'(\hat{c}_1) = \frac{n_1 q + n_2 f'(n_1 \hat{h}_1/x) + n_2 (n_1 \hat{h}_1/x) f''(n_1 \hat{h}_1/x)}{q\left[(n_1 + 2n_2) f'(n_1 \hat{h}_1/x) - n_2 q + n_2 (n_1 \hat{h}_1/x) f''(n_1 \hat{h}_1/x)\right]}, \tag{10}$$

$$U'(\hat{c}_2) = \frac{1}{q}, \tag{11}$$

where the second-order conditions for a maximum imply that the denominator on the right-hand-side of (10) is strictly positive.

Conditions (10) and (11) show that for each type of worker the marginal utility of consumption is equal to the marginal social cost of producing the consumption. Thus, the consumption of the high-productivity workers is at its first-best level. At the same time, the binding incentive-compatibility constraint of the high-productivity workers implies that the marginal social cost of producing $\hat{c}_1$ (the right-hand-side of (10)) exceeds $1/f'\left(n_1 \hat{h}_1/x\right)$, which is a low-productivity worker's marginal disutility from increasing his production.[7] The explanation is that an increase in the consumption of the low-productivity workers tightens the incentive-compatibility constraint of the high-productivity workers by making the consumption-income bundle of the low-productivity workers more attractive to the high-productivity workers, and to counteract this tightening of the constraint, the working hours of the high-productivity workers must decrease. Therefore, to increase the consumption of the low-productivity workers, their labor input must increase in order to both produce all their additional consumption and make up for the decrease in the labor input of the high-productivity workers. It is this need to substitute the less efficient low-productivity workers for the more efficient high-productivity workers that causes the marginal social cost of a low-productivity worker's consumption to exceed $1/f'\left(n_1\hat{h}_1/x\right)$. As a consequence of the high marginal social cost of $\hat{c}_1$, the low-productivity workers will consume less than the first-best level that is obtained by the high-productivity workers, i.e., $\hat{c}_1 < \hat{c}_2$. The low-productivity workers will also work more, while the high-productivity workers will work less than in the first-best allocation.[8]

Also for later reference, we now derive the marginal tax rates in the social optimum. Following e.g. Stiglitz (1982) and assuming differentiability of the income tax schedule, we differentiate a low-productivity worker's utility with respect to h_1 to obtain that his marginal

[7] It follows from $f'(n_1\hat{h}_1/x) < q$ and $f''(n_1\hat{h}_1/x) < 0$ that

$$\begin{aligned} & \left[f'(n_1\hat{h}_1/x) - q\right]^2 + (n_1\hat{h}_1/x)\left[f'(n_1\hat{h}_1/x) - q\right] f''(n_1\hat{h}_1/x) > 0 \\ \Leftrightarrow \quad & \frac{n_1 q + n_2 f'(n_1\hat{h}_1/x) + n_2(n_1\hat{h}_1/x)f''(n_1\hat{h}_1/x)}{q\left[(n_1+2n_2)f'(n_1\hat{h}_1/x) - n_2 q + n_2(n_1\hat{h}_1/x)f''(n_1\hat{h}_1/x)\right]} > \frac{1}{f'(n_1\hat{h}_1/x)}. \end{aligned}$$

[8] A sufficient condition for $\hat{h}_1 > 0$ is that the denominator on the right-hand-side of (10) is positive for $n_1 = 0$, which requires that $f'(0) > n_2 q/(n_1+2n_2)$.

tax rate in the social optimum, denoted by $\hat{t}_1$, is given by

$$\begin{aligned} & w_1(1-\hat{t}_1)U'(\hat{c}_1) - 1 = 0 \\ \Leftrightarrow \quad & \hat{t}_1 = 1 - \frac{1}{U'(\hat{c}_1)f'\left(n_1\hat{h}_1/x\right)}. \end{aligned}$$

Similarly, differentiating a high-productivity worker's utility with respect to h_2, we obtain that his marginal tax rate in the social optimum, denoted by $\hat{t}_2$, is given by

$$\begin{aligned} & q(1-\hat{t}_2)U'(\hat{c}_2) - 1 = 0 \\ \Leftrightarrow \quad & \hat{t}_2 = 0. \end{aligned}$$

As with a constant-returns-to-scale technology (Stiglitz, 1982), in our framework with decreasing returns to scale the optimal marginal tax rate for the low-productivity workers is positive and the optimal marginal tax rate for the high-productivity workers is zero.

4 A Graduated Minimum Wage Rate

The only way that government interference in the wage setting may lead to increased social welfare would be if it were to loosen the incentive-compatibility constraint of the high-productivity workers. It would then be less attractive for the high-productivity workers to choose the consumption-income bundle intended for the low-productivity workers. This would give the government more leeway in increasing the consumption of the low-productivity workers and possibly also decreasing their working hours. As we will show, a loosening of the incentive-compatibility constraint of the high-productivity workers can be obtained by allowing the government to not only set income taxes, but also to introduce a graduated minimum wage rate. To this end, we now let the government enact a minimum wage schedule that determines how the minimum wage rate a low-productivity firm must pay will depend on the total working hours of the workers it employs at the minimum wage rate. In other words, in the same way as the government offers the workers a choice between different consumption-work bundles, it also offers the firms a choice between different minimum wage-hours bundles. Accordingly, the wage rate for low-productivity workers is no longer competitively determined and may exceed their marginal product. In contrast, the wage rate for the high-productivity workers continues to be competitively determined and is equal to q.

Just as with the optimal tax system where we need only be concerned with the consumption-income bundles intended for the two types of workers, with the optimal graduated minimum wage rate we need only consider the minimum wage-hours bundle for the low-productivity workers that the government intends the firms to choose.[9] Observe that since firms can always choose not to hire low-productivity workers at all, the profit from employing low-productivity workers at the intended minimum wage-hours bundle must be nonnegative. Therefore, letting m denote the minimum wage rate, the graduated minimum wage rate generates the constraint that

$$f(\ell_1) - m\ell_1 \geq 0.$$

Since the government will set the income taxes and the graduated minimum wage rate so that the labor market for low-productivity workers clears, (3)-(4) continue to hold.[10] The high-productivity workers are still paid their marginal product, so also (6) continues to hold. Likewise, the government's budget constraint is again given by (8). The incentive-

[9] The government can set a prohibitively high minimum wage rate, for example one equal to q, in all minimum wage-hours bundles that have different working hours. It can be shown that, since low-productivity workers cannot be forced to take employment, the concavity of their utility function implies that it would not be socially optimal to provide different low-productivity workers with different consumption-work bundles. In particular, it would not be socially optimal that some low-productivity workers be unemployed.

[10] If the minimum wage constraint is not binding, then excess demand for low-productivity workers, $\ell_1 > n_1 h_1 / x$, would be incompatible with competition among firms, while if the minimum wage constraint is binding, then by definition there would be no excess demand for low-productivity workers. In both cases, excess supply of low-productivity workers, $\ell_1 < n_1 h_1 / x$, would lead to rationing of low-productivity workers which would be incompatible with the maximization of social welfare.

compatibility constraint of the high-productivity workers is modified to[11]

$$U(c_2) - h_2 = U(c_1) - \frac{mh_1}{q}, \tag{12}$$

while the minimum wage constraint can be written as

$$f\left(\frac{n_1 h_1}{x}\right) - \frac{mn_1 h_1}{x} \geq 0. \tag{13}$$

With a graduated minimum wage rate, then, the government's optimization problem can be formulated as how to determine (c_1, h_1, m) and (c_2, h_2) so as to maximize social welfare (1). The solution has to satisfy the budget constraint (8), the modified incentive-compatibility constraint of the high-productivity workers (12), and the minimum-wage constraint (13).

We now show that, starting from an optimal allocation with nonlinear income taxes, a graduated minimum wage rate can always be used to achieve a Pareto improvement. Observe that $f''\left(n_1\hat{h}_1/x\right) < 0$ implies that the minimum wage constraint is not binding at an optimal allocation with nonlinear taxes. Without violating the minimum wage constraint, therefore, the minimum wage rate can be set slightly above the wage rate of the optimal allocation with nonlinear taxes. At unchanged consumption and work levels, this will loosen the incentive-compatibility constraint of the high-productivity workers as they will have to work more in order to earn the income of the low-productivity workers. Accordingly, starting from an optimal allocation with nonlinear taxation, without violating any constraint the government can slightly increase the minimum wage rate, the consumption, and the working hours of the low-productivity workers, with the increase in their consumption being equal to the increase in their production. Since the marginal utility of a low-productivity worker's consumption exceeds his marginal disutility from an increase in production, i.e., $U'(\hat{c}_1) > 1/f'\left(n_1\hat{h}_1/x\right)$ (see condition (10)), the low-productivity workers' utility will increase. The government can furthermore transfer a little of the low-productivity workers' consumption

[11] As in the case without a graduated minimum wage rate, (only) the incentive-compatibility constraint for the high-productivity workers binds.

to the high-productivity workers without violating the incentive-compatibility constraint of the low-productivity workers and the minimum wage constraint, thereby increasing the high-productivity workers' utility, while still making sure that the low-productivity workers' utility increases. We have therefore proved:

Proposition 1: *A graduated minimum wage rate can provide a Pareto improvement of an optimal allocation with nonlinear income taxation.*

A Pareto improvement is possible because the graduated minimum wage rate is more efficacious than income taxation (and profit taxation) in directing firms' profits directly to the low-productivity workers. The rationale is that the graduated minimum wage rate serves as a mechanism that funnels more of the firms' profits to the low-productivity workers, thereby raising their pretax income at the expense of the high-productivity workers' pretax income. This makes it easier for the government to separate between the low- and high-productivity workers, which mitigates the incentive-compatibility constraint of the high-productivity workers and hence facilitates the Pareto improvement.

The driving force underlying the Pareto improvement is solely the capability of the graduated minimum wage rate to mitigate the incentive-compatibility constraint of the high-productivity workers. In particular, the ability of the government to provide a Pareto improvement is independent of its taste for redistribution as embodied in the utilitarian social welfare function. Proposition 1 will therefore be valid beyond the specific assumptions of the model and hold also in a much more general setting as long as there are decreasing returns to scale and the low-productivity workers have a decreasing marginal product.

The graduation of the minimum wage rate is crucial, as a constant minimum wage rate cannot be used to achieve a Pareto improvement. This is because, on the one hand, a constant minimum wage rate that does not exceed the competitive wage rate is ineffectual, while, on the other hand, a constant minimum wage rate that exceeds the competitive wage rate will reduce the low-productivity workers' working hours. However, the latter will necessarily be counterproductive as it tightens (rather than loosens) the modified incentive-compatibility constraint of the high-productivity workers (Allen, 1987; Guesnerie and Roberts, 1987).[12]

Proposition 1 immediately implies that an optimal graduated minimum wage rate will increase social welfare. Furthermore, the logic underlying the proof of Proposition 1 indicates that the minimum wage rate will be raised as long as the minimum wage constraint is not binding. Consequently, in a social optimum the firms' profits are driven to zero so that the minimum wage constraint will bind. Hence, (13) becomes an equality constraint,

$$f\left(\frac{n_1 h_1}{x}\right) - \frac{m n_1 h_1}{x} = 0. \tag{14}$$

That is, in a social optimum the minimum wage rate equals the average product of the low-productivity workers (and hence exceeds their marginal product).

5 Social Welfare Optimum with a Graduated Minimum Wage Rate

In a social welfare optimum, the consumption and working hours of the low- and high-productivity workers and the minimum wage rate will satisfy (assuming an internal solution)

$$n_1 U'(c_1^*) - \lambda_m n_1 - \mu_m U'(c_1^*) = 0,$$
$$-n_1 + \lambda_m n_1 f'\left(\frac{n_1 h_1^*}{x}\right) + \frac{\mu_m m^*}{q} + \frac{\eta n_1}{x}\left[f'\left(\frac{n_1 h_1^*}{x}\right) - m^*\right] = 0,$$
$$n_2 U'(c_2^*) - \lambda_m n_2 + \mu_m U'(c_2^*) = 0,$$
$$-n_2 + \lambda_m q n_2 - \mu_m = 0,$$

[12] Formally, the inability of a constant minimum wage rate to bring about a Pareto improvement follows from the fact that if workers are paid their marginal product, then setting a constant minimum wage rate is equivalent to setting the working hours for the low-productivity workers. Therefore, the possibility of setting a constant minimum wage rate does not provide the government with an extra instrument to affect the feasible consumption-work bundles.

$$\frac{\mu_m}{q} - \frac{\eta n_1}{x} = 0,$$

where an asterisk denotes the optimal value of a variable, $\lambda_m > 0$ is the multiplier of the budget constraint (8), $\mu_m > 0$ is the multiplier of the modified incentive-compatibility constraint of the high-productivity workers (12), and $\eta > 0$ is the multiplier of the minimum wage constraint (14).

By eliminating the multipliers, we obtain that the marginal utilities of consumption satisfy

$$U'(c_1^*) = \frac{n_1 q + n_2 f'(n_1 h_1^*/x)}{qA}, \tag{15}$$

$$U'(c_2^*) = \frac{1}{q}, \tag{16}$$

where

$$A \equiv (n_1 + 2n_2) f'(n_1 h_1^*/x) - n_2 q,$$

and the second-order conditions for a maximum imply that $A > 0$.

As shown by conditions (16) and (11), the consumption of the high-productivity workers is not affected by the graduated minimum wage rate and remains at the first-best level. This reflects the fact that the marginal social cost of their consumption is unchanged. However, comparing condition (15) with condition (10) reveals that the marginal social cost of the low-productivity workers' consumption, and consequently their consumption, is changed. Specifically, we have:[13]

Proposition 2: $\hat{c}_1 < c_1^* < \hat{c}_2 = c_2^*$.

The graduated minimum wage rate loosens the incentive-compatibility constraint of the high-productivity workers, thereby reducing the marginal social cost of the low-productivity

[13] The proofs of the propositions are in Appendix B.

workers' consumption (the right-hand-side of condition (15)). As a consequence, the introduction of a graduated minimum wage rate raises the consumption of the low-productivity workers.

Since the modified incentive-compatibility constraint of the high-productivity workers remains binding, the graduated minimum wage rate does not raise the consumption of the low-productivity workers all the way to the first best. Hence, $c_1^* < c_2^*$ even with a graduated minimum wage rate. Furthermore, the low-productivity workers work more than in the first-best allocation.[14] Indeed, on the margin, an increase in the consumption of the low-productivity workers still requires an increase in their labor input in order to produce the additional consumption as well as to allow for the decrease in the labor input of the high-productivity workers that assures that the modified incentive-compatibility constraint of the high-productivity workers is satisfied. Accordingly, $U'(c_1^*)$ exceeds $1/f'(n_1 h_1^*/x)$ in a social optimum with a graduated minimum wage rate,[15] signifying that it is still true that the marginal social cost of a low-productivity worker's consumption exceeds his marginal disutility from increasing his production.

While the graduated minimum wage rate increases a low-productivity worker's consumption to bring it closer to the first best, it need not reduce a low-productivity worker's hours to bring them closer to the first best. The intuition is that, on the one hand, the introduction of a graduated minimum wage rate increases the low-productivity workers' income with unchanged working hours. This makes it less attractive for the high-productivity workers to mimic their income, which would tend to decrease the working hours of the low-productivity workers that are necessary to satisfy the modified incentive-compatibility constraint of the

[14] A sufficient condition for $h_1^* > 0$ is that $A > 0$ for $n_1 = 0$, i.e., that $f'(0) > n_2 q/(n_1 + 2n_2)$ (the same as for $\hat{h}_1 > 0$).

[15] Formally, the marginal social cost of a low-productivity worker's consumption increases with his working hours (see the proof of Proposition 2), and since the right-hand-side of condition (15) would equal $1/f'(n_1 h_1^*/x)$ if $f'(n_1 h_1^*/x) = q$, it follows that $U'(c_1^*) > 1/f'(n_1 h_1^*/x)$.

high-productivity workers (12). On the other hand, for a given level of working hours, the graduated minimum wage rate also boosts the gain of the low-productivity workers' income associated with an increase in their working hours. Thus, an increase in the working hours of the low-productivity workers loosens the incentive-compatibility constraint of the high-productivity workers more in the presence of a graduated minimum wage rate than without, and this provides a counteracting influence which would tend to increase the working hours of the low-productivity workers. Hence, the graduated minimum wage has both a negative and a positive effect on the working hours, and whereas it is possible that $h_1^* < \hat{h}_1$ and hence $h_2^* > \hat{h}_2$ (because total consumption increases), it is also possible that $h_1^* > \hat{h}_1$. Indeed, the working hours of the low-productivity workers may increase so much that it allows not only for the production of the additional consumption, but also for a reduction in the working hours of the high-productivity workers, i.e., $h_2^* < \hat{h}_2$. Consequently, the effects of the graduated minimum wage rate on the working hours of both types of workers are ambiguous.

Relatedly, although the introduction of the graduated minimum wage rate increases social welfare (and can provide a Pareto improvement), it does not necessarily benefit the low-productivity workers at the expense of the high-productivity workers. Indeed, it may be socially preferable to raise the working hours of the low-productivity workers so much that their utility decreases while the utility of the high-productivity workers increases. So the graduated minimum wage rate also has ambiguous effects on the utilities of the two types of workers.

The following three examples illustrate these ambiguities. All the examples assume that $n_1 = n_2 = x = q = 1$ and $U(c) = -1/c$, but have different production functions.

Example 1: Let $f(\ell_1) = \ell_1 - \ell_1^2/10$, where $\ell_1 < 5$. With income taxes and no minimum wage rate, the social optimum has $\hat{c}_1 = 0.865$, $\hat{h}_1 = 1.006$, $\hat{c}_2 = 1$, and $\hat{h}_2 = 0.960$. A low-productivity worker's utility is $-1/0.865 - 1.006 = -2.162$, and a high-productivity worker's utility is $-1 - 0.960 = -1.960$. With income taxes and a graduated minimum wage rate, the social optimum has $c_1^* = 0.886$, $h_1^* = 0.973$, $c_2^* = 1$, and $h_2^* = 1.007$. A low-productivity worker's utility is $-1/0.886 - 0.973 = -2.102$, and a high-productivity worker's utility is $-1 - 1.007 = -2.007$. In this example, the working hours of the low-productivity workers decrease while the working hours of the high-productivity workers increases. The low-productivity workers benefit from both the higher consumption and the lower working hours, so their utility increases. Since the high-productivity workers have increased working hours, their utility decreases. Thus, although there is an overall gain of social welfare from the graduated minimum wage rate, the additional redistribution of resources to the low-productivity workers causes a welfare loss for the high-productivity workers.

Example 2: Let $f(\ell_1) = \frac{1}{2}\ell_1/(1+\ell_1)$. With income taxes and no minimum wage rate, the social optimum has $\hat{c}_1 = 0.424$, $\hat{h}_1 = 0.072$, $\hat{c}_2 = 1$, and $\hat{h}_2 = 1.390$. A low-productivity worker's utility is $-1/0.424 - 0.072 = -2.431$, and a high-productivity worker's utility is $-1 - 1.390 = -2.390$. With income taxes and a graduated minimum wage rate, the social optimum has $c_1^* = 0.427$, $h_1^* = 0.092$, $c_2^* = 1$, and $h_2^* = 1.385$. A low-productivity worker's utility is $-1/0.427 - 0.092 = -2.434$, and a high-productivity worker's utility is $-1 - 1.385 = -2.385$. In this example, the working hours of the low-productivity workers increase while the working hours of the high-productivity workers decrease. The low-productivity workers benefit from the higher consumption, but lose so much from the higher working hours that their utility decreases. Since the high-productivity workers work less hours, their utility increases. Thus, the effects are opposite those in Example 1 and illustrate that the graduated minimum wage rate may cause so much redistribution of resources from the low- to the high-productivity workers that the former lose and the latter gain from the graduated minimum wage rate.

Example 3: Let $f(\ell_1) = \ell_1 - \ell_1^2$, where $\ell_1 < \frac{1}{2}$. With income taxes and no minimum wage rate, the social optimum has $\hat{c}_1 = 0.462$, $\hat{h}_1 = 0.220$, $\hat{c}_2 = 1$, and $\hat{h}_2 = 1.290$. A low-productivity worker's utility is $-1/0.462-0.220 = -2.387$, and a high-productivity worker's utility is $-1 - 1.290 = -2.290$. With income taxes and a graduated minimum wage rate, the social optimum has $c_1^* = 0.481$, $h_1^* = 0.278$, $c_2^* = 1$, and $h_2^* = 1.280$. A low-productivity worker's utility is $-1/0.481 - 0.278 = -2.357$, and a high-productivity worker's utility is $-1 - 1.280 = -2.280$. Accordingly, the effects on the working hours are the same as in the previous example. However, the utility of both the low- and high-productivity workers increases so that the graduated minimum wage rate constitutes a Pareto improvement.

To sum up, the three examples show that the opposing effects on the working hours of the low-productivity workers indeed make the direction of the change in their working hours ambiguous. The examples also show that the gain from the graduated minimum wage rate may accrue to the low-productivity workers only, to the high-productivity workers only, or to both groups of workers.

6 A Differentiable Graduated Minimum Wage Schedule

Up to now, in our formulation the government only needed to be concerned with firms deviating from the intended minimum wage-hours bundle by declining to hire any low-productivity workers. This is because it was assumed that all other minimum wage-hours bundles entailed a prohibitively high minimum wage rate. Thus, the socially optimal minimum wage schedule that we have determined is not differentiable. To shed further light on the nature of the optimal schedule and its impact on the marginal tax rates, we note that the same social optimum can be implemented also with a differentiable graduated minimum wage schedule. To see this, assume now that the socially optimal minimum wage schedule is a differentiable function of the working hours of the low-productivity workers, denoted by $m(\ell_1)$. A firm would then determine the working hours such that the marginal labor cost,

$d[m(\ell_1)\ell_1]/d\ell_1 = m(\ell_1) + m'(\ell_1)\ell_1$, equals the marginal product. Accordingly, in a social optimum with a graduated minimum wage rate we have that

$$m\left(n_1 h_1^*/x\right) + m'\left(n_1 h_1^*/x\right)\left(n_1 h_1^*/x\right) = f'\left(n_1 h_1^*/x\right).$$

Since $m\left(n_1 h_1^*/x\right) > f'\left(n_1 h_1^*/x\right)$, the marginal labor cost is less than the minimum wage rate. Therefore, the marginal minimum wage rate, $m'\left(n_1 h_1^*/x\right)$, is negative in a social optimum. Thus, we have proved:

Proposition 3: $m'\left(n_1 h_1^*/x\right) < 0$.

That is, a socially optimal differentiable minimum wage schedule is downward sloping at the intended minimum wage rate. Essentially, the graduated minimum wage rate transforms the low-productivity workers into second-degree price discriminating monopolists that confront the firms with a downward-sloping labor supply. The result is that the low-productivity workers extract all of the firms' profits and that the firms hire more labor than they would if the minimum wage rate were constant at its level in the social optimum.

Assuming differentiability of the tax schedule, the marginal tax rate of a low-productivity worker in a social optimum, denoted by t_1^*, is given by

$$\begin{aligned} & w_1(1-t_1^*)\left[m(n_1 h_1^*/x) + m'(n_1 h_1^*/x) n_1 h_1^*/x\right] U'(c_1^*) - 1 = 0 \\ \Leftrightarrow\quad & w_1(1-t_1^*) f'(n_1 h_1^*/x) U'(c_1^*) - 1 = 0 \\ \Leftrightarrow\quad & t_1^* = 1 - \frac{1}{U'(c_1^*) f'\left(n_1 h_1^*/x\right)}, \end{aligned}$$

while the marginal tax rate of a high-productivity worker in a social optimum, denoted by t_2^*, is given by

$$\begin{aligned} & q(1-t_2^*)U'(c_2^*) - 1 = 0 \\ \Leftrightarrow\quad & t_2^* = 0. \end{aligned}$$

Comparing the marginal tax rates with and without a graduated minimum wage rate, we have:

Proposition 4: $0 = t_2^* = \hat{t}_2 < t_1^* < \hat{t}_1$.

The optimal marginal tax rate of low-productivity workers is positive, but smaller with a graduated minimum wage rate than without. That $0 < t_1^*$ reflects that the marginal social cost of a low-productivity worker's consumption exceeds his marginal disutility from an increase in his production, $1/f'(n_1 h_1^*/x)$, due to the modified incentive-compatibility constraint of the high-productivity workers. The reason for $t_1^* < \hat{t}_1$ is that the incentive-compatibility constraint of the high-productivity workers is less inhibitive with a graduated minimum wage rate than without, which reduces the marginal social cost of a low-productivity worker's consumption. The lower marginal social cost enhances the gain from increasing the work of the low-productivity workers, and their marginal tax rate is therefore decreased to encourage an increase in their work.

The optimal marginal tax rate of high-productivity workers is zero, which is the same as without a graduated minimum wage rate. The explanation is that the high-productivity workers have a constant marginal product so that the marginal social cost of a high-productivity worker's consumption is unchanged and equal to his marginal disutility from an increase in his production. As the marginal social cost of their consumption is the lowest possible, the high-productivity workers obtain their first-best consumption level in both cases. Consequently, in neither case should their work incentives be distorted.[16]

[16] If it were the case that $t_2^* > 0$ ($t_2^* < 0$), then high-productivity workers could obtain the same utility by working more (less) and having their consumption increased (lowered) by the full amount that output is increased (reduced). As this would increase government revenue and hence provide the means to increase the consumption of both types of workers without increasing their work, it follows that $t_2^* = 0$. See Stiglitz (1982).

7 Comparative Statics

We now determine the effect of the different parameters of the model on the minimum wage rate in a social optimum.

Proposition 5: $dm^*/dR < 0$.

The higher the government expenditure on public goods, the less output is available for the consumption of low-productivity workers for given working hours of both types of labor. As the marginal utility of low-productivity workers' consumption is inversely related to their consumption, the social gain from additional work increases. Accordingly, the working hours of the low-productivity workers (as well as of the high-productivity workers) increase, thereby lowering the average product of the low-productivity workers and hence the minimum wage rate.

Proposition 6: $dm^*/dn_1 < 0$.

An increase in the supply of low-productivity workers entails that more of these workers will be employed in each firm. As a result, the total working hours of the low-productivity workers increases within each firm. This reduces the average product of the low-productivity workers and hence the minimum wage rate.

Proposition 7: $dm^*/dn_2 > 0$.

High-productivity workers consume less than they produce, so an increase in the supply of these workers facilitates a transfer of resources to the low-productivity workers. This reduces the social gain from the work of the low-productivity workers. Hence, the low-productivity workers work less, which increases their average product and the minimum wage rate.

Proposition 8: $dm^*/dx > 0$.

The more firms there are, the more thinly the low-productivity workers will be spread over firms, which results in a reduction of the total working hours of the low-productivity workers in each firm. This increases the average product of the low-productivity workers and hence also the minimum wage rate.

Proposition 9: $dm^*/dq > 0$.

The higher the productivity of the high-productivity workers, the more resources can be transferred from high- to low-productivity workers, which reduces the social gain from the work of the low-productivity workers. As in the case of an increase in the supply of high-productivity workers, an increase in the productivity of the high-productivity workers leads to the low-productivity workers working less, which increases their average product and hence the minimum wage rate.

8 Conclusion

In this paper we present a normative analysis of a graduated minimum wage rate in a model with two types of workers and optimal income taxation. We have proved that a graduated minimum wage rate can provide a Pareto improvement of an optimal allocation with nonlinear taxation. The explanation is that a graduated minimum wage rate can simultaneously increase the income of the low-productivity workers and reduce the incentive of the high-ability workers to mimic the income of the low-productivity workers. Thus, in sharp contrast to the existing minimum wage literature, we show that a minimum wage policy can be justified without appealing to specific welfare functions or to the government's desire for redistribution.

Effectively, the graduated minimum wage rate allows the low-productivity workers to practice second-degree price discrimination, which explains why it is possible to increase both the wage rate and the income of these workers. This, in turn, weakens the incentive of the high-productivity workers to mimic the income of the low-productivity workers and thereby decreases the marginal social cost of the low-productivity workers' consumption. The latter is also reflected in our finding that the introduction of a graduated minimum wage leads to a reduction in the marginal tax rate of the low-productivity workers.

Among our results for a utilitarian social welfare function is that the introduction of a graduated minimum wage rate leads to an increase in the consumption of the low-productivity workers, thereby bringing it closer to the first best. However, it does not necessarily lead to a reduction in their work to bring it closer to the first best. It is possible that their work increases so much that their utility is reduced, with the low-productivity workers' loss of utility being more than outweighed by the high-productivity workers' gain of utility. We also describe how the minimum wage rate in the social optimum depends on, among other things, the supply of the low- and high-productivity workers, the mass of firms, and productivity.

The model in this paper is a first step in formulating the theoretical underpinnings of a nonlinear minimum wage rate and analyzing the potential benefits of such a policy. Further research should expand the basic model to account for firm heterogeneity, which would require the government to design the minimum wage schedule to distinguish between different types of firms similarly to how the income tax schedule distinguishes between different types of workers. We wish to stress, however, that our finding that a graduated minimum wage rate can provide a Pareto improvement is independent of many of the specific assumptions in our model and therefore holds quite generally.

Appendix A

Workers' Incentive-Compatibility Constraints

The incentive-compatibility constraints of the low- and high-productivity workers are

$$\begin{aligned} U(c_1) - h_1 &\geq U(c_2) - \frac{w_2 h_2}{w_1}, \\ U(c_2) - h_2 &\geq U(c_1) - \frac{w_1 h_1}{w_2}. \end{aligned}$$

By adding the constraints, we obtain that

$$\frac{w_1 h_1}{w_2} - h_2 \geq U(c_1) - U(c_2) \geq h_1 - \frac{w_2 h_2}{w_1}$$
$$\Leftrightarrow w_1 (w_1 h_1 - w_2 h_2) \geq w_1 w_2 [U(c_1) - U(c_2)] \geq w_2 (w_1 h_1 - w_2 h_2).$$

If $c_1 > c_2$, these inequalities would not be possible as $w_1 < w_2$. If $c_1 = c_2$, both incentive-compatibility constraints would have to be binding and $w_1 h_1 = w_2 h_2$. If also $c_1 = c_2 < (U')^{-1}(1/q)$, the government could then increase c_1, c_2, and h_2 by small amounts and keep h_1 unchanged such that $U(c_1)$ and $U(c_2) - h_2$ increase to the same extent. This will increase social welfare, and since it loosens the incentive-compatibility constraint of the low-productivity workers (because $U(c_2) - w_2 h_2 / w_1$ increases less than $U(c_2) - h_2$) and does not affect the incentive-compatibility constraint of the high-productivity workers, it follows that $c_1 = c_2 < (U')^{-1}(1/q)$ is not possible. Instead, if $c_1 = c_2 \geq (U')^{-1}(1/q)$, the government could decrease c_1 and h_1 by small amounts, increase c_2 by a small amount, and keep h_2 unchanged such that $U(c_1) - h_1$ and $U(c_2)$ increase to the same extent. This will increase social welfare, and since it does not affect the incentive-compatibility constraint of the low-productivity workers and loosens the incentive-compatibility constraint of the high-productivity workers (because $U(c_1) - w_1 h_1 / w_2$ increases less than $U(c_1) - h_1$), it follows that $c_1 = c_2 \geq (U')^{-1}(1/q)$ is also not possible. Accordingly, it must be the case that $c_1 < c_2$.

If the incentive-compatibility constraint of the high-productivity workers would not bind, then social welfare could be improved by a small increase in c_1 balanced by a small decrease in c_2 determined such that h_1 and h_2 are unchanged and the budget constraint continues to be satisfied. Since this would not violate the incentive-compatibility constraint of the low-productivity workers, it follows that the incentive-compatibility constraint of the high-productivity workers must be binding, i.e., that

$$U(c_2) - h_2 = U(c_1) - \frac{w_1 h_1}{w_2}.$$

We then have that

$$\begin{aligned} & U(c_1) - h_1 \\ = \;& U(c_2) - \frac{w_2 h_2}{w_1} + \frac{w_2 h_2 - w_1 h_1}{w_1} \\ > \;& U(c_2) - \frac{w_2 h_2}{w_1}, \end{aligned}$$

so that the incentive-compatibility constraint of the low-productivity workers is not binding.

□

Appendix B

Proof of Propositions

Proposition 2: The marginal social cost of a low-productivity worker's consumption with a graduated minimum wage rate, given by the right-hand side of condition (15), is

$$\frac{n_1 q + n_2 f'(n_1 h_1^*/x)}{q\left[(n_1+2n_2)f'(n_1 h_1^*/x) - n_2 q\right]}, \tag{B1}$$

and the marginal social cost of a low-productivity worker's consumption without a graduated minimum wage rate, given by the right-hand side of condition (10), is

$$\frac{n_1 q + n_2 f'(n_1 \hat{h}_1/x) + n_2(n_1\hat{h}_1/x)f''(n_1\hat{h}_1/x)}{q\left[(n_1+2n_2)f'(n_1\hat{h}_1/x) - n_2 q + n_2(n_1\hat{h}_1/x)f''(n_1\hat{h}_1/x)\right]}.$$

If it were the case that $h_1^* = \hat{h}_1$, then the two social costs of a low-productivity worker's consumption would differ due to the term $n_2(n_1\hat{h}_1/x)f''(n_1\hat{h}_1/x)$ which is negative and appears additively in the numerator and denominator of the latter but not in the former. Since the derivative of

$$\frac{n_1 q + n_2 f'(n_1 h_1/x) + z}{q\left[(n_1+2n_2)f'(n_1 h_1/x) - n_2 q + z\right]}$$

with respect to z is

$$\frac{(n_1+n_2)\left[f'(n_1 h_1/x) - q\right]}{q\left[(n_1+2n_2)f'(n_1 h_1/x) - n_2 q + z\right]},$$

which is negative, it follows that if $h_1^* = \hat{h}_1$, then the marginal social cost of a low-productivity worker's consumption would be less with a graduated minimum wage rate than without.

The derivative of (B1) with respect to h_1^* is

$$-\frac{n_1(n_1+n_2)^2 f''(n_1 h_1^*/x)}{x\left[(n_1+2n_2)f'(n_1 h_1^*/x) - n_2 q + z\right]^2},$$

and hence positive, so the marginal social cost of c_1^* increases with h_1^*. As the marginal utility of a low-productivity worker's consumption decreases with c_1^*, it follows that either $c_1^* > \hat{c}_1$ or $h_1^* > \hat{h}_1$ or both $c_1^* > \hat{c}_1$ and $h_1^* > \hat{h}_1$. However, the loss of social welfare in the absence of a graduated minimum wage rate stems from the low-productivity workers consuming less and working more than in the first best, and the graduated minimum wage rate therefore increases social welfare by increasing their consumption and/or reducing their working hours. Formally, the total differential of the social welfare (1) given the budget constraint (8) and an unchanged consumption of the high-productivity workers (since their consumption is the same with and without a graduated minimum wage rate) is

$$n_1\left\{\left[U'(c_1)-\frac{1}{q}\right]dc_1+\left[\frac{f'(n_1h_1/x)}{q}-1\right]dh_1\right\}.$$

Since $f'(n_1h_1/x) < q$ implies that (B1) exceeds $1/q$ and hence that $U'(c_1) > 1/q$ for all relevant values of c_1, it is not possible that $dc_1 \leq 0$ and $dh_1 \geq 0$ as social welfare would then not increase. It follows that $dc_1 > 0$ so that $\hat{c}_1 < c_1^*$. Consequently, it can be concluded that $\hat{c}_1 < c_1^* < \hat{c}_2 = c_2^*$. □

Proposition 4: It has been shown in the main text that $0 = t_2^* = \hat{t}_2$. It has also been shown there that $U'(c_1^*) > 1/f'(n_1h_1^*/x)$, which implies that $0 < t_1^*$. If $h_1^* \geq \hat{h}_1$, then $f'(n_1h_1^*/x) \leq f'(n_1\hat{h}_1/x)$. Since Proposition 2 shows that $c_1^* > \hat{c}_1$ and hence $U'(c_1^*) < U'(\hat{c}_1)$, it follows that $U'(c_1^*)f'(n_1h_1^*/x) < U'(\hat{c}_1)f'(n_1\hat{h}_1/x) \Rightarrow t_1^* < \hat{t}_1$. If $h_1^* < \hat{h}_1$, we can then rewrite Condition (15) as

$$U'(c_1^*)f'(n_1h_1^*/x) = \frac{f'(n_1h_1^*/x)\left[n_1q+n_2f'(n_1h_1^*/x)\right]}{qA}.$$

The right-hand-side of this expression increases in h_1^* since its derivative with respect to h_1^* is

$$\frac{n_1n_2\left[f'(n_1h_1^*/x)-q\right]f''(n_1h_1^*/x)\left[(n_1+n_2)f''(n_1h_1^*/x)+n_1q\right]}{qxA^2},$$

which is positive. The proof of Proposition 2 shows that $U'(c_1^*)f'(n_1h_1^*/x) < U'(\hat{c}_1)f'(n_1\hat{h}_1/x)$ if $h_1^* = \hat{h}_1$, from which it follows that $h_1^* < \hat{h}_1 \Rightarrow U'(c_1^*)f'(n_1h_1^*/x) < U'(\hat{c}_1)f'(n_1\hat{h}_1/x) \Rightarrow t_1^* < \hat{t}_1$. Consequently, it is always true that $t_1^* < \hat{t}_1$. Hence, $0 = t_2^* = \hat{t}_2 < t_1^* < \hat{t}_1$. □

For brevity, we omit the argument $n_1 h_1^*/x$ of the f-function in the proofs of Propositions 5-9.

Proposition 5: Differentiating conditions (15) and (16) and constraints (8), (12), and (14) totally with respect to R yields

$$\frac{dm^*}{dR} = \frac{x\left(m^* - f'\right) A^2 U''(c_1^*)}{B},$$

where

$$B \equiv -(n_1 + n_2) h_1^* \left\{ n_1 (n_1 + n_2) \left[n_1 + n_2 q U'(c_1^*) \right] f'' + x f' A^2 U''(c_1^*) \right\}$$

is positive since $f'' < 0$ and $U''(c_1^*) < 0$. It follows that dm^*/dR has the same sign as $x\left(m^* - f'\right) A^2 U''(c_1^*)$, which is negative. Hence, $dm^*/dR < 0$. □

Proposition 6: Differentiating (15), (16), (8), (12), and (14) totally with respect to n_1 yields

$$\frac{dm^*}{dn_1} = \frac{x\left(m^* - f'\right) \left\{ -n_1 n_2 \left(q - f'\right)^2 \left[n_1 + n_2 q U'(c_1^*) \right] + (c_1^* n_1 + m^* h_1^* n_2) q A^2 U''(c_1^*) \right\}}{n_1 q B},$$

which has the same sign as

$$-n_1 n_2 \left(q - f'\right)^2 \left[n_1 + n_2 q U'(c_1^*) \right] + (c_1^* n_1 + m^* h_1^* n_2) q A^2 U''(c_1^*).$$

Since both terms in this expression are negative, it follows that $dm^*/dn_1 < 0$. □

Proposition 7: Differentiating (15), (16), (8), (12), and (14) totally with respect to n_2 yields

$$\frac{dm^*}{dn_2} = \frac{x\left(m^* - f'\right) \left\{ n_1 \left(q - f'\right)^2 \left[n_1 + n_2 q U'(c_1^*) \right] - q (h_2^* q - c_2^*) A^2 U''(c_1^*) \right\}}{qB},$$

which has the same sign as

$$n_1 \left(q - f'\right)^2 \left[n_1 + n_2 q U'(c_1^*) \right] - q (h_2^* q - c_2^*) A^2 U''(c_1^*). \tag{17}$$

If $qh_2^* \leq c_2^*$, then the budget constraint implies that $m^* h_1^* \geq c_1^*$. The modified incentive-compatibility constraint of the high-productivity workers would then be violated since

$$\begin{aligned} & U(c_2^*) - h_2^* - U(c_1^*) + \frac{m^* h_1^*}{q} \\ \geq\ & U(c_2^*) - \frac{c_2^*}{q} - U(c_1^*) + \frac{c_1^*}{q} \\ >\ & 0, \end{aligned}$$

where the last inequality follows from the fact that $U(c) - c/q$ increases in c for $c < c_2^*$ (in view of condition (16)) and that $c_1^* < c_2^*$. It follows that $qh_2^* > c_2^*$. Hence, (17) is positive which implies that $dm^*/dn_2 > 0$.

Proposition 8: Differentiating (15), (16), (8), (12), and (14) totally with respect to x yields

$$\frac{dm^*}{dx} = -\frac{(h_1^* m^* n_2 + xf)(m^* - f') A^2 U''(c_1^*)}{B},$$

which is positive. □

Proposition 9: Differentiating (15), (16), (8), (12), and (14) totally with respect to q yields

$$\frac{dm^*}{dq} = \frac{n_2 x (m^* - f') \{(q - f') [n_1 q + (n_1 + 2n_2) f'] [n_1 + n_2 q U'(c_1^*)] - A^2 U''(c_1^*) q (h_2^* q - h_1^* m^*)\}}{q^2 B},$$

which has the same sign as

$$(q - f') [n_1 q + (n_1 + 2n_2) f'] [n_1 + n_2 q U'(c_1^*)] - A^2 U''(c_1^*) q (h_2^* q - h_1^* m^*).$$

This expression is positive since the modified incentive-compatibility constraint of the high-productivity workers implies that $h_2^* q - h_1^* m^* > 0$. Consequently, $dm^*/dq > 0$. □

References

Allen, Stephen P. 1987. "Taxes, Redistribution, and the Minimum Wage: A Theoretical Analysis." *Quarterly Journal of Economics,* 102(3): 477-89.

Basu, Arnab K., Nancy H. Chau, and Ravi Kanbur. 2010. "Turning a Blind Eye: Costly Enforcement, Credible Commitment and Minimum Wage Laws." *Economic Journal,* (forthcoming).

Bhaskar, V., Alan Manning, and Ted To. 2002. "Oligopsony and Monopsonistic Competition in Labor Markets." *Journal of Economic Perspectives*, 16(2): 155-74.

Blumkin, Tomer and Efraim Sadka. 2005. "Income Taxation and Wage Policy: An Application to Minimum Wage." *International Tax and Public Finance*, 12(6): 713-22.

Boadway, Robin, and Katherine Cuff. 2001. "A Minimum Wage Can Be Welfare-Improving and Employment-Enhancing." *European Economic Review*, 45(3): 553-76.

Burkhauser, Richard V., Kenneth A. Couch, and David C. Wittenburg. 2000. "A Reassessment of the New Economics of the Minimum Wage Literature with Monthly Data from the Current Population Survey." *Journal of Labor Economics*, 18(4): 653-80.

Cahuc, Pierre and Guy Laroque. 2009. "Optimal Taxation and Monopsonistic Labor Market: Does Monopsony Justify the Minimum Wage?" Working Paper.

Card, David, and Alan B. Krueger. 1995. *Myth and Measurement: The New Economics of the Minimum Wage.* Princeton, NJ: Princeton University Press.

Danziger, Leif. 2009. "The Elasticity of Labor Demand and the Minimum Wage." *Journal of Population Economics*, 22(3): 757-72.

Flinn, Christopher J. 2006. "Minimum Wage Effects on Labor Market Outcomes under Search, Matching, and Endogenous Contact Rates." *Econometrica*, 74(4): 1013-62.

Guesnerie, Roger; Roberts, Kevin. 1987. "Minimum Wage Legislation as a Second Best Policy." *European Economic Review,* 31(1/2): 490-98.

Hungerbühler, Mathias, and Etienne Lehmann. 2009. "On the Optimality of a Minimum

Wage: New Insights from Optimal Tax Theory." *Journal of Public Economics,* 93(3-4): 464-81.

Lee, David and Emmanuel Saez. 2009. "Optimal Minimum Wage Policy in Competitive Labor Markets." Working Paper.

Marceau, Nicolas, and Robin Boadway. 1994. "Minimum Wage Legislation and Unemployment Insurance as Instruments for Redistribution." *Scandinavian Journal of Economics*, 96(1): 67-81.

Mortensen, Dale T., and Christopher A. Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies*, 61(3): 397-415.

Neumark, David, and William L. Wascher. 2007. "Minimum Wages and Employment." *Foundations and Trends in Microeconomics*, 3(1-2): 1-186.

Neumark, David, Mark Schweitzer, and William Wascher. 2004. "Minimum Wage Effects throughout the Wage Distribution." *Journal of Human Resources*, 39(2): 425-50.

Rebitzer, James B., and Lowell J. Taylor. 1995. "The Consequences of Minimum Wage Laws: Some New Theoretical Ideas." *Journal of Public Economics*, 56(2): 245-55.

Robinson, Joan. 1933. *The Economics of Imperfect Competition.* London: MacMillan.

Stiglitz, Joseph E. 1982. "Self-Selection and Pareto Efficient Taxation." *Journal of Public Economics*, 17(2): 213-40.

Zavodny, Madeline. 2000. "The Effect of the Minimum Wage on Employment and Hours." *Labour Economics*, 7(6): 729-50.