A Pareto-Improving Minimum Wage*

Eliav Danziger[1] and Leif Danziger[2]

Abstract

This paper shows that a graduated minimum wage, in contrast to a constant minimum wage, can provide a Pareto improvement over what can be achieved with an optimal income tax. The reason is that a graduated minimum wage requires high-productivity workers to work more to earn the same income as low-productivity workers, which makes it more difficult for the former to mimic the latter. In effect, a graduated minimum wage allows the low-productivity workers to benefit from second-degree price discrimination which increases their income.

JEL Classification: J30
Keywords: Graduated minimum wage, Pareto improvement

---

* We thank Pierre Cahuc and two anonymous referees for helpful comments.
[1] Department of Economics, Princeton University, Princeton, NJ 08544, USA

[2] Department of Economics, Ben-Gurion University, Beer-Sheva 84105, Israel

1 Introduction

In a competitive economy, a constant minimum wage cannot be used to make everyone better off and thereby provide a Pareto improvement over what can be achieved by an optimal income tax alone. Indeed, as shown by Allen (1987) and Guesnerie and Roberts (1987), a binding constant minimum wage would be counterproductive as it would reduce the hours of low-productivity workers. The reason is that the government could have attained the same lower working hours by an appropriately designed income tax alone, but has refrained from doing so. In other words, the power to set a constant minimum wage does not provide the government with an extra instrument to affect the feasible consumption-work bundles.

Subsequent papers have tried to justify the minimum wage by modifying the competitive framework and changing the objective to be to increase social welfare rather than to provide a Pareto improvement. Thus, if the government's social welfare function expresses a sufficiently strong taste for redistribution toward the less productive workers, combining a constant minimum wage with an optimal income tax may lead to an increase in social welfare if it forces some low-productivity workers to become unemployed (Marceau and Boadway, 1994); if it is combined with a welfare policy that obliges workers to accept job offers (Boadway and Cuff, 2001); if it prevents workers from signalling their earning ability (Blumkin and Sadka, 2005); and if workers with the highest disutilities of work are laid off first (Lee and Saez, 2012).

In the absence of optimal income taxation, a constant minimum wage may increase social welfare both in a competitive environment (Danziger, 2009; and Lee and Saez, 2012)[1] and in a monopsonistic environment (Robinson, 1933).[2] Since search frictions and informational asymmetries may lead to monopsonistic features, a constant minimum wage may be rationalized with other types of labor markets if income taxation is not optimal (Rebitzer

[1] See also Stewart and Swaffield (2008), De Fraja (1999), and Strobl and Walsh (2011).

[2] However, with optimal income taxation, monopsony cannot justify a constant minimum wage (Cahuc and Laroque, in press).

and Taylor, 1995; Bhaskar and To, 1999; Cahuc et al., 2001; Bhaskar et al., 2002; Manning, 2003; Flinn, 2006; Hungerbühler and Lehmann, 2009; and Basu et al., 2010). Additionally, politics, the power of unions, and cultural factors may affect the setting of the minimum wage (Sobel, 1999; Belot et al., 2007; Boeri and Burda, 2009; Brown, 2009; Checchi and Garcia-Penalosa, 2010; Aghion et al., 2011; and Boeri, 2012).

In the current paper we pursue a different path and provide a novel justification for the minimum wage without modifying the competitive framework or changing the objective function. Our approach is designed to be as close as possible to the standard competitive model with two types of workers and optimal income taxes (Stiglitz, 1982). As we will show, our minimum wage design can provide a Pareto improvement of the second-best optimum achievable with taxation alone.

There is an obvious dissonance between insisting on the minimum wage being constant while allowing for a nonlinear income tax schedule. Thus, in the spirit of the optimal income tax literature, this paper analyzes a graduated minimum wage that depends on the total employment of a firm's minimum-wage workers. We show that a graduated minimum wage severs the tight connection between the wage and a workers' marginal product, thereby providing a tool for compelling firms to hire low-productivity workers at a minimum wage that exceeds their marginal product.

A graduated minimum wage can be designed so that a high-productivity worker needs to work more hours in order to earn the same income as a low-productivity worker. This makes it less attractive for high-productivity workers to mimic the low-productivity workers by earning their income. Consequently, it is easier to distinguish between workers with different productivities, and therefore also to redistribute toward the less productive workers. We prove that even if the economy is initially in a second-best optimum with an optimal income tax, the introduction of a graduated minimum wage can benefit all individuals and thus constitute a Pareto improvement.[3]

[3] As pointed out by Guesnerie and Roberts (1987), there is an apparent informational inconsistency

If the graduated minimum wage is differentiable, we show that a firm's marginal labor cost is less than the minimum wage. Effectively, therefore, firms are facing a downward-sloping labor supply of low-productivity workers so that these workers are employed more than they would be at a constant minimum wage equal to their marginal product. In other words, the graduated minimum wage endows the workers with monopoly power that allows the practicing of second-degree price discrimination which increases their income.

2 The Standard Competitive Model with Taxes

We consider an economy with a continuum of low-productivity workers of measure $n_1 > 0$ and a continuum of high-productivity workers of measure $n_2 > 0$. A worker's utility function is $U(c, h)$, where $c \geq 0$ is the worker's consumption and $h \geq 0$ is the working hours. The utility function satisfies $U_c > 0$, $U_{cc} < 0$, $U_h < 0$, $U_{hh} \leq 0$, and the agent monotonicity property, i.e., hU_h/U_c is decreasing in h.

The economy also includes a unit continuum of homogenous firms that produce a single good whose price is unity. A firm's production function is $f(\ell_1, \ell_2)$, where ℓ_1 and ℓ_2 are the total working hours of the low- and high-productivity workers, respectively. The production functions satisfy $f(0, 0) = 0$ and exhibit decreasing returns to scale, a decreasing marginal product of the low-productivity workers, and an imperfect substitutability between the two types of workers (i.e., the isoquants are not straight lines). Due to a bias in the production functions in favor of the high-productivity workers and/or a large ratio of low- to high-

in assuming that while the government cannot observe the workers' wages (and hence condition taxes on wages), it can nevertheless enforce a minimum wage. However, Lee and Saez (2012, p. 746) argue that even though it may not be possible for the government to observe wages, enforcement of a minimum wage can be brought about by a system of audits triggered by worker complaints and hefty penalties for noncompliant firms. As we show that there exists a graduated minimum wage schedule which is downward sloping at the intended working hours, a worker's complaint of being paid less than the specified minimum wage (given the firm's total employment of low-productivity workers) is equivalent to a complaint that the firm's total employment of low-productivity workers is improper (given the minimum wage). The issue of noncompliance is studied in Ashenfelter and Smith (1979), Weil (2005), and Danziger (2010). Alternatively, it may be the case that a governmental enforcement agency can observe workers' wages, but that political or institutional constraints limit taxes to depend on only income.

productivity workers, in all relevant allocations a low-productivity worker's marginal product is less than that of a high-productivity worker. This justifies our classification of workers as having either low or high productivity.

The low- and high-productivity workers' competitive wages are w_1 and w_2, respectively, where $w_1 < w_2$. Assuming that working hours are positive for both types of workers, the decreasing returns to scale imply that profits are positive in a competitive equilibrium, i.e., $f(\ell_1, \ell_2) - w_1\ell_1 - w_2\ell_2 > 0$.[4] Also, we let (c_1, h_1) and (c_2, h_2) denote the consumption and working hours of the low- and high-productivity workers, respectively, where $h_1 \equiv \ell_1/n_1$ and $h_2 \equiv \ell_2/n_2$.

In addition to workers and firms, there is a government that wants to transfer consumption from the high- to the low-productivity workers (i.e., $c_1 - w_1h_1 > c_2 - w_2h_2$). As is standard in the tax literature, taxation is based on incomes and cannot depend directly on wages. Since this rules out a first-best allocation in which lump-sum taxation is based on productivity levels, the government chooses a second-best allocation by determining a nonlinear income tax that transfers consumption from the high- to the low-productivity workers. The government's choice has to satisfy the incentive-compatibility constraints that workers of one type cannot make themselves better off by choosing their working hours so that they earn the same income as workers of the other type. Let a circumflex over a variable denote its value in a second-best optimum with nonlinear income taxes and redistribution toward the low-productivity workers, but without a minimum wage. Then the low- and high-productivity workers' incentive-compatibility constraints are

$$U(\hat{c}_1, \hat{h}_1) \geq U\left(\hat{c}_2, \frac{\hat{w}_2\hat{h}_2}{\hat{w}_1}\right), \tag{1}$$

$$U(\hat{c}_2, \hat{h}_2) \geq U\left(\hat{c}_1, \frac{\hat{w}_1\hat{h}_1}{\hat{w}_2}\right), \tag{2}$$

where $\hat{w}_2\hat{h}_2/\hat{w}_1$ (resp. $\hat{w}_1\hat{h}_1/\hat{w}_2$) are the hours that a low-productivity (resp. high-productivity)

[4] The government taxes all profits away.

worker would have to work in order to reach a high-productivity (resp. low-productivity) worker's income. The agent monotonicity property implies that the incentive-compatibility constraint of the low-productivity workers (1) is slack in a second-best optimum, while that of the high-productivity workers (2) binds (Salanié, 2003).

The government's choice must also satisfy the resource constraint

$$f(\hat{\ell}_1, \hat{\ell}_2) - n_1 \hat{c}_1 - n_2 \hat{c}_2 = 0. \tag{3}$$

In sum, in a second-best redistributive optimum the government maximizes $U(c_1, h_1)$ subject to a given level of $U(c_2, h_2)$ and the binding constraints (2) and (3).

3 A Graduated Minimum Wage

Suppose that the low-productivity workers must be paid according to a graduated minimum wage schedule $m(\ell_1)$, which would make the lowest wage a firm can pay a function of the total working hours of all the low-productivity workers employed in the firm. Any potential minimum wage must be the lowest wage in the economy (even if never chosen). Therefore, $m(\ell_1)$ cannot exceed a high-productivity worker's wage for any ℓ_1; that is, $m(\ell_1) \leq w_2$ for any ℓ_1. Letting an asterisk indicate the value of a variable in an equilibrium with a graduated minimum wage, the low- and high-productivity workers' incentive-compatibility constraints become

$$U(c_1^*, h_1^*) \geq U\left(c_2^*, \frac{w_2^* h_2^*}{w_1^*}\right), \tag{4}$$

$$U(c_2^*, h_2^*) \geq U\left(c_1^*, \frac{w_1^* h_1^*}{w_2^*}\right). \tag{5}$$

Just as the income-tax function offers the workers a choice between different consumption-work bundles, the graduated minimum wage offers the firms a choice between different minimum wage-hour bundles. With the optimal income tax one need only be concerned with determining the intended consumption-work bundles for the two types of workers (because all other bundles can be made arbitrarily unattractive). However, with the optimal graduated minimum wage one need be concerned not only with determining the intended minimum wage-hours bundle, but also with making sure that a firm's profit with that bundle is at least as high as the profit that can be attained with any other minimum wage-hours bundle. That is, a firm's choice of working hours of low- and high-productivity workers must satisfy the minimum wage constraint

$$f(\ell_1^*, \ell_2^*) - m(\ell_1^*)\ell_1^* - w_2^*\ell_2^* \geq \max_{\ell_1 \neq \ell_1^*, \ell_2} [f(\ell_1, \ell_2) - m(\ell_1)\ell_1 - w_2^*\ell_2]. \tag{6}$$

Since a firm always has the option of not producing by setting $\ell_1 = \ell_2 = 0$ and hence earning zero profit, the right-hand-side of (6) is nonnegative. Thus, there exist working hours (ℓ_1^*, ℓ_2^*) for which a firm's profit is nonnegative and the minimum wage constraint is satisfied.

The resource constraint becomes

$$f(\ell_1^*, \ell_2^*) - n_1 c_1^* - n_2 c_2^* = 0, \tag{7}$$

and in an equilibrium the government's choice of a nonlinear income tax and a graduated minimum wage must satisfy the low- and the high-productivity workers' incentive-compatibility constraints (4) and (5), the firms' minimum wage constraint (6), and the resource constraint (7).

We will now show that even if the economy is initially in a second-best optimum with a nonlinear income tax that redistributes toward the less productive workers, the introduction of a graduated minimum wage can achieve a Pareto improvement:

Proposition 1: *A graduated minimum wage can provide a Pareto improvement of a second-best optimum with nonlinear income taxation that redistributes consumption toward the less productive workers.*

Proof: In three steps we construct a graduated minimum wage that increases the utility of all workers in equilibrium.

Step 1: *A Graduated Minimum Wage that Leaves the Utility of All Workers Unchanged*

Suppose that the graduated minimum wage be given by

$$m(\ell_1) = \begin{cases} \hat{w}_1 + \alpha & \text{for } \ell_1 = \hat{\ell}_1, \\ \hat{w}_2 & \text{for } \ell_1 \neq \hat{\ell}_1, \end{cases}$$

where $0 < \alpha < \hat{w}_2 - \hat{w}_1$. That is, the wage is α higher than $\hat{w}_1$ at the same total hours of low-productivity workers as in the second-best allocation with only nonlinear income taxes, and $\hat{w}_2 - \hat{w}_1$ higher than $\hat{w}_1$ at all other hours.

Let the intended consumption-work bundle for the low-productivity workers remain unchanged, which requires that their tax payment increases by $\alpha\hat{h}_1$. Also, let the intended consumption-work bundle for the high-productivity workers remain unchanged, which requires that the income tax when earning $\hat{w}_2\hat{h}_2$ be unchanged. The low-productivity workers' incentive-compatibility constraint (4) will then be the same as (1) with $\hat{w}_1 + \alpha$ substituted for $\hat{w}_1$, and it will remain slack if α is small enough; that is,

$$U(\hat{c}_1, \hat{h}_1) > U\left(\hat{c}_2, \frac{\hat{w}_2\hat{h}_2}{\hat{w}_1 + \alpha}\right). \tag{8}$$

The high-productivity workers have to work more in order to earn the income of the low-productivity workers. Therefore, their incentive-compatibility constraint (5) will be the same as (2) with $\hat{w}_1 + \alpha$ substituted for $\hat{w}_1$. As the right-hand-side of (2) decreases, their incentive-compatibility constraint is loosened and will be slack; that is

$$U(\hat{c}_2, \hat{h}_2) > U\left[\hat{c}_1, \frac{(\hat{w}_1 + \alpha)\hat{h}_1}{\hat{w}_2}\right]. \tag{9}$$

If it were the case that $\alpha = 0$ so that $m(\hat{\ell}_1) = \hat{w}_1$, but $m(\ell_1) = \hat{w}_2$ for $\ell_1 \neq \hat{\ell}_1$, then the left-hand-side of the minimum wage constraint (6) would equal a firm's profit in the absence of a minimum wage, but the right-hand-side would be less since the hourly wage of low-productivity workers have increased from $\hat{w}_1$ to $\hat{w}_2$ for any $\ell_1 \neq \hat{\ell}_1$. The constraint would therefore hold with a strict inequality. In reality, though, the intended minimum wage exceeds $\hat{w}_1$ by $\alpha > 0$ so that a firm's profit at $(\hat{\ell}_1, \hat{\ell}_2)$ decreases by $\alpha\hat{\ell}_1$. However, for a sufficiently small α and any given $\ell_1 > 0$ we have that $\alpha\hat{\ell}_1 < (\hat{w}_2 - \hat{w}_1)\ell_1$ so that for small α's the left-hand-side of the minimum wage constraint (6) is less than the right-hand-side. This implies that for a sufficiently small α the minimum wage constraint becomes slack and a firm's profit is positive; that is

$$f(\hat{\ell}_1, \hat{\ell}_2) - (\hat{w}_1 + \alpha)\hat{\ell}_1 - \hat{w}_2\hat{\ell}_2 > \max_{\ell_1 \neq \hat{\ell}_1, \ell_2} [f(\ell_1, \ell_2) - \hat{w}_2\ell_1 - \hat{w}_2\ell_2]. \tag{10}$$

In actuality, therefore, the graduated minimum wage compels the firms to pay low-productivity workers more than the competitive wage while preventing them from shrinking the labor input of these workers.

Since the workers' consumption and working hours do not change, the resource constraint will not be affected. As a result, the graduated minimum wage schedule detailed above leads to an equilibrium in which the incentive-compatibility constraints of both low- and high-productivity workers as well as the minimum wage constraint are slack. Nonetheless, the consumption and working hours of each type of worker, and hence their utilities, will be the same as in a second-best optimum with nonlinear income taxes and redistribution toward the low-productivity workers.

Step 2: *A Graduated Minimum Wage Increasing the Utility of Low-Productivity Workers*

In a second-best optimum with nonlinear income taxes and redistribution toward the low-productivity workers, their marginal rate of substitution between consumption and leisure will be less than their marginal product (see Appendix). Therefore, the low-productivity workers would benefit if their working hours were increased and they would get to consume all the additional production. One implication is that there exists an increase in each low-productivity worker's hours $\gamma > 0$ such that $U(\hat{c}_1 + \delta, \hat{h}_1 + \gamma) > U(\hat{c}_1, \hat{h}_1)$, where $\delta \equiv [f(\hat{\ell}_1 + n_1\gamma, \hat{\ell}_2) - f(\hat{\ell}_1, \hat{\ell}_2)]/n_1$. Here, $n_1\gamma$ is the increase in the total working hours of low-productivity workers in a firm (because there is a unit continuum of firms) and $n_1\delta$ is the corresponding increase in the output.

We now let the graduated minimum wage for $\hat{\ell}_1 + n_1\gamma$ working hours of low-productivity workers in a firm be equal to $[(\hat{w}_1+\alpha)\hat{h}_1+\delta]/(\hat{h}_1+\gamma)$, and the income tax when earning $m(\hat{\ell}_1+n_1\gamma)(\hat{h}_1+\gamma)=(\hat{w}_1+\alpha)\hat{h}_1+\delta$ be such that the low-productivity workers' consumption is $\hat{c}_1+\delta$. We also let the high-productivity workers' income tax be adjusted such that if their wage changes to $\tilde{w}_2 \equiv \partial f(\hat{\ell}_1+n_1\gamma, \hat{\ell}_2)/\partial\hat{\ell}_2$, then their consumption and working hours remain unchanged. Finally, we let the graduated minimum wage be equal to $\tilde{w}_2$ for $\ell_1 \neq \hat{\ell}_1+n_1\gamma$.

The low- and high-productivity workers' incentive-compatibility constraints are

$$U(\hat{c}_1+\delta, \hat{h}_1+\gamma) \geq U\left[\hat{c}_2, \frac{\tilde{w}_2\hat{h}_2}{m(\hat{\ell}_1+n_1\gamma)}\right], \tag{11}$$

$$U(\hat{c}_2, \hat{h}_2) \geq U\left[\hat{c}_1+\delta, \frac{(\hat{w}_1+\alpha)\hat{h}_1+\delta}{\tilde{w}_2}\right], \tag{12}$$

and the minimum wage constraint is

$$f(\hat{\ell}_1+n_1\gamma, \hat{\ell}_2) - m(\hat{\ell}_1+n_1\gamma)\hat{\ell}_1 - \tilde{w}_2\hat{\ell}_2 \geq \max_{\ell_1\neq\hat{\ell}_1+n_1\gamma, \ell_2} [f(\ell_1, \ell_2) - \tilde{w}_2\ell_1 - \tilde{w}_2\ell_2]. \tag{13}$$

Since the constraints (11), (12), and (13) would be identical to (8), (9), and (10) for $\gamma=\delta=0$, and the latter are slack, there exist positive values of γ for which (11), (12), and (13) are also slack. The resource constraint would be satisfied. Hence, there is a graduated minimum wage leading to an equilibrium in which the low-productivity workers's utility increases while the high-productivity workers' utility remains unchanged.

Step 3: *A Graduated Minimum Wage that Increases the Utility of All Workers*

Consider a value of γ for which (11), (12), and (13) are slack. The fact that the low-productivity workers gain and the high-productivity workers are indifferent implies that there exists a small $\epsilon > 0$ such that if, for the same intended minimum wage as in step 2, the low-productivity workers' consumption is reduced from $\hat{c}_1+\delta$ to $\hat{c}_1+(1-\epsilon)\delta$ and the high-productivity workers' consumption is increased from $\hat{c}_2$ to $\hat{c}_2+\epsilon\delta n_1/n_2$, then:

- both the low-productivity workers' utility $U[\hat{c}_1+(1-\epsilon)\delta, \hat{h}_1+\gamma]$ and the high-productivity workers' utility $U(\hat{c}_2+\epsilon\delta n_1/n_2, \hat{h}_2)$ will be higher than in the second-best optimum with nonlinear income taxes and redistribution toward the low-productivity workers;

- both the low- and the high-productivity workers' incentive-compatibility constraints will be slack, i.e.,

$$\begin{aligned} U[\hat{c}_1 + (1-\epsilon)\delta, \hat{h}_1 + \gamma] &> U\left[\hat{c}_2 + \frac{\epsilon\delta n_1}{n_2}, \frac{\tilde{w}_2\hat{h}_2}{m(\hat{\ell}_1 + n_1\gamma)}\right], \\ U\left(\hat{c}_2 + \frac{\epsilon\delta n_1}{n_2}, \hat{h}_2\right) &> U\left[\hat{c}_1 + (1-\epsilon)\delta, \frac{(\hat{w}_1+\alpha)\hat{h}_1 + \delta}{\tilde{w}_2}\right]; \end{aligned}$$

- the minimum wage constraint will be slack, i.e.,

$$f(\hat{\ell}_1 + n_1\gamma, \hat{\ell}_2) - m(\hat{\ell}_1 + n_1\gamma)(\hat{\ell}_1 + n_1\gamma) - \tilde{w}_2\hat{\ell}_2 > \max_{\ell_1 \neq \hat{\ell}_1, \ell_2} [f(\ell_1, \ell_2) - \tilde{w}_2\ell_1 - \tilde{w}_2\ell_2].$$

Since the resource constraint continues to hold, we have proved that there exists an equilibrium with the graduated minimum wage

$$m(\ell_1) = \begin{cases} w_1^* & \text{for } \ell_1 = \ell_1^*, \\ \tilde{w}_2 & \text{for } \ell_1 \neq \ell_1^*, \end{cases}$$

where $w_1^* = [(\hat{w}_1 + \alpha)\hat{h}_1 + \delta]/(\hat{h}_1 + \gamma)$. In this equilibrium, $c_1^* = \hat{c}_1 + (1-\epsilon)\delta$; $h_1^* = \hat{h}_1 + \gamma$; $c_2^* = \hat{c}_2 + \epsilon\delta n_1/n_2$; $h_2^* = \hat{h}_2$; and $w_2^* = \tilde{w}_2$. All workers obtain a higher utility than in the second-best optimum with only nonlinear income taxes and redistribution toward the low-productivity workers. This completes the proof. □

As the construction of the proof of Proposition 1 shows, a Pareto improvement is possible because a graduated minimum wage is more efficacious than income taxes in directing resources to the low-productivity workers. Specifically, the graduated minimum wage funnels more of the firms' revenues to the low-productivity workers, thereby raising their pretax income without affecting that of the high-productivity workers (see step 1 of the proof). High-productivity workers will then have to work more in order to earn the same pretax income as low-productivity workers, and this mitigates the high-productivity workers' incentive-compatibility constraint. Consequently, the graduated minimum wage makes it easier for the government to distinguish between the low- and high-productivity workers, which facilitates a Pareto improvement.[5]

The graduation of the minimum wage is essential for achieving a Pareto improvement, as Allen (1987) and Guesnerie and Roberts (1987) have shown that a constant minimum wage cannot do so. This is because, on the one hand, a constant minimum wage that does not exceed the competitive wage is ineffectual, while, on the other, a constant minimum wage that does exceed the competitive wage will reduce the working hours of the low-productivity workers. However, the same lower working hours could also have been obtained with a nonlinear income tax alone. The fact that the government has chosen not to do so reveals that it would be deleterious.

4 A Continuous Piecewise Differentiable Graduated Minimum Wage

In our formulation until now, the graduated minimum wage is not continuous at the intended working hours of the low-productivity workers (since it equals $[(\hat{w}_1 + \alpha)\hat{h}_1 + \delta]/(\hat{h}_1 + \gamma)$ for the intended working hours and equals $\tilde{w}_2$ for other working hours). To shed further light on the source of the Pareto improvement, assume now that the minimum wage schedule, which we still write as $m(\ell_1)$, is locally differentiable at the intended working hours of the low-productivity workers. To maximize its profit, a firm would then set the low-productivity workers' hours such that the marginal labor cost, $\partial[m(\ell_1)\ell_1]/\partial\ell_1 = m(\ell_1) + m'(\ell_1)\ell_1$, equals their marginal product. Accordingly, if the minimum wage constraint is satisfied, profit

[5] An alternative to the assumption that the government taxes all profits away would be to assume that there is a given unit continuum of firm owners each owning one firm, and that only profits above the normal level $\overline{\pi} > 0$ are taxed away. The resource constraint for the economy would then be modified by substituting $\overline{\pi}$ instead of zero at the right-hand-side of (3), and Proposition 1 would state that a graduated minimum wage can provide a Pareto improvement for the workers while leaving the firm owners no worse off. By applying a similar argument as in step 3 of the proof of Proposition 1, the proposition could be generalized to state that there exists a graduated minimum wage which can provide a Pareto improvement for both the workers and the firm owners.

maximization implies that[6]

$$m(\ell_1) + m'(\ell_1)\,\ell_1 = f_{\ell_1}. \tag{14}$$

In an equilibrium with a Pareto-improving-graduated minimum wage, each firm employs low-productivity workers for more hours than it would choose if the wage were constant at $\hat{w}_1$ and the firm could freely decide the hours. Accordingly, the minimum wage exceeds a low-productivity worker's marginal product, i.e., $m(\ell_1) > f_{\ell_1}$. By (14), this implies that the marginal labor cost of low-productivity workers, $m(\ell_1) + m'(\ell_1)\ell_1$, is less than the minimum wage, $m(\ell_1)$, and hence that the marginal minimum wage, $m'(\ell_1)$, is negative. That is, a Pareto improvement would be achieved by a locally differentiable minimum wage schedule that is downward sloping at the intended hours. In effect, the graduated minimum wage transforms the low-productivity workers into second-degree price discriminating monopolists that confront firms with a downward-sloping labor supply at the intended hours. The result is that the low-productivity workers are able to extract some of the firms' revenues and that they earn more than they would if the firms were permitted to pay the same minimum wage for fewer working hours.

We can now establish:

Proposition 2: *There exists a continuous piecewise differentiable graduated minimum wage schedule and income taxes that can implement the same Pareto improvement as the non-continuous graduated minimum wage constructed in the proof of Proposition 1.*

Proof: Let the intended minimum wage $m(\ell_1^*)$ and the income taxes be the same as in the proof of Proposition 1. Further, let the graduated minimum wage schedule be given by

$$m(\ell_1) = \begin{cases} f_{\ell_1^*} + \dfrac{\beta}{\ell_1} & \text{for } \ell_1 > \dfrac{\beta}{\tilde{w}_2 - f_{\ell_1^*}}, \\ \tilde{w}_2 & \text{for } \ell_1 \leq \dfrac{\beta}{\tilde{w}_2 - f_{\ell_1^*}}, \end{cases}$$

[6] The minimum wage constraint implies that the second-order condition for profit maximization holds, i.e., that $2m'(\ell_1) + m''(\ell_1)\,\ell_1 > f_{\ell_1\ell_1}$.

where $\beta \equiv (w_1^* - f_{\ell_1^*})\ell_1^*$ and $f_{\ell_1^*}$ is evaluated at (ℓ_1^*, ℓ_2^*). Note that the minimum wage is always continuous and that the marginal labor cost is $f_{\ell_1^*}$ for $\ell_1 > \beta/(\tilde{w}_2 - f_{\ell_1^*})$ and is $\tilde{w}_2$ for $\ell_1 \leq \beta/(\tilde{w}_2 - f_{\ell_1^*})$. Thus, profit maximization implies that a firm sets the low-productivity workers' hours to ℓ_1^* and that the corresponding (intended) graduated minimum wage is w_1^*.

Since income taxes are the same as before, it is immediate that the low- and high-productivity workers' incentive-compatibility constraints and the minimum wage constraint will be slack. It is also clear that the resource constraint still holds. Hence, the same Pareto-improving equilibrium as constructed in the proof of Proposition 1 is achieved. $\square$

The graduated minimum wage used in the proof of Proposition 2 is, in the relevant range, a hyperbolic function of the low-productivity workers' hours. This implies that the marginal labor cost is a constant that equals $f_{\ell_1^*}$ independently of the value of β. Therefore, the hours are the same as they would be with a competitive wage w_1^*, but the low-productivity workers nevertheless succeed in increasing their total pay by an additional β.

5 Heterogenous Firms

We now generalize the model by allowing for heterogenous firms. Thus, suppose that there are two categories of firms, with x_j being the measure of firms in category $j \in \{1, 2\}$ and $x_1 + x_2 = 1$. The production function of a firm in category j is $f_j(\ell_{1j}, \ell_{2j})$, where ℓ_{1j} and ℓ_{2j} are the total hours of labor input of the low- and high-productivity workers, respectively. The production functions satisfy $f_j(0, 0) = 0$ and exhibit decreasing returns to scale, a decreasing marginal product of the low-productivity workers, and an imperfect substitutability between the two types of workers. Each worker can only work in one firm.

In the competitive framework with nonlinear income taxes but no graduated minimum wage, we continue to denote the low- and high-productivity workers' competitive wages by w_1 and w_2, respectively, and the high-productivity workers' consumption and working hours by (c_2, h_2). We let $n_{1j} > 0$, where $n_{11} + n_{12} = n_1$, denote the measure of low-productivity workers that work for a firm in category j, and c_{1j} and $h_{1j} \equiv \ell_{1j}/n_{1j}$ denote the consumption and working hours of low-productivity workers that work for a firm in category j. The working hours are assumed positive for both types of workers in both categories of firms. It follows that in equilibrium all low-productivity workers will obtain the same consumption-work bundle no matter where they work, i.e., $(\hat{c}_{11}, \hat{h}_{11}) = (\hat{c}_{12}, \hat{h}_{12}) = (\hat{c}_1, \hat{h}_1)$. Furthermore, the decreasing returns to scale imply that in a second-best optimum with nonlinear income taxes, the profits are positive for both firm categories.

The low- and high productivity workers' incentive-compatibility constraints are again given by (1) and (2), with only the latter being binding, while the resource constraint is

$$f_1(\hat{\ell}_{11}, \hat{\ell}_{21})x_1 + f_2(\hat{\ell}_{12}, \hat{\ell}_{22})x_2 - n_1\hat{c}_1 - n_2\hat{c}_2 = 0. \tag{15}$$

In a second-best optimum with redistribution favoring the low-productivity workers, the government chooses the nonlinear income tax to maximize $U(c_1, h_1)$ for a given level of $U(c_2, h_2)$ subject to the binding constraints (2) and (15).

After the introduction of the graduated minimum wage, with heterogenous firms the incentive-compatibility constraints that the low- and high-productivity workers cannot improve their utility by choosing working hours that yield the income of the other type of worker are generalized to

$$U(c_{1j}^*, h_{1j}^*) \geq U\left[c_2^*, \frac{w_2^* h_2^*}{m(\ell_{1j}^*)}\right] \quad \text{for } j \in \{1, 2\}, \tag{16}$$

$$U(c_2^*, h_2^*) \geq U\left[c_{1j}^*, \frac{m(\ell_{1j}^*)h_{1j}^*}{w_2^*}\right] \quad \text{for } j \in \{1, 2\}. \tag{17}$$

Further, the firms' choices of working hours for the low- and high-productivity workers must satisfy the minimum wage constraint

$$f_j(\ell_{1j}^*, \ell_{2j}^*) - m(\ell_{1j}^*)\ell_{1j}^* - w_2^*\ell_{2j}^* \geq \max_{\ell_{1j} \neq \ell_{1j}^*, \ell_{2j}} [f_j(\ell_{1j}, \ell_{2j}) - m(\ell_{1j})\ell_{1j} - w_2^*\ell_{2j}] \text{ for } j \in \{1, 2\}. \tag{18}$$

Since it is possible for a firm to set $\ell_{1j} = \ell_{2j} = 0$ and obtain zero profit, for $j \in \{1, 2\}$ there exist labor inputs $(\ell_{1j}^*, \ell_{2j}^*)$ which yield a nonnegative profit and satisfy the minimum wage constraint.

In the presence of heterogenous firms, however, the graduated minimum wage introduces an additional incentive-compatibility constraint for the low-productivity workers. The intended consumption-work bundles of the low-productivity workers employed in the different categories of firms must be such that those employed in one category of firms cannot make themselves better off by earning the same income as those employed in the other category.[7] Thus, the additional incentive-compatibility constraint for low-productivity workers is

$$U(c_{1j}^*, h_{1j}^*) \geq U\left[c_{1j'}^*, \frac{m(\ell_{1j'}^*)h_{1j'}^*}{m(\ell_{1j}^*)}\right] \text{ for } j \neq j'. \tag{19}$$

The resource constraint is

$$f_1(\ell_{11}^*, \ell_{21}^*)x_1 + f_2(\ell_{12}^*, \ell_{22}^*)x_2 - n_{11}c_{11}^* - n_{12}c_{12}^* - n_2c_2^* = 0. \tag{20}$$

The government's choice of a nonlinear income tax and a graduated minimum wage must satisfy the low-productivity workers' incentive-compatibility constraints (16) and (19), the high-productivity workers' incentive-compatibility constraint (17), the firms' minimum wage constraint (18), and the resource constraint (20).

We now extend Proposition 1 to the case of heterogenous firms:

Proposition 3: *With heterogenous firms there exists a graduated minimum wage which can provide a Pareto improvement of a second-best optimum with nonlinear income taxation that redistributes consumption toward the less productive workers.*

Proof: Following the same logic as in the proof of Proposition 1, we use three steps to construct an equilibrium with a graduated minimum wage that increases the utility of all workers.

[7] All low-productivity workers prefer to work for the firms that pay the higher minimum wage. However, while some workers have the good luck to be employed in those firms, the other workers have no choice but to resign themselves to work for the lower paying firms.

Step 1: Suppose the graduated minimum wage schedule is

$$m(\ell_{1j}) = \begin{cases} \hat{w}_1 + \sigma & \text{for } \ell_{1j} \in \{\hat{\ell}_{11}, \hat{\ell}_{12}\}, \\ \hat{w}_2 & \text{for } \ell_{1j} \notin \{\hat{\ell}_{11}, \hat{\ell}_{12}\}, \end{cases}$$

where $0 < \sigma < \hat{w}_2 - \hat{w}_1$ and σ is sufficiently small that a firm in category j chooses $\hat{\ell}_{1j}$ of low-productivity workers.

In order for the low- and high-productivity workers' intended consumption-work bundles to be unchanged, let the low-productivity workers' tax increase by $\sigma\hat{h}_1$ and the high-productivity workers' tax be unchanged. Then, the low- and high-productivity workers' incentive-compatibility constraints (16) and (17) will be slack, and since the minimum wage is the same in all firms, the low-productivity workers' incentive-compatibility constraint (19) is satisfied. The firms' minimum wage constraint (18) will be slack and the resource constraint (20) is satisfied. As a result, we have an equilibrium with a graduated minimum wage in which both the incentive-compatibility constraints that one type of worker does not want to mimic the other type and the minimum wage constraint are slack. In this equilibrium, the utilities of all the workers are the same as in a second-best optimum with nonlinear income taxes and redistribution toward the low-productivity workers.

Step 2: Recall that the low-productivity workers' marginal rate of substitution between consumption and leisure is less than their marginal product in a second-best equilibrium with nonlinear income taxes and redistribution toward the low-productivity workers. Therefore, letting the total demand for high-productivity workers be kept unchanged by adjusting their wage, for sufficiently small positive μ_j's there is a graduated minimum wage with $m(\hat{\ell}_{1j}+\mu_j) > \hat{w}_1+\sigma$ such that firms in category j prefer to employ $\hat{\ell}_{1j}+\mu_j$ of low-productivity workers.

Let the income tax be such that the low-productivity workers in both firm categories get to consume all their additional production. The low-productivity workers will then be better off with their new consumption-work bundles. Furthermore, let the income tax provide the high-productivity workers with the same consumption-work bundle as before. Accordingly, there is an equilibrium with a graduated minimum wage in which the low- and high-productivity workers' incentive-compatibility constraints (16) and (17) as well as the firms' minimum wage constraint (18) are slack, and the constraints (19) and (20) hold. In the equilibrium, all the low-productivity workers obtain a higher utility and the high-productivity workers obtain the same utility as in the a second-best equilibrium without a graduated minimum wage.

Step 3: Keeping the low-productivity workers' intended minimum wage and working hours in the different categories of firms as well as the high-productivity workers' wage and working hours the same as in step two, it is feasible to transfer some of the low-productivity workers' consumption to the high-productivity workers without violating any of the constraints (16)-(20). Consequently, there exists a graduated minimum wage which can provide a Pareto improvement. □

The Pareto-improving graduated minimum wage distinguishes between different categories of firms similarly to how a nonlinear income tax distinguishes between different types of workers. Consequently, the minimum wage is not the same for all workers, and the graduated minimum wage therefore benefits some of the low-productivity workers more than others.

6 Conclusion

This paper has shown that a graduated minimum wage – in contrast to a constant minimum wage – can provide a Pareto improvement even if the economy is initially in a second-best optimum with a nonlinear income tax that redistributes toward the low-productivity workers. The explanation is that a graduated minimum wage allows the less productive workers to benefit from second-degree price discrimination which increases their wage and income simultaneously. This reduces the attractiveness of trying to mimic their income by the more productive workers whose incentive-compatibility constraint is thereby loosened. As a result, a minimum wage policy can be justified without appealing to a particular social welfare function.

Our model is a first step in studying the merits of a graduated minimum wage and further research is needed to generalize the framework. It would be of interest to investigate how, for different social welfare functions, a graduated minimum wage interacts with the optimal income tax in affecting the consumption and working hours of the different types of workers in a social optimum. Likewise, it would be worthwhile to determine how a graduated minimum wage in a social optimum depends on the production technologies and the supply of the different types of workers.

Appendix

The relevant Lagrangian is[8]

$$\mathcal{L} = U(c_1, h_1) + \lambda_1 U(c_2, h_2) + \lambda_2 \left[f(\ell_1, \ell_2) - n_1 c_1 - n_2 c_2\right] + \lambda_3 \left[U(c_2, h_2) - U\left(c_1, \frac{w_1 h_1}{w_2}\right)\right],$$

where $\lambda_1,\ \lambda_2,\ \lambda_3 > 0$. The first-order conditions for a second-best optimum with nonlinear taxes include

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial c_1} &= \frac{\partial U(c_1, h_1)}{\partial c_1} - \lambda_2 n_1 - \lambda_3 \frac{\partial U\left(c_1, w_1 h_1 / w_2\right)}{\partial c_1} = 0, \\
\frac{\partial \mathcal{L}}{\partial h_1} &= \frac{\partial U(c_1, h_1)}{\partial h_1} + \lambda_2 n_1 f_{\ell_1} - \lambda_3 \frac{\partial U\left(c_1, w_1 h_1 / w_2\right)}{\partial h_2} \left[\frac{w_1}{w_2} + h_1 \frac{\partial\left(w_1/w_2\right)}{\partial h_1}\right] = 0.
\end{aligned}$$

Accordingly,

$$\begin{aligned}
\frac{\partial U(c_1, h_1)}{\partial c_1} &= \lambda_2 n_1 + \lambda_3 \frac{\partial U\left(c_1, w_1 h_1 / w_2\right)}{\partial c_1}, \\
\frac{\partial U(c_1, h_1)}{\partial h_1} &= -\lambda_2 n_1 f_{\ell_1} + \lambda_3 \frac{\partial U\left(c_1, w_1 h_1 / w_2\right)}{\partial h_2} \left[\frac{w_1}{w_2} + h_1 \frac{\partial\left(w_1/w_2\right)}{\partial h_1}\right],
\end{aligned}$$

and hence

$$-\frac{\partial U(c_1, h_1)/\partial h_1}{\partial U(c_1, h_1)/\partial c_1} = \frac{\lambda_2 n_1 f_{\ell_1} - \lambda_3 \left[\partial U\left(c_1, w_1 h_1 / w_2\right)/\partial h_1\right]\left[w_1/w_2 + h_1 \partial\left(w_1/w_2\right)/\partial h_1\right]}{\lambda_2 n_1 + \lambda_3 \partial U\left(c_1, w_1 h_1 / w_2\right)/\partial c_1}. \tag{A1}$$

Using that $w_1 = f_{\ell_1}$ and $w_2 = f_{\ell_2}$ we obtain that

$$\frac{\partial\left(w_1/w_2\right)}{\partial h_1} = \frac{\partial\left(f_{\ell_1}/f_{\ell_2}\right)}{\partial h_1} = \frac{\left(f_{\ell_1 \ell_1} f_{\ell_2} - f_{\ell_1} f_{\ell_2 \ell_1}\right) n_1}{f_{\ell_2}^2} < 0.$$

Hence, (A1) implies that

$$-\frac{\partial U(c_1, h_1)/\partial h_1}{\partial U(c_1, h_1)/\partial c_1} < \frac{\lambda_2 n_1 f_{\ell_1} - \lambda_3 \left[\partial U\left(c_1, w_1 h_1 / w_2\right)/\partial h_1\right] f_{\ell_1}/f_{\ell_2}}{\lambda_2 n_1 + \lambda_3 \partial U\left(c_1, w_1 h_1 / w_2\right)/\partial c_1},$$

which may be written as

$$\frac{\theta}{f_{\ell_1}} < \frac{1 + \psi\phi/f_{\ell_2}}{1 + \psi}, \tag{A2}$$

[8] The proof is based on Stiglitz (1982).

where

$$
\begin{aligned}
\theta &\equiv -\frac{\partial U(c_1, h_1)/\partial h_1}{\partial U(c_1, h_1)/\partial c_1}, \\
\phi &\equiv \frac{\partial U\left(c_1, w_1 h_1/w_2\right)/\partial h_1}{\partial U\left(c_1, w_1 h_1/w_2\right)/\partial c_1}, \\
\psi &\equiv \frac{\partial U\left(c_1, w_1 h_1/w_2\right)}{\partial c_1}\frac{\lambda_3}{\lambda_2 n_1}.
\end{aligned}
$$

Since the agent monotonicity property implies that $\theta/f_{\ell_1} > \phi/f_{\ell_2}$, it follows from (A2) that $\theta < f_{\ell_1}$. That is, the low-productivity workers' marginal rate of substitution between consumption and leisure is less than their marginal product.

References

Aghion, P., Algan, Y., Cahuc, P., 2011. Civil society and the state: The interplay between cooperation and minimum wage regulation. Journal of the European Economic Association 9, 3-42.

Allen, S.P., 1987. Taxes, redistribution, and the minimum wage: A theoretical analysis. Quarterly Journal of Economics 102, 477-489.

Ashenfelter, O., Smith, R.S., 1979. Compliance with the minimum wage law. Journal of Political Economy 87, 333–350.

Basu, A.K., Chau, N.H., Kanbur, R., 2010. Turning a blind eye: Costly enforcement, credible commitment and minimum wage laws. Economic Journal 120, 244-269.

Belot, M., Boone, J., van Ours, J., 2007. Welfare-improving employment protection. Economica 74, 381–396.

Bhaskar, V., Manning, A., To, T., 2002. Oligopsony and monopsonistic competition in labor markets. Journal of Economic Perspectives 16, 155-174.

Bhaskar, V., To, T., 1999. Minimum wages for Ronald McDonald monopsonies: A theory of monopsonistic competition. The Economic Journal 109, 190–203.

Blumkin, T., Sadka, E., 2005. Income taxation and wage policy: An application to minimum wage. International Tax and Public Finance 12, 713-722.

Boadway, R., Cuff, K., 2001. A minimum wage can be welfare-improving and employment-enhancing. European Economic Review 45, 553-576.

Boeri, T., 2012. Setting the minimum wage. Labour Economics 19, 281-290.

Boeri, T., Burda, M.C., 2009. Preferences for collective versus individualised wage setting. Economic Journal 119, 1440-1463.

Brown, W., 2009. The process of fixing the British national minimum wage, 1997-2007. British Journal of Industrial Relations 47, 429-443.

Cahuc, P., Laroque, G., in press. Optimal taxation and monopsonistic labor market: Does monopsony justify the minimum wage? Journal of Public Economic Theory.

Cahuc, P., Zylberberg, A., Saint-Martin, A., 2001. The consequences of the minimum wage when other wages are bargained over. European Economic Review 45, 337–352.

Checchi, D., and Garcia-Penalosa, C., 2010. Labour market institutions and the personal distribution of income in the OECD. Economica 77, 413–450.

Danziger, L., 2009. The elasticity of labor demand and the minimum wage. Journal of Population Economics 22, 757-772.

Danziger, L., 2010. Endogenous monopsony and the perverse effect of the minimum wage in small firms. Labour Economics 17, 224-229.

De Fraja, G., 1999. Minimum wage legislation, productivity and employment. Economica 66, 473–488.

Flinn, C., 2006. Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates. Econometrica 74, 1013-1062.

Guesnerie, R., Roberts, K., 1987. Minimum wage legislation as a second best policy. European Economic Review 31, 490-498.

Hungerbühler, M., Lehmann, E., 2009. On the optimality of a minimum wage: New insights from optimal tax theory. Journal of Public Economics 93, 464-481.

Lee, D., Saez, E., 2012. Optimal minimum wage policy in competitive labor markets. Journal of Public Economics 96, 739-749.

Manning, A., 2003. Monopsony in motion: Imperfect competition in labor markets. Princeton University Press, New Jersey.

Marceau, N., Boadway, R., 1994. Minimum wage legislation and unemployment insurance as instruments for redistribution. Scandinavian Journal of Economics 96, 67-81.

Rebitzer, J.B., Taylor, L.J., 1995. The consequences of minimum wage laws: Some new theoretical ideas. Journal of Public Economics 56, 245–255.

Robinson, J., 1933. The economics of imperfect competition. London: MacMillan.

Salanié, B., 2003. The economics of taxation. Cambridge: MIT Press.

Sobel, R.S., 1999. Theory and evidence on the political economy of the minimum wage.

Journal of Political Economy107, 761-785.

Stewart, M.B., Swaffield, J.K., 2008. The other margin: Do minimum wages cause working hours adjustments for low-wage workers? Economica 75, 148–167.

Stiglitz, J.E., 1982. Self-selection and Pareto efficient taxation. Journal of Public Economics 17, 213-240.

Strobl, E., Walsh, F., 2011. The ambiguous effect of minimum wages on hours. Labour Economics 18, 218-228.

Weil, D., 2005. Public enforcement/private monitoring: Evaluating a new approach to regulating the minimum wage. Industrial and Labor Relations Review 58, 238–257.