



Queen's Economics Department Working Paper No. 980

A Minimum Wage can be Welfare-Improving and Employment-Enhancing

Robin Boadway
Queen's University

Katherine Cuff
Queen's University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

2-1999

A MINIMUM WAGE CAN BE WELFARE-IMPROVING AND EMPLOYMENT-ENHANCING

by

Robin Boadway, *Queen's University*

Katherine Cuff, *Queen's University*

January 1999

ABSTRACT

We examine whether minimum wages can fulfill a useful role as part of an optimal non-linear income tax scheme. In this setting, governments cannot observe household abilities, only their incomes. Redistributing according to income, the government is constrained by a set of incentive constraints. Firms, on the other hand, are able to identify abilities of workers. To exploit that, the government imposes a minimum wage. This will preclude firms from offering a job to anyone below the minimum wage. The use of the minimum wage policy combined with the institutional features of typical welfare systems allows the incentive constraints to be severed at the ability level associated with the minimum wage. If such a scheme can be enforced, the government can increase the amount of redistribution from those working to those not working. Moreover, the optimal minimum wage may actually lower the number of unemployed.

KEY WORDS: minimum wage, optimal income tax, unemployment

JEL CLASSIFICATION: H21, H23, J38

We have benefitted from discussions with Patrick Francois, Nicolas Marceau, Maurice Marchand, Steeve Mongrain, and Russell Davidson. Financial support from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

A MINIMUM WAGE CAN BE WELFARE-IMPROVING AND EMPLOYMENT-ENHANCING

‘A prerequisite for an increase in the minimum wage to raise the incomes of poor families...is that minimum wages redistribute earnings toward low-wage workers.’ One factor to consider is that ‘changes in labour income may affect government transfers received by the family. In general, those family experiencing job loss as a result of a minimum wage hike may receive an increase in government assistance.’ (OECD, 1998, p. 49)

I. Introduction

In North America, minimum wages are a basic component of government redistributive policy, and represent a significant portion of average earnings. In 1997, the minimum wage in the United States was 38% of median earnings for all workers and in Canada it was 39%. Currently, the British government is in the process of introducing a statutory minimum wage. As with any redistributive policy, the objective is to increase the well-being of some segment of the population. In the case of the minimum wage, it is to increase the welfare of low-income earners by increasing their wages. However, the jury is still out on whether or not this form of price intervention actually makes individuals better off.

Traditionally, economists would argue that by interfering with the immutable laws of supply and demand, minimum wages are counterproductive: by creating excess supply in the low-wage labour market, some jobs of those persons whom the policy was intended to benefit are destroyed.¹ Neoclassical economics does not support the use of the minimum wage as an efficient redistributive policy. How then can economists explain the legislation of minimum wages by governments?

Economists have investigated the use of minimum wage policy by relaxing the three main assumptions of the neoclassical argument, all of which are restrictive and unrealistic. They are as follows: labour markets are perfectly competitive, workers are homogeneous and there exists perfect information in the economy. Under these assumptions, any reduction in employment necessarily results in welfare losses. Allowing for monopsonistic

¹ The classic argument is by Stigler (1946). For an early statement of the case against minimum wages, see Pigou (1920).

labour markets gives a more realistic portrayal of the labour market and suggests a role for minimum wage policy.² Profit-maximizing firms with wage-setting power pay wages less than workers' marginal productivities. Imposition of a minimum wage (above the current wage and below the market-clearing wage) increases employment and welfare of both the currently employed and the new employees. Minimum wage policy may be optimal in this setting. The second approach is to allow for heterogeneous labour. Under such an assumption, it is possible that reductions in a given minimum wage can actually reduce growth and possibly decrease the welfare of individuals in the economy (Cahuc and Michel, 1996).³ The last approach is to allow for different forms of imperfect information. In efficiency wage models, employers are assumed to be unable to observe worker effort, but can imperfectly monitor it. Using the threat of dismissal, employers ensure workers do not shirk by paying them a wage above the market-clearing wage. A minimum wage increases the cost of shirking and releases resources which may be used to hire additional workers (Rebitzer and Taylor, 1995). In the short-run, in which the number of firms is fixed, the minimum wage increases welfare as those working earn a higher wage and there are more individuals employed. This effect might be mitigated in the long-run with the exit of firms.

All of the above arguments exhibit a positive correlation between the minimum wage's effect on employment and its effect on welfare. Alternatively, if it is assumed that workers possess imperfect information about job opportunities and must exert effort to find work, this correlation no longer necessarily holds and whether or not the minimum wage is an efficient policy is inconclusive. In a unilateral search model with homogeneous workers, the minimum wage can have ambiguous employment effects, but still be welfare-improving (Swinerton, 1996).⁴ In a bilateral search model with heterogeneous workers, a minimum

² See Chapter 12, Gunderson and Riddell (1988) for a formal exposition of a minimum wage in a monopsony setting.

³ Using an endogenous growth model, Cahuc and Michel (1995) show a minimum wage induces unemployment in the unskilled labour market and, thus, increases the incentive of unskilled workers to invest in human capital.

⁴ Swinerton (1996) assumes that firms are competitive, exhibit diminishing returns to labour, and differ in their productivities which suggests some highly productive firms may be labour-supply constrained. Imposing a minimum wage induces a reallocation of workers from unconstrained firms (low productivity) to constrained firms (high productivity) and therefore,

wage is employment-enhancing but not welfare-improving (Lang and Kahn, 1998).⁵

Evidently, the relaxation of the unrealistic assumptions of the neoclassical model has not given rise to a consensus of the effect of minimum wages on welfare or employment. Likewise there remains debate as to the actual effects of minimum wage legislation. Empirically, the focus of minimum wage research has been on its employment effects, particularly for young workers, who as a group have higher unemployment rates than adult workers.⁶ Mixed evidence has spark intense debate about empirically specifications and data selection from which no general consensus has emerged.⁷

The focus of this paper is the more fundamental normative issue of whether minimum wages can fulfill any useful role as part of the menu of optimal redistribution policies. The impetus for re-visiting the issue derives from a pervasive finding in second-best theory: policies which rely on quantity controls can often be found to be welfare-improving in second-best settings. In the seminal paper, Guesnerie and Roberts (1984) show that in an economy with given distortions between price and marginal cost, either in-kind transfers or forced consumption of particular goods can be welfare-improving. They suggest that minimum wages which cause low-income individuals to reduce their hours work (under-employment) might be one such policy (Guesnerie and Roberts, 1987). However, their analysis is limited to linear distortions, a setting in which tax policy is sub-optimal.

The obvious next question is: can quantity controls still be optimal when the government imposes optimal non-linear taxes? An appropriate framework is the optimal non-linear income taxation model of Mirrlees (1971, 1974). In this setting, governments cannot observe household abilities, only their incomes. Redistributing according to income runs up against a set of incentive constraints: higher-ability persons might prefer to pre-

is welfare-improving.

⁵ A minimum wage induces higher quality applicants to move to low-wage jobs which exhibit greater vacancies. However, increased competition for higher-ability individuals makes lower-ability individuals worse-off and the higher-ability individuals are not made any better off.

⁶ For an extensive survey of recent empirical studies on the effect of the minimum wage on employment and on earnings and income distributions, see OECD (1998).

⁷ See Card and Krueger (1995), who provide evidence of zero or positive employment effects, and both Burkhauser, Couch and Wittenburg (1998) and Neumark and West (1996) who challenge their results.

tend to be of lower-ability. By earning the same income as the lower-ability individual, the high-ability individual is observed to be lower-ability by the government. Faced with this informational constraint, governments must use non-linear income taxes to efficiently separate higher from lower-ability persons. Surprisingly, even in the presence of optimal non-linear income taxes, quantity controls, such as in-kind transfers and mandated consumption, can be a useful supplementary policy (Boadway and Marchand, 1995).⁸ Such quantity controls work by relaxing the incentive constraints, and thereby, expanding the second-best utility possibilities frontier.

This result was exploited in the context of minimum wage models by Allen (1987) and Marceau and Boadway (1998). Extending Guesnerie and Roberts' (1987) model into a non-linear income tax setting and allowing for general equilibrium effects on wages, Allen (1987) argues that the potential usefulness of a minimum wage which induces under-employment no longer holds. Using a similar setting, Marceau and Boadway (1998) assume the minimum wage causes involuntary unemployment, rather than under-employment, and reach the opposite conclusion. A minimum wage supplemented by unemployment insurance can be welfare-improving.

This paper takes the possibility of relaxing the incentive constraints further. We argue that the use of a minimum wage combined with the institutional features of typical welfare systems allows the incentive constraints to be severed at the ability level associated with the minimum wage. This has dramatic consequences which, to the extent that they are true, put a somewhat different light on minimum wage policy. We argue that not only can much higher transfers be made to those unemployed, but the amount of unemployment can actually be *lower* than it would be without the minimum wage legislation. The intuition is as follows. Begin with an optimal income tax allocation of the sort first portrayed by Mirrlees (1971). For a range of ability-types at the bottom of the distribution, labour supply and thus income can be zero. The level of transfers going to the idle is constrained by an incentive constraint required that no one working must prefer to mimic the unemployed. Now, recall that in this setting, firms must be able to identify abilities of workers, even though the government cannot. To exploit that, the government imposes a minimum

⁸ Moreover, such policies can dominate policies implemented through the price system, such as subsidies on the same goods, as shown in Boadway, Marchand and Sato (1998).

wage. This will preclude firms from offering a job to anyone below the minimum wage. Combine the minimum wage policy with a requirement that to be eligible for the transfer, any job offered must be accepted, a requirement which is a feature of typical welfare systems. If such a scheme can be enforced, it will succeed in nullifying the incentive constraint applying to workers at the minimum wage. The government can now increase the amount of redistribution from those working to those not working. Moreover, as we show, the optimal minimum wage may actually be lower than the wage at the top end of the bunching interval. That is, the number who are unemployed may fall.

An implication of combining the minimum wage with the requirement that job offers must be accepted to be eligible for welfare or unemployment benefits is that those on receiving benefits are better off than those working.⁹ This seems to be a relevant feature of most welfare and unemployment insurance systems. If it were not, it would not be necessary to devote the considerable resources that are devoted to monitoring transfer recipients to ensure that they are actively seeking employment and accepting any job offers that come along.

The model we use to illustrate these possibilities is based on a simplified version of the standard Mirrlees optimal income tax model, a version which admits of an explicit solution. In particular, we assume that household preferences are quasi-linear in leisure, a case first analyzed by Lollivier and Rochet (1983) and subsequently by Weymark (1986) and Ebert (1992). This turns out to be a very convenient model for our purpose, since not only does it yield an explicit and intuitive solution which can easily be calculated, but it also allows us readily to compare optimal policy regimes with and without minimum wages.

To understand better the forces at work, we begin in the next section with an even simpler case — that in which household labour supplies are fixed. Though the fixed labour supply assumption strains credulity, this case allows us to present the nature of our argument, and to illustrate the importance of the informational assumptions, in the simplest possible way. We then turn to the variable labour supply case in Section III. This

⁹ We are assuming that all households have the same preferences as in the standard optimal taxation framework. Actual welfare systems distinguish between employable persons and the disabled who might have different preferences. Thus, our analysis can be interpreted as applying to the employable segment of the population.

involves first setting out the non-linear income tax problem for the quasi-linear preference case in some detail, and then using it to derive the effects on the tax-transfer structure and employment from imposing the minimum wage when abilities are private information. The qualitative effects on employment are ambiguous in both the fixed and variable labour supply cases, so numerical examples are presented to show how optimal (un)employment can either rise or fall with minimum wage legislation. In the final section, we discuss our results and some of the limiting assumptions, especially the informational ones, that we have adopted.

II. An Expository Case: Fixed Labour Supply

Individual preferences are represented by $u(c) - d$, where c is a composite consumption good, d is the disutility individuals receive from working, and $u(\cdot)$ is increasing and strictly concave. Each individual who works supplies a given amount of labour and receives a fixed amount of income y , which varies according to the individual's productive ability. Productive ability is assumed to be perfectly observable to all firms, so that y also reflects the output an individual produces. We are implicitly assuming that both firms and the labour market are perfectly competitive and there exists a sufficient number of jobs for all individuals to be employed. If an individual decides to go into the labour market, then they are necessarily offered a job that pays them their marginal productivity conditional on providing a fixed amount of labour to their employer. Therefore, all unemployment is voluntary in this setting. It is further assumed that y is distributed over the interval $[\underline{y}, \bar{y}]$, with $\underline{y} > 0$ and $\bar{y} < \infty$, and with a density of $f(y)$ and a cumulative distribution function $F(y)$. The population size is normalized to unity, so $F(\bar{y}) = 1$. The individual's budget constraint is $c = y - t(y)$, where $t(y)$ is the tax-transfer schedule.

If there were no tax-transfer system in place, so $t(y) = 0$ for all y , then individuals consume what they earn. They would chose to work only if the additional utility they receive from consuming a positive amount of consumption is at least as great as the disutility they receive from working. Since all individuals receive the same disutility from working there may exist some cut-off level of income, such that all individuals who earn more than this income level work and all individuals who earn less do not work. Assuming that this cut-off is in the interior of the distribution, the marginal individual working is defined as the individual who earns the cut-off level of income.

The feasibility and optimality of the various forms of redistributive policy in this model depend critically on the information available to the government. To develop our argument in a way which is both intuitive and illuminates the more general Mirrlees setting considered in the next section, we consider a sequence of three cases here. The first, which serves as a benchmark, is that of full information. The government, like firms, can observe individuals abilities (and thus income). A first-best outcome can be achieved by a lump-sum tax-transfer system.

In the second case, the government can observe incomes of those who work, but cannot observe what the incomes would be of those who are not working, i.e. potential incomes. This implies that redistribution is limited by an incentive constraint that entails that those working can be no worse off than if they were not working. In this context, even though the government can make optimal lump-sum transfers among the working, it cannot transfer as much to those not working as it might like. In this context, a minimum wage will be welfare-improving. But, it will necessarily increase the number of persons not working.

The third case, which most closely parallels the informational assumptions of Section III, assumes that the government cannot observe even the incomes earned by those working. All it can observe is whether or not a household is working, so it must condition its tax-transfer system only on that. Obviously, social welfare in this case is less than in the previous two cases. As in the second case, a minimum wage will be unambiguously welfare-improving. Moreover, it may entail a reduction in the number of persons working, a result which anticipates the one obtained in the more general Mirrlees model considered in the next section.

The minimum wage legislation is independent of the informational assumptions discussed above and has two components. First, all firms must pay wages that are greater or equal to the stated minimum wage. Since firms can observe individuals' abilities, they only offer jobs to individuals with a productivity greater or equal to the minimum wage, or equivalently in this fixed labour supply setting, only offer jobs to individuals whose outputs are at least, say, y_m . Second, all individuals must accept a job if they are offered one. This is a common feature of most welfare systems. To qualify for benefits, individuals must conduct job searches and accept any jobs that are offered. It is assumed that both components of the legislation are perfectly enforceable and that enforcement and monitor-

ing costs are zero. This is obviously an extreme assumption. In the real world, there are a variety of enforcement problems. As in the standard redistributive tax literature, there will be incentives to misrepresent incomes to evade taxes. In addition, in our setting firms will have an incentive to hire workers illegally at wages less than the minimum wage. And, transfer recipients will have an incentive not to look for jobs or accept job offers. Any of these evasive activities will reduce the efficiency of redistributive policy. By assuming costless enforcement, we are able to present our argument in the clearest way.¹⁰

Case 1: Full Information

Consider then the case where the government can perfectly identify individuals' abilities, i.e. actual and potential incomes. Assume the government chooses the tax-transfer schedule by maximizing a utilitarian social welfare function¹¹ subject to its revenue constraint, which is given by:

$$\int_{y_f}^{\bar{y}} t(y)f(y)dy - \int_{\underline{y}}^{y_f} T(y)f(y)dy \geq R$$

where R is the government's revenue requirement, y_f is the cut-off level of income/ability, and $T(y) > 0$ is the transfer schedule for the non-working individuals. The government offers $T(y)$ only to individuals who would earn income $y \in [\underline{y}, y_f)$. All individuals who earn income $y \in (y_f, \bar{y}]$ are subject to the tax-transfer schedule $t(y)$. Given that the government can observe abilities, it will choose y_f such that $u(y_f) - d = u(0)$: the working segment of the population is the same as in the laissez-faire case. However, non-working individuals are better off in this case as the government redistributes income. From the choice of social welfare function, it follows that the government will choose the tax-transfer schedules to equalize individuals' marginal utility of consumption, $u'(T(y)) = u'(y - t(y))$. From the strict concavity assumption on $u(\cdot)$, this implies that optimally all individuals receive the same consumption, which from the binding budget constraint, and recalling

¹⁰ Adding costly enforcement of evasion and other forms of illegal behaviour will complicate the analysis considerably, but the same qualitative results will apply.

¹¹ Throughout we assume the objective function to be the sum of utilities. This is just for illustrative purposes. Other objective functions exhibiting aversion to inequality will serve as well, except one with a complete aversion to inequality, i.e. maxi-min. As shown below, introduction of a minimum wage reduces the utility of some working individuals below the utility level of the initially 'worst-off' individuals and thus, precludes the use of a maxi-min function as a welfarist objective.

$F(\bar{y}) = 1$, is $c_f = \int_{\underline{y}_f}^{\bar{y}} yf(y)dy - R$. The tax-transfer schedule, $t(y)$, is increasing in earned income with a 100% marginal tax rate, and the transfer schedule, $T(y)$, is equal to c_f for all $y \in [\underline{y}, y_f)$. Optimally, the government perfectly “socially insures” the consumption of individuals against the possibility of zero income. As a result, non-working individuals are better off than working individuals.¹²

Case 2: Unobservable Potential Incomes

In this case, it is assumed that the government cannot observe the income an unemployed individual would earn if they were working. Thus, the informationally constrained government must give all non-working individuals the same transfer, T . But earned income is observable, so the government can condition the tax-transfer schedule for those working on their income, $t(y)$. These informational assumptions imply that working individuals have the option of mimicking non-working individuals to avoid the disutility from working. The government must also satisfy incentive-compatibility constraints to ensure that working individuals are at least as well-off as if they were not working. Individuals of a given ability cannot mimic the earned income of any individual of a different ability since firms can perfectly identify working individuals’ abilities.

The working segment of the population is determined endogenously by the government’s choice of T and $t(y)$. Given the tax-transfer schedule $(T, t(y))$, individuals work only if they receive a lower utility from not working. The government must satisfy the participation constraint for all individuals working. Define the cut-off level of earned income y_p such that all individuals who work earn $y \geq y_p$ (their participation constraint is satisfied and binding since $t(y)$ is variable) and all individuals who would earn $y < y_p$ do not work (their participation constraints are violated). The problem is set up such that the government chooses y_p as an artificial variable subject to the participation constraints.¹³ The Lagrangian of the government’s problem is then given by:

$$\max_{T, t(y), y_p} \mathcal{L} = \int_{\underline{y}}^{y_p} u(T) f(y) dy + \int_{y_p}^{\bar{y}} [u(y - t(y)) - d] f(y) dy$$

¹² This parallels the well-known result in the variable labour supply model that when optimal lump-sum taxes are available and the objective function is utilitarian, utility will be decreasing in ability (Mirrlees, 1974).

¹³ Equivalently, the participation constraints can be suppressed by treating y_p as a function of T and $t(y)$ as determined implicitly by the constraints.

$$\begin{aligned}
& + \lambda \left(\int_{y_p}^{\bar{y}} t(y)f(y)dy - \int_{\underline{y}}^{y_p} Tf(y)dy - R \right) \\
& + \int_{y_p}^{\bar{y}} \phi(y)[u(y - t(y)) - d - u(T)]dy
\end{aligned}$$

where λ is the marginal cost of public funds in terms of social welfare and $\phi(y)$ is the multiplier on the participation constraint of the individual earning y . The first-order conditions are as follows:

$$T : \quad u'(T)F(y_p) - \lambda F(y_p) - \int_{y_p}^{\bar{y}} \phi(y)u'(T)dy = 0$$

$$t(y) : \quad -u'(y - t(y))f(y) + \lambda f(y) - \phi(y)u'(y - t(y)) = 0$$

$$y_p : \quad u(T) - [u(y_p - t(y_p)) - d] - \lambda[T + t(y_p)] - \phi(y_p)[u(y_p - t(y_p)) - d - u(T)] = 0$$

Denote the solution to this problem by $t^p(y)$, T^p and y_p . In equilibrium, all the participation constraints bind, so $u(y - t^p(y)) - d = u(T^p)$ for all $y \in [y_p, \bar{y}]$. All individuals receive the same utility, which implies working individuals receive more consumption than non-working individuals to compensate them for their disutility from working. Since this disutility is constant, all working individuals must receive the same consumption and, therefore, the government optimally equalizes the marginal utility of consumption, $u'(y - t(y))$, across the working population.

From the first-order condition on $t(y)$, we can solve for the non-negative multiplier on the participation constraints

$$\phi(y) = f(y) \left(\frac{\lambda}{u'(y - t(y))} - 1 \right)$$

In the optimum, this will be positive for all $y \in [y_p, \bar{y}]$. Substitute it into the first-order condition on T to obtain,

$$\lambda = \frac{u'(T)}{F(y_p) + (1 - F(y_p)) \frac{u'(T)}{u'(y - t(y))}}$$

From the expression for $\phi(y)$ and the binding participation constraints, it follows that $u'(y - t(y)) < \lambda < u'(T)$. As well, the first-order condition on y_p reduces to $T^p = -t(y_p) > 0$. The marginal individual working receives a transfer.

The inability of the government to observe potential incomes of those individuals not working prevents it from implementing the first-best redistributive scheme, one in which marginal utilities of consumption are equalized across the entire population. Examining the first-order conditions on T and $t(y)$ indicates that if $\phi(y) = 0$ for all $y \in [y_p, \bar{y}]$, i.e. these constraints were not binding, social welfare could be increased by incremental increases in both T^p and $t^p(y)$, that is, by more redistribution. This motivates us to investigate the use of a minimum wage to relax the incentive constraints applying between those individuals working and those individuals not working.

Minimum Wage Legislation

Recall, that the minimum wage legislation entails both firms offering jobs only to those individuals who produce $y \geq y_m$ and individuals accepting any job that they are offered. We then proceed in two steps. First, we illustrate that a legislated minimum wage increases social welfare when the government imposes it at the minimum income level y_p , the cut-off level when there is no minimum wage. Second, we show that the optimal choice of the minimum income will always be greater than y_p . In other words, the optimal minimum wage legislation will always reduce employment in this case.

Suppose the government imposes a minimum wage such that $y_m = y_p$. Effectively, the minimum wage legislation allows the government to ignore the participation constraints of all individuals working. Suppose an individual of ability greater than y_p was unemployed. To qualify for the unemployment transfer T , they would have to conduct a job search. Given their ability, some firm will offer them a job and they must accept it.¹⁴ The government's problem is identical to the one given above except $\phi(y)$ is equal to zero for all $y \in [y_p, \bar{y}]$ and y_p is fixed. Re-optimizing social welfare, the government selects a different tax-transfer schedule. This implies that social welfare is at least as great as when there was no minimum wage since the government could have chosen the same tax-transfer as it had before. In fact, the government increases both the amount it transfers to the non-working individuals and the amount it taxes the working individuals such that consumption is equalized across the entire population.

To see this, note that when $\phi(y)$ is set to zero for all $y \in [y_p, \bar{y}]$ the right-hand side

¹⁴ Implicitly, it is assumed that the minimum wage is not prohibitive. That is, $u(y - t(y)) - d > u(0)$ for all $y \in [y_m, \bar{y}]$. All working individuals prefer to be working than consuming nothing.

of the first-order conditions on both T and $t(y)$ fall at the original tax-transfer schedule $(T^p, t^p(y))$. Equality is restored at a new tax-transfer schedule $(T^m, t^m(y))$, such that all individuals receive the same consumption, T^m , that is $u'(T^m) = u'(y - t^m(y))$ for all $y \in [y_p, \bar{y}]$. To satisfy the balance budget requirement, this implies T and $t(y)$ both increase from $(T^p, t^p(y))$. At this new optimum, working individuals are worse off (they receive lower consumption) and non-working individuals are better off (they receive higher consumption). In addition, the marginal cost of public funds is equal to the marginal utility of the unemployed individuals, $\lambda = u'(T^m)$.¹⁵

The next question to address is: what is the optimal minimum income? To answer this, we examine the optimal direction of change in y_m starting from $y_m = y_p$. By the Envelope Theorem, differentiating the maximized Lagrangian with respect to y_m evaluated at $y_m = y_p$ gives us the change in social welfare, SW , from a small change in y_m :

$$\frac{dSW}{dy_m} = \left. \frac{\partial \mathcal{L}}{\partial y_m} \right|_{y_m=y_p} = [u(T^m) - u(y_m - t^m(y_m)) + d] - u'(T^m)[t^m(y_m) + T^m]$$

The first term represents the net utility gain from marginally increasing y_m . The second term represents the cost in utility terms (recall, $\lambda = u'(T^m)$) of marginally increasing y_m (the lost tax revenue $t^m(y_m)$ plus the additional transfer T^m). Given that consumption is equalized across the population, the first term reduces to d , that is, the utility the marginal individual receives from no longer working, and the second term reduces to $u'(T^m)y_m$, that is, the social value of the income the marginal individual is no longer earning. From the no minimum wage case, it can be shown that $d > u'(T^p)y_p$.¹⁶ Imposing a minimum wage at $y_m = y_p$, increases the transfer to the unemployed individuals ($T^m > T^p$). From the strict concavity assumption on utility, this implies $u'(T^p)y_p > u'(T^m)y_m$ and the above expression is unambiguously positive. Social welfare is improved if y_m is increased from $y_m = y_p$. More generally, if the government can optimally chose y_m , it will do so by setting the above expression to zero and thus, the optimal minimum wage is such that $y_m = d/u'(T^m) > y_p$.

¹⁵ This follows from the first-order condition on T when $\phi = 0$.

¹⁶ Using the binding participation constraint at $y = y_p$ and the first-order condition on y_p , to obtain $u(y_p + T^p) - u(T^p) = d$. Taking a second-order Taylor expansion of the first term around T^p and subtracting $u(T^p)$ (noting that $u''(\cdot) < 0$) gives this result.

Thus, in this case legislating an optimal minimum wage increases social welfare and at the same time reduces employment from the no minimum wage case. This unambiguous employment effect depends critically on the assumption that earned incomes are observable. If this is no longer true, then the optimal minimum wage can increase employment. We turn to this case next.

Case 3: Unobservable Actual Incomes

Suppose we further restrict the government's information and assume incomes of those working are unobservable. The government is only able to observe an individual's employment status. Thus, all working individuals must be taxed by the same amount, t , and all non-working individuals must receive the same transfer, T . Working individuals have an incentive to mimic non-working individuals to avoid the disutility of labour and as in the previous case, the government is constrained by incentive-compatibility constraints, such that those working cannot be any worse off than those not working. However, since t is a constant, the participation constraint of the marginal individual working will be the only binding, and therefore relevant, participation constraint. Denote the income earned by this individual as y_o and define it such that $u(y_o - t) - d = u(T)$. Thus, it is determined endogenously by the government's choice of t and T and is assumed to be in the interior of the distribution. Given the tax-transfer schedule, individuals work if $u(y - t) - d \geq u(T)$. As the left-hand side of this expression is increasing in y , the participation constraints of all individuals who earn more than y_o are slack and they choose to work. The participation constraints are violated for all individuals who earn less than y_o and they choose not to work.

The problem is again set up such that the government chooses y_o as an artificial variable subject to the participation constraint. The Lagrangian of the government's problem is then given by:

$$\begin{aligned} \max_{T, t, y_o} \mathcal{L} &= \int_{\underline{y}}^{y_o} u(T) f(y) dy + \int_{y_o}^{\bar{y}} [u(y - t) - d] f(y) dy \\ &+ \lambda \left(\int_{y_o}^{\bar{y}} t f(y) dy - \int_{\underline{y}}^{y_o} T f(y) dy - R \right) \\ &+ \phi [u(y_o - t) - d - u(T)] \end{aligned}$$

where ϕ is the multiplier on the participation constraint of the individual who earns y_o .

The first-order conditions are as follows:

$$T : \quad u'(T)F(y_o) + \lambda F(y_o) - \phi u'(T) = 0$$

$$t : \quad - \int_{y_o}^{\bar{y}} u'(y-t)dF(y) + \lambda(1-F(y_o)) - \phi u'(y_o-t) = 0$$

$$y_o : \quad u(T)f(y_o) - [u(y_o-t_w) - d]f(y_o) - \lambda f(y_o)(T+t) + \phi u'(y_o-t) = 0$$

From the first-order condition on T , we obtain $\lambda = u'(T)(1 - \phi/F(y_o)) < u'(T)$. Substituting this expression into the first-order condition on t :

$$\frac{\int_{y_o}^{\bar{y}} u'(y-t)dF(y)}{1-F(y_o)} = u'(T) - \phi \left(\frac{u'(T)}{F(y_o)} + \frac{u'(y_o-t)}{1-F(y_o)} \right)$$

In equilibrium, the participation constraint will be binding, $\phi > 0$. From the above expression, this implies that the average marginal utility of working individuals is less than the marginal utility of non-working individuals. Thus, on average working individuals receive more consumption than individuals not working. Denote the solution to this problem by t^o , T^o and y_o . The path of utility is shown in Figure 1.

—FIGURE 1 NEAR HERE—

The inability of the government to observe incomes is reflected both in the requirement to apply the same tax on all working individuals and in the need to abide by the incentive constraint. The latter requires that the marginal working person have the same utility as those not working, which in turn constrains the choices of T^o , t^o , and y_o . In the absence of that constraint, ϕ would be zero. Examining the first-order conditions on T and t indicates that social welfare could be increased by incremental increases in both T^o and t^o , that is, in more redistribution. At the same time, incrementally *decreasing* y_o starting at (T^o, t^o) , and thus increasing the number employed could increase social welfare: this would increase the tax revenue that the government raises and make available more funds to transfer to the remaining unemployed. This motivates us to investigate the use of a minimum wage to relax the incentive constraint applying between those working and those not working.

Minimum Wage Legislation

As in the previous case, we proceed in two steps. First, we illustrate that a legislated minimum wage increases social welfare when the government imposes it at the minimum income level y_o , the cut-off level when there is no minimum wage. Second, we show that the optimal choice of the minimum income can be greater or less than y_o . This is in stark contrast to the unambiguous negative employment effect obtained in the previous case. That is, the optimal minimum wage in this case can either decrease or increase unemployment.

Suppose the government imposes a minimum wage such that $y_m = y_o$. All individuals who would earn less than y_o will not be offered a job and all individuals who are offered a job must work. Effectively, the minimum wage legislation allows the government to ignore the participation constraint of the marginal individual working. The government's problem is identical to the one given above except ϕ is equal to zero and y_o is fixed. Given the minimum income, the government re-optimizes social welfare and selects a different tax-transfer schedule to the case without a minimum wage: thus, social welfare is at least as great as when there was no minimum wage. In fact, the government optimally increases both the amount it transfers to the non-working individuals and the amount it taxes the working individuals.

To see this, note that legislating the minimum income at y_o , allows the government to ignore the participation constraint, so $\phi = 0$. This implies that the right-hand side of the expression derived from the first-order conditions on T and t given above, rises relative to the left-hand side at the original tax-transfer schedule (T^o, t^o) . Equality is restored at a new tax-transfer schedule (T^m, t^m) , such that $u'(T)$ is equal to the average marginal utility of those individuals working. To satisfy the balance budget requirement, this implies T and t both increase from (T^o, t^o) . This is illustrated in Figure 1. The government optimally transfers more income to the unemployed individuals and imposes higher taxes on working individuals. As a result, the individual who earns $y_o = y_m$ will be worse off when the government imposes the minimum wage. It can also be shown that the marginal cost of public funds with the minimum wage is equal to the marginal utility of the unemployed individuals.¹⁷

¹⁷ This follows from the first-order condition on T when $\phi = 0$.

The next question to address is: what is the optimal minimum income? To answer this, we examine the optimal direction of change in y_m starting from $y_m = y_o$ using the Envelope Theorem. Then we allow the government to explicitly optimize over y_m .

By the Envelope Theorem, we differentiate the maximized Lagrangian to find the change in social welfare, SW , with a small change in y_m :

$$\frac{dSW}{dy_m} = \frac{\partial \mathcal{L}}{\partial y_m} \Big|_{y_m=y_o} = [u(T^m) - u(y_m - t^m) + d] - u'(T^m)[T^m + t^m]$$

The above expression has the same interpretation as in the previous case. The first term represents the net utility gain from marginally increasing y_m and is positive since the working individual earning the minimum wage is worse off than the non-working individuals. Likewise, the second term represents the cost in utility terms of this marginal increase and is also positive since the government imposes a positive tax on earned income and gives a positive transfer to non-working individuals. Therefore, the sign of the expression is ambiguous: the government might want to increase or decrease y_m from $y_m = y_o$. More generally, if the government can optimally choose y_m it will do so by setting the above expression to zero. It is possible that the optimal minimum wage decreases unemployment, $y_m < y_o$, or increases it, $y_m \geq y_o$. The conditions under which $y_m <, > y_o$ are not apparent. In order to illustrate that either is possible, a simple numerical example is given below.

Numerical Illustration

In this example, we adopt a uniform distribution of income with $y \in [1, 10]$ and a quadratic utility function $u(c) = c - bc^2$, $b > 0$. The disutility of labour, d , is assigned the value of 1.98. The results on the cut-off level of income, the tax rate, the level of transfers and social welfare are given in Table 1 and Table 2 for $b = .01$ and $b = .05$, respectively.

Table 1

$b = .01$	$y_{cut-off}$	t	T	SW
Full Information	2.02		5.33	3.29
Asymmetric Information	2.15	.016	.113	3.21
Minimum Wage	2.37	.945	5.24	3.25

Table 2

$b = .05$	$y_{cut-off}$	t	T	SW
Full Information	2.23		5.23	2.17
Asymmetric Information	4.39	.67	1.11	1.81
Minimum Wage	4.34	2.66	4.51	2.16

From the above tables, we see that legislating a minimum wage increases social welfare from the asymmetric information case without a minimum wage and can increase or decrease unemployment. The effect on employment depends on the value of b or the curvature of the utility function, and of course on the information available to the government.¹⁸

The assumptions in this section are obviously strong, especially the presumption that labour supply is fixed. The example was structured so as to illustrate in a simple, intuitive way the nature of the results and the implication of informational assumptions in the more complicated version of the model. In the next section, we follow the conventions of the optimal non-linear taxation literature by allowing individuals to choose how much labour to supply and assuming the government can observe income, but not wages (abilities). We illustrate that the same forces are at work as in this section. Imposing a minimum wage unambiguously increases social welfare because it relaxes the self-selection constraints faced by the government and has an ambiguous effect on employment.

III. Variable Labour Supply

We begin by outlining the government's optimal redistribution problem in the absence of a minimum wage, that is, in the standard optimal income taxation setting. Our choice of a quasi-linear utility function allows us characterize the optimal allocation explicitly, and considerably simplifies the treatment of the minimum wage, which we take up subsequently.

Asymmetric Information with No Minimum Wage

Individuals have the same quasi-linear preferences as in the previous section, but d is replaced by the amount of labour they supply, ℓ , which is also the disutility they receive from working. Individuals differ in their ability to convert labour, ℓ , (in terms of time spent

¹⁸ Note, $u'(c) = 1 - 2bc$ and $u''(c) = -2b$. Increasing b increases the curvature of the individuals' indifference curves.

working) into output, y (in terms of the composite good). The composite consumption good, c , is the numeraire and individuals' abilities (or marginal productivities) are represented by their wage rates, w , as firms can perfectly observe abilities. It is assumed that wage rates in the population are distributed according to the density function $f(w)$ and have a corresponding cumulative distribution function, $F(w)$ for $w \in [\underline{w}, \bar{w}]$, where $\underline{w} > 0$ and $\bar{w} < \infty$. We assume $F(\bar{w}) = 1$, so the population is normalized to unity, and $f(w) > 0$ for all w .

The individual's budget constraint is, as before, $c = y - t(y)$, where $y = w\ell$ is labour income and $t(y)$ is a non-linear tax schedule chosen by the government. Individual utility is then given by $u(c) - y/w$. Individuals of type w will choose their consumption-income bundle $(c(w), y(w))$. It turns out to be convenient for analytical purposes to rescale individual utility by defining $V(w) \equiv wu(c(w)) - y(w)$. The variable $V(w)$ will be used as a state variable in the government's problem below.¹⁹ Note that for a person of wage rate w , the definition of $V(w)$ implicitly defines $c(w)$ as a function of $y(w)$ and $V(w)$, such that $\partial c(w)/\partial y(w) = \partial c(w)/\partial V(w) = 1/(wu'(c(w)))$. This is just the marginal rate of substitution of $c(w)$ for $y(w)$, $MRS_{c,y}$. A household facing a budget constraint $c = y - t(y)$ will select a (c, y) -bundle such that $MRS_{c,y} = 1 - t'(y)$, or $1/(wu'(c)) = 1 - t'(y)$, where $t'(y)$ is the marginal tax rate.

The marginal rate of substitution between consumption and income is strictly decreasing in ability, $\partial MRS_{c,y}/\partial w = -1/(w^2 u'(c)) < 0$. This is the single-crossing property: indifference curves of individuals with different abilities cross only once in consumption-income space. This implies that if an individual with $w = w_0$ finds it optimal not to work, $\ell(w_0) = 0$, any individual with $w < w_0$ will also find it optimal not to work. This follows from the utility-maximization behaviour of the individual and was first proved by Mirrlees (1971). The result that the optimal nonlinear tax schedule may generate unemployment is an accepted aspect of the optimal non-linear income tax schedule, and one that will be important to our analysis.

¹⁹ This procedure of rescaling the utility function by w follows Lollivier and Rochet (1983), who studied the bunching properties of the optimal income tax problem using quasi-linear preferences. The use of the function $V(w)$ essentially simplifies the form of the incentive-compatibility conditions, as shown below and allows us to substitute y out of the formal optimization problem of the government. See also Weymark (1986).

As in the previous section, it is assumed that the government maximizes a utilitarian social welfare function, here given by:

$$SW = \int_{\underline{w}}^{\bar{w}} \frac{V(w)}{w} f(w) dw \quad (1)$$

The government is only able to observe individuals' earned incomes. The government's choice of an optimal income tax schedule is effectively equivalent to its choice of a menu of (c, y) -bundles from which the individuals may choose. For an individual of ability w , the government would like them to select the tax bundle intended for their true type, $(c(w), y(w))$. However, to ensure that this is the case, the government's policy must satisfy incentive-compatibility constraints. These constraints ensure that individuals pick the (c, y) -bundle intended for them and thereby reveal their true type to the government. To represent the incentive constraints, we can think about an individual who minimizes the difference between their maximized utility if they report their true type (i.e. pick the bundle intended for them) and their maximized utility if they report any other type (i.e. pick any other consumption-income bundle). This individual's problem is given by:

$$\min_{\hat{w}} V(w) - wu(c(\hat{w}) + y(\hat{w}))$$

The first-order condition is:

$$-wu'(c)\dot{c}(\hat{w}) + \dot{y}(\hat{w}) = 0$$

and the second-order condition is:

$$-wu''(c(\hat{w}))\dot{c}(\hat{w})^2 - wu'(c(\hat{w}))\ddot{c}(\hat{w}) + \ddot{y}(\hat{w}) \geq 0$$

where a single dot denotes the first derivative of a variable with respect to ability and two dots denote the second derivative. To ensure truth-telling behaviour, the first and second-order conditions of the individual's problem must be satisfied at $w = \hat{w}$. In this case, the first-order condition can be rewritten as the following:²⁰

$$\dot{V}(w) = u(c(w)) \quad (2)$$

²⁰ Differentiating $V(w) = wu(c(w)) - y(w)$, we obtain $\dot{V}(w) = wu'(c)\dot{c}(w) - \dot{y}(w) + u(c(w))$. Evaluating the first-order condition of the individual's problem at $\hat{w} = w$, and substituting it into the expression for $\dot{V}(w)$ given above, we obtain (2).

and the second-order can be rewritten as:²¹

$$\dot{c}(w) \geq 0 \tag{3}$$

That is, consumption must be non-decreasing in ability. From the individual's first-order condition, this implies that income must also be non-decreasing in ability if the tax schedule is to be incentive compatible.²² Note that if the second-order incentive constraint is binding (so the so-called first-order approach is not sufficient), $\dot{c}(w) = 0$, and therefore $\dot{y}(w) = 0$ from (3). That is, households of different abilities will be *bunched*. This will be of some relevance in what follows. The government must also satisfy its revenue constraint which is given by:

$$\int_{\underline{w}}^{\bar{w}} [y(w) - c(w)]f(w)dw \geq R \tag{4}$$

Substituting for income from the definition of $V(w)$, we can rewrite the revenue constraint as,

$$\int_{\underline{w}}^{\bar{w}} [wu(c(w)) - V(w) - c(w)]f(w)dw - R \geq 0 \tag{5}$$

and thereby, substitute out income from the government's problem. (This is the simplifying feature of the quasi-linear utility function.)

The government's problem is then to maximize (1), subject to the revenue constraint (5), and the incentive-compatibility constraints, which are represented by the first-order incentive condition (2) and the second-order incentive condition (3). Since all the relevant variables are functions of ability, w , we frame the government's problem as a dynamic optimization problem with ability as the running variable. Define $x(w) = \dot{c}(w)$. Using $x(w)$ as a control variable allows us to impose a non-negativity constraint on it and ensure

²¹ Differentiating $\dot{V}(w) = wu'(c)\dot{c}(w) - \dot{y}(w) + u(c(w))$ and differentiating the first-order condition given by (2), we obtain two expressions for $\ddot{V}(w)$. Equating these two expressions and substituting this resulting expression into the second-order condition evaluated at $\hat{w} = w$, we obtain $u'(c(w))\dot{c}(w) \geq 0$, which is equivalent to (3) since $u'(c(w)) > 0$ for all $c(w)$.

²² An equivalent way to write these incentive constraints is given by:

$$V(w) \geq wu(c(w')) - y(w') = V(w') + (w - w')u(c(w))$$

for all w and w' . In other words, if $V(w)$ is convex and $dV/dw = u(c(w))$, then the incentive compatibility constraints are satisfied. See Lollivier and Rochet (1983).

that the second-order conditions are satisfied. Let $g(x(w))$ be an arbitrary increasing function with $g(0) = 0$ and $g'(x(w)) > 0$.²³ The state variables are $c(w)$ and $V(w)$.

The Hamiltonian of the government's problem is given by:

$$H = V(w) \frac{f(w)}{w} - \lambda[wu(c(w)) - V(w) - c(w)]f(w) + \pi(w)u(c(w)) + \mu(w)x(w) + \kappa(w)g(x(w))$$

where the co-state variables on the first-order and second-order incentive conditions are $\pi(w)$ and $\mu(w)$ respectively, and λ is the marginal cost of public funds (which is constant).

The optimal solution must satisfy the following necessary conditions:

$$\frac{\partial H}{\partial x} = 0 = \mu(w) + \kappa(w)g'(x(w))$$

$$\frac{\partial H}{\partial c} = -\dot{\mu}(w) = \lambda(wu' - 1)f(w) + \pi(w)u'(c(w)) \quad (6)$$

$$\frac{\partial H}{\partial V} = -\dot{\pi}(w) = \frac{f(w)}{w} - \lambda f(w) \quad (7)$$

$$\pi(\underline{w}) = \pi(\bar{w}) = \mu(\underline{w}) = \mu(\bar{w}) = 0, \quad \kappa \geq 0$$

If $\kappa(w)g'(x(w)) > 0 \rightarrow \dot{c}(w) = 0$, and if $\dot{c}(w) > 0 \rightarrow \kappa(w)g'(x(w)) = 0$. If there is no bunching, we can ignore the second-order incentive compatibility conditions and $\kappa(w)g'(x(w)) = 0$ for all w . If there is bunching over some interval, then $\kappa(w)g'(x(w))$ will be positive within the interval and zero at the endpoints of the interval. Those individuals in the bunched interval will receive the same consumption-income bundle.

To solve for the optimal paths of consumption, utility, and income both in the bunched and unbunched intervals, we first note that the value of the co-state variable $\pi(w)$ can be derived by integrating the necessary condition (7) and using the transversality condition, $\pi(\underline{w}) = 0$, to find:

$$\pi(w) = \lambda F(w) - G(w)$$

where $G(w) = \int_{\underline{w}}^w (f(m)/m)dw$ and is defined as the expected value of $1/m$ over the interval $\underline{w} \leq m \leq w$ multiplied by the number of individuals on the interval, $F(w)$. Evaluating

²³ Ebert (1992) introduced this transformation to avoid singular solutions. See also Myles (1995).

$\pi(w)$ at $w = \bar{w}$, using the transversality condition, $\pi(\bar{w}) = 0$, and recalling that $F(\bar{w}) = 1$, we can solve for the marginal cost of public funds,

$$\lambda = G(\bar{w})$$

Note that λ and therefore $\pi(w)$ depend only on the distribution of abilities. Over any unbunched intervals, the second-order incentive compatibility conditions are satisfied, so $\mu(w) = 0$. The optimal path of consumption can then be derived by substituting in our expressions for $\pi(w)$ and λ into the necessary condition (6), to obtain,

$$u'(c(w)) = \frac{f(w)}{wf(w) + F(w) - \frac{G(w)}{G(\bar{w})}} \quad (8)$$

Note that this, too, depends only on the ability distribution: it is, for example, independent of the revenue requirement, R . As long as $\dot{c}(w) > 0$, or equivalently $u'(c(w))$ is declining in w , the second-order condition will be satisfied and no bunching will occur. That clearly depends on the properties of $F(w)$. It is straightforward to show that bunching only occurs at the bottom of the distribution if skills are uniformly distributed. If skills are distributed according to a unimodal distribution, such as the lognormal distribution, bunching may also occur in the interior of the distribution.²⁴ However, for sufficient small \underline{w} , the second-order conditions will necessarily be violated at the bottom of any unimodal distribution.²⁵ In what follows, we assume that bunching occurs only at the bottom, i.e. there is no interior bunching. Thus, if bunching does occur, there will be a range of skills in $[\underline{w}, w_o]$ where all individuals receive the same bundle $(c(w_o), y(w_o))$.

The value of $c(w_o)$ is obtained from evaluating the above expression at w_o . To determine the end of the bunched interval, w_o , we note that in the bunched range, $\mu(w)$ is

²⁴ These results follow from differentiating (8) with respect to w . For a uniform distribution, this differential expression is monotonically decreasing in w . Therefore, if the second-order condition is violated at some w , i.e. $\partial u'(c(w))/\partial w > 0$, then the condition will be violated for all wages less than that w . This is not the case for a unimodal distribution. For example, with a lognormal distribution (given the modal skill level is greater than the mean skill level), there may be violations of the second-order conditions for w below $1/G(\bar{w})$ and for w above the mode.

²⁵ This follows from differentiating (8) with respect to w and evaluating the resulting expression at $w = \underline{w}$, to obtain $\partial u'(c(w))/\partial w = 1/(\underline{w}G(\bar{w})) - 2$ which will necessarily be positive for sufficiently small \underline{w} .

positive and can be determined by integrating (6) to give:

$$\mu(w) = \int_{\underline{w}}^w [\lambda(1 - mu'(c(w_o)))f(m) - \pi(m)u'(c(w_o))] dm$$

Substituting in the expressions for λ and $\pi(w)$ and using the fact that $\mu(w_o) = 0$, we derive the following expression,

$$\frac{F(w_o)}{u'(c(w_o))} = \int_{\underline{w}}^{w_o} \frac{f(w)}{u'(c(w))} dw$$

where $u'(c(w))$ is given by the expression for marginal utility in the unbunched interval.²⁶ This determines w_o . The path of consumption thus depends only on the distribution of skills and the form of the utility of consumption function, and is independent of required revenue.

We now turn to the determination of utility and income. Using the first-order incentive compatibility constraint (2), we can rewrite total utility of the individual of skill w as the following,

$$V(w) = V(\underline{w}) + \int_{\underline{w}}^w u(c(w))dw \quad (9)$$

which also depends only on the distribution of skills and $u(\cdot)$. To determine income, recall $V(w) = wu(c(w)) - y(w)$ so,

$$y(w) = wu(c(w)) - V(\underline{w}) - \int_{\underline{w}}^w u(c(w))dw$$

It can then be shown that,

$$V(\underline{w}) = \int_{\underline{w}}^{\bar{w}} K(w)dw - R$$

where $K(w) = [wu(c(w)) - c(w)]f(w) - u(c(w))[1 - F(w)]$.²⁷ Substituting this expression into $y(w)$ we find that,

$$y(w) = w(u(c(w))) - \int_{\underline{w}}^w u(c(w))dw - \int_{\underline{w}}^{\bar{w}} K(w)dw + R$$

²⁶ For more details on the characteristics of bunching in this setting, see Boadway, Cuff, and Marchand (1999).

²⁷ Integrating (9) over the entire distribution, substituting in the definition of $V(w)$ and using Fubini's Theorem to evaluate the double integral, we derive an expression for $V(\underline{w})$:

$$V(\underline{w}) = \int_{\underline{w}}^{\bar{w}} ([wu(c(w)) - y(w)]f(w) - u(c(w))[1 - F(w)]) dw$$

Substituting out $y(w)$ using the revenue constraint (4), we obtain the expression given above.

Taking the derivative of $y(w)$ with respect to R , we obtain

$$\frac{dy(w)}{dR} = 1 \quad (10)$$

Any increase in R will shift the optimal path of income parallel upwards: an increase in revenue requirement of \$1 will cause all households to increase their incomes by \$1 (recall, the population is normalized to unity). Consumption for all households will remain unchanged in both the bunched and the non-bunched intervals.

For individuals in the bunched region, $V(w) = wu(c(w_o)) - y(w_o)$. Generally, those individuals in the bunched region might supply a positive amount of labour.²⁸ However, in order to see more clearly the effect of a minimum wage and how it affects the level of unemployment, we would like to compare it to a situation where there is optimally some unemployment. To do this, we adjust the amount of required revenue so that those individuals in the bunched interval just supply zero labour. This is a valid approach because the optimal path of income changes on a one-for-one basis with any change in the required revenue and the path of consumption is unaffected. The change in revenue needed to reduce the income earned by the bunched individuals to zero can be determined by the following,

$$y(w_o) = \underline{w}u(c(w_o)) - \int_{\underline{w}}^{\bar{w}} K(w)dw + R^o > \underline{w}u(c(w_o)) - \int_{\underline{w}}^{\bar{w}} K(w)dw + R^n = 0$$

where R^o is the old required revenue and R^n is the new required revenue. The change in required revenue needed is $\Delta R = R^n - R^o = -y(w_o)$. All individuals reduce their income by $y(w_o)$ and those individuals in the bunched interval earn zero income. Suppose then that R^n is the required revenue. The revenue constraint can now be re-written as the following:

$$\int_{w_o}^{\bar{w}} [y(w) - c(w)]f(w)dw = R^n + c_0F(w_o)$$

²⁸ Actually, it is also possible that the solution to the above problem using the first-order approach could lead to *negative* amounts of labour being supplied by lower-skilled individuals. To preclude that, one would have to add a non-negative income constraint to the problem. As shown in Boadway, Cuff, and Marchand (1999), this complicates the problem considerably and has typically been neglected in the literature. We avoid it here by simply assuming the non-negative income constraint is not binding. But, the argument in the text would apply to this case as well. We can adjust the revenue requirement, R , so that bunching just occurs where incomes are zero in the bunched range.

This completes our description of the reference economy we shall use to study the effects of a legislated minimum wage. Its key features are summarized in Figure 2.

—FIGURE 2 NEAR HERE—

Minimum Wage Legislation

The next step is to introduce a minimum wage. If a firm employs a worker, then it must pay a wage that is no less than the minimum wage. The minimum wage is optimally determined by the government. It is assumed that firms can perfectly identify individuals' abilities, before they are offered a job. Therefore, a firm will offer an individual a job only if they are at least as productive as the minimum wage.

As in Section II, it is assumed that if an individual is offered a job, they must accept it: eligibility for a transfer is ruled out if the job offer is refused, a feature of typical welfare systems. As a result, individuals need not be better off working than not working in this context. The working individuals still decide on the number of hours to work, and therefore, the government is still constrained by the incentive compatibility constraint applying to them. However, the minimum wage legislation will identify all those individuals with $w < w_m$ since they will not be offered a job. The government will give them all the same transfer, T .

We proceed in a manner parallel to that of the fixed labour supply case. In the first stage, we assume that the government arbitrarily sets the minimum wage at the limit of the bunching interval w_o and re-optimizes the tax-transfer schedule to obtain the required revenue R^n . In this case, it is no longer constrained by the incentive constraints for those individuals below w_o because of our assumptions on the minimum wage legislation given above. We show that the marginal cost of public funds, λ , is reduced under this legislation. This implies that the path of consumption for those individuals working will be higher. This also implies that social welfare must be at least as great as when there was no minimum wage as the government could have chosen the same consumption allocation, but optimally does not. And, we show that the government chooses the transfer to the non-working population such that their consumption equates the marginal utility of these individuals to the marginal cost of public funds. This optimal transfer will necessarily be higher than c_o . Thus, non-working individuals are made better off. We then demonstrate

that the utility level of the individuals with skill level w_o is lower under the minimum wage legislation. This follows from the fact that the government is no longer constrained by the incentive compatibility constraints below w_o . In the second stage, we set out the conditions for the optimal choice of the minimum wage and demonstrate numerically that it can actually be less than w_o .

Stage 1: Minimum Wage, $w_m = w_o$

The problem of the government is to choose a menu of (c, y) -bundles and transfer, T , to maximize the sum of utilities, which can now be written as the following:

$$SW = \int_{w_m}^{\bar{w}} \frac{V(w)}{w} f(w) dw + u(T)F(w_m) \quad (11)$$

where $u(T)$ is independent of ability. The government's revenue constraint can also be rewritten as:

$$\int_{w_m}^{\bar{w}} (y(w) - c(w)) f(w) dw - TF(w_m) \geq R^n \quad (12)$$

Given this formulation of the problem, it is more convenient to use a calculus of variations approach to characterize the optimal nonlinear tax schedule.

The government's problem is to maximize social welfare, (11), subject to its revenue constraint, (12), and the incentive compatibility constraint, (3). It is assumed that the second-order incentive compatibility constraints (2) are non-binding for all working persons, an assumption which is not too restrictive. Unlike the previous case, we are explicitly allowing the government to choose the transfer, T , to the unemployed individuals, those with $w < w_m$.

In the calculus of variations approach, the objective function must be a function of the running variable, the state variable, and the derivative of the state variable with respect to the running variable. To formulate the government's problem in this manner, let $c(w)$ be determined by $V(w) = wu(c(w)) - y(w)$ or $c(y(w), V(w))$. Then substitute out the variable $y(w)$ using the first-order incentive constraint (3), i.e., $y(w) = wu(c(w)) - V(w) = w\dot{V} - V(w)$. The Lagrange expression for the government can be written as:

$$\mathcal{L} = \int_{w_m}^{\bar{w}} J(w, V, \dot{V}) dw + u(T)F(w_m) - \lambda TF(w_m) \quad (13)$$

where

$$J(w, V(w), \dot{V}(w)) = \left(\frac{V(w)}{w} + \lambda[w\dot{V} - V - c(w\dot{V} - V, V)] \right) f(w) \quad (14)$$

and λ is the Lagrange multiplier on the revenue constraint (12).

The optimal path of the state variable, $V(w)$, must satisfy the Euler equation, which is given by:

$$J_V = \frac{d}{dw} J_{\dot{V}}$$

where $J_V = \partial J / \partial V$ and $J_{\dot{V}} = \partial J / \partial \dot{V}$. Using (14), this becomes:

$$\left(\frac{1}{w} - \lambda \right) f(w) = \lambda \frac{d}{dw} \left[\left(w - \frac{1}{u'(c)} \right) f(w) \right] \quad (15)$$

The transversality conditions are as follows:

$$J_{\dot{V}} \Big|_{w_m} = \lambda \left(w_m - \frac{1}{u'(c(w_m))} \right) f(w_m) = 0$$

and

$$J_{\dot{V}} \Big|_{\bar{w}} = \lambda \left(\bar{w} - \frac{1}{u'(c(\bar{w}))} \right) f(\bar{w}) = 0$$

which imply that the highest ability individuals and those individuals earning the minimum wage face a zero marginal tax rate. This second result differs from the case when a minimum wage is not legislated. In that case, individuals with ability w_o have a positive marginal tax rate. The reason is that when the distribution is not truncated, the first transversality condition must hold at \underline{w} rather than w_o .

Denote the solution to the above problem by $V^m(w)$, $y^m(w)$ and $c^m(w)$. Integrating the Euler equation (15), we can rewrite the second of these transversality conditions as,

$$J_{\dot{V}} \Big|_{\bar{w}} = \int_{w_m}^{\bar{w}} \left(\frac{1}{w} - \lambda \right) f(w) dw = 0$$

From this we can solve for the marginal cost of public funds,

$$\lambda = \frac{G(\bar{w}) - G(w_m)}{1 - F(w_m)}$$

The marginal cost of public funds is thus equal to the sum of the $1/w$ of those individuals working weighted by their numbers, i.e. $E((1/w)|w \geq w_m)$. Recall, that without a minimum wage, $\lambda = G(\bar{w})$, which is necessarily higher than the marginal cost of public funds with a minimum wage at $w_m = w_o$.²⁹ The marginal cost of public funds is reduced when

²⁹ This follows from the fact that $G(w_o) > F(w_o)G(\bar{w})$.

the government truncates the distribution and re-optimizes over social welfare. This will affect the optimal path of consumption.

From the Euler equation (15), we can also solve for the optimal path of marginal utility and thereby, the optimal path of consumption. Integrating (15) over w and using the above expression for λ we derive,

$$u'(c(w)) = \frac{f(w)}{wf(w) + F(w) - F(w_m) - (1 - F(w_m)) \frac{G(w) - G(w_m)}{G(\bar{w}) - G(w_m)}} \quad (16)$$

Comparing this to the path of marginal utility of individuals with skill $w \in [w_o, \bar{w}]$ in our reference economy, expression (8), we find that marginal utility is lower with a minimum wage except at $w = \bar{w}$ where both expressions are equivalent.³⁰ Therefore, the optimal path of consumption with the minimum wage is everywhere higher over the interval $[w_o, \bar{w}]$ as shown in Figure 3.

—FIGURE 3 NEAR HERE—

The government will choose T so the follow necessary condition is satisfied.

$$\frac{\partial \mathcal{L}}{\partial T} = -\lambda F(w_m) + u'(T)F(w_m) = 0$$

This condition states that the optimal choice of T for a given w_m is such that the marginal utility of the unemployed individuals is equal to the marginal cost of public funds, $\lambda = u'(T^m)$. In other words, the revenue loss from a marginal increase in T should equal the benefit from that increase. The optimal transfer T^m will be greater than the transfer unemployed individuals receive when there is no minimum wage, c_o . To see this, evaluate expression (16) at $w = w_o$ to obtain, $u'(c(w_o)) = 1/w_o$ which is greater than the marginal cost of public funds, λ as defined above. We also know that the marginal utility of consumption with a minimum wage is lower then without a minimum wage at $w = w_o$. These observations allow us to conclude that, $\lambda = u'(T^m) < 1/w_o = u'(c^m(w_o)) < u'(c_o)$. The marginal utility of unemployed individuals is lower with a minimum wage and, thus, they must be receiving a higher transfer, so $T^m > c_o$.

³⁰ To see this take the inverse of both expression and put over the positive common denominator $G(\bar{w})[G(\bar{w}) - G(w_o)]$. (The sign of this expression is positive since $G(w)$ is increasing in w .) The resulting expression in the numerator reduces to $G(w_o) > F(w_o)G(\bar{w})$.

We next show that the government makes the individual with skill level w_o worse off when it imposes the minimum wage at w_o and re-optimizes the tax-transfer structure. To do this, we use the same procedure as in the previous section to solve for the individual's utility,

$$V^m(w_o) = \frac{1}{1 - F(w_o)} \left[\int_{w_o}^{\bar{w}} K^m(w) dw - R^n - TF(w_o) \right]$$

The rescaled utility of this individual without a minimum wage is given by the following;

$$V(w_o) = \int_{w_o}^{\bar{w}} K(w) dw + u(c(w_o)) \int_{\underline{w}}^{w_o} [wf(w) + F(w)] dw - R^n - c_o F(w_o)$$

Multiplying the both expressions by $1 - F(w_o)$ and subtracting, we obtain

$$\begin{aligned} (V^m(w_o) - V(w_o))(1 - F(w_o)) &= \int_{w_o}^{\bar{w}} (K^m(w) - K(w)) dw - (T - c_o)F(w_o) \\ + F(w_o) \left[\int_{w_o}^{\bar{w}} K(w) dw - R^n - c_o F(w_o) + u(c(w_o)) \int_{\underline{w}}^{w_o} [wf(w) + F(w)] dw \right] \\ &\quad - u(c(w_o)) \int_{\underline{w}}^{w_o} [wf(w) + F(w)] dw \end{aligned}$$

From the definition of $K(w)$, we know that

$$\frac{\partial K(w)}{\partial c(w)} = u'(c(w)) [wf(w) + F(w) - 1] + f(w)$$

This expression will be negative if $u'(c(w)) < f(w)[wf(w) + F(w) - 1]^{-1}$. We know from expressions (8) and (12) that with or without a minimum wage, this condition will hold. We also know that consumption increases when the government imposes a minimum wage at w_o . Since $u(\cdot)$ is monotonically increasing in c it must be that this condition holds for all $c(w)$ which implies the first term in the above expression is negative. We also know that $T > c_o$ and therefore, the second term is also negative. From the definition of $V(\underline{w})$, the term in the square brackets reduces to $w_o u(c(w_o))$.³¹ The last two terms can then be

³¹ Note, $V(\underline{w})$ can be re-written as

$$\begin{aligned} V(\underline{w}) &= \underline{w}u(c(w_o)) = u(c(w_o)) \int_{\underline{w}}^{w_o} [wf(w) + F(w)] dw \\ &\quad + \underline{w}u(c(w_o)) - w_o u(c(w_o)) + \int_{w_o}^{\bar{w}} K(w) dw - R^n - c_o F(w_o) \end{aligned}$$

Subtracting $\underline{w}u(c(w_o))$ from both sides gives us this result.

combined into the following expression:

$$u(c(w_o)) \left[w_o F(w_o) - \int_{\underline{w}}^{w_o} [wf(w) + F(w)] dw \right]$$

This expression will be zero if skills are distributed uniformly and negative if they are distributed according to a distribution that has a concave cumulative distributive function over the interval $[\underline{w}, w_o]$. This then confirms that individuals with ability w_o are worse off when the government imposes a minimum wage equal to w_o . In the latter case, the government is no longer constrained to ensure these individuals are just as well off as those individuals not working.³²

Stage 2: Choice of w_m

To fully characterize the effect of a minimum wage legislation, we now allow the government to optimally choose the level of the minimum wage. In fact, the government's Lagrange expression is identical to the one above, and all of the stated results hold at a given w_m . The only additional feature of the problem is the following necessary condition that must be satisfied by the optimal choice of w_m :

$$J|_{w_m} - \dot{V}(w_m) J_{\dot{V}}|_{w_m} - u(T)f(w_m) + \lambda T f(w_m) = 0$$

Using the first transversality condition, this condition reduces to:

$$u(T) - \frac{V(w_m)}{w_m} - u'(T)[y(w_m) - c(w_m) + T] = 0$$

The interpretation of the above is identical to the one given in the fixed labour supply case. The optimal choice of w_m will equate the net utility gain of having one more unemployed person (the utility the individual receives once they become unemployed, $u(T)$, minus the utility they received while they were working, $V(w_m)/w_m$) to the cost in utility terms of having one more unemployed person (the loss tax revenue, $y(w_m) - c(w_m)$, plus the additional transfer, T , multiplied by the marginal cost of public funds, $\lambda = u'(T)$). As in the fixed labour supply case, the sign of the above expression is ambiguous when evaluated

³² This result implies a government with a maxi-min objective function would not impose a minimum wage. The government must be willing to tolerate lower utility for those at w_o in order to make those below w_o better off.

at $w_m = w_o$. Unemployment may optimally increase or decrease. This ambiguity motivates the use of a simple numerical example to illustrate the possibility that unemployment can decrease when a minimum wage is imposed.

Numerical Illustration

In this subsection, we adopt a specific utility function and skill distribution and apply it to the solution derived above to highlight two results. First, individuals of ability w_o will receive lower utility when a minimum wage is legislated and second, a minimum wage can increase employment. In this example, we use a logarithmic utility function, $u(c) = \log(c)$ and a uniform distribution of skills with $w \in [1, 10]$ and population normalized to one. The optimal implementable tax schedule will give those individuals with ability in the interval $[1, 4.575]$ the same consumption-income bundle (2.2, 1.148). We adjust the required revenue so the bunched individuals just supply zero labour. We then impose the minimum wage at $w_0 = 4.575$ and given this new revenue requirement re-optimize the tax schedule for those individuals with ability $w \in [4.575, 10]$ and the transfer, T , given to individuals with ability $w \in [1, 4.575]$. The optimal transfer is 6.94, which is significantly higher than the consumption unemployed individuals received when there was no minimum wage, that is, $c_o = 2.2$. Individuals of ability $w_o = 4.575$ are worse-off with a minimum wage. Prior to the imposition of the minimum wage, $V(w_o) = 3.61$ and after its imposition, $V^m(w_o) = -0.14$. The additional consumption transferred to the unemployed individuals is financed entirely through an increase in the income earned by working individuals which reduces their total utility.

Using the Envelope Theorem, we evaluate the necessary condition on w_m at $w_m = w_o$ and find that it is negative (-.604). This implies that the optimal minimum wage will be lower than the cut-off ability level of the bunched individuals. Imposing the optimal minimum wage will increase employment in this simple example.

IV. Concluding Remarks

We have argued that the minimum wage combined with an ability to monitor the acceptance of job offers can be a useful adjunct to the optimal non-linear income tax-transfer system. By severing the incentive constraint that requires the working population to be no better off than those not working, more transfers can be directed to the unemployed.

Moreover, it is quite possible that the optimal level of unemployment decreases with the imposition of the minimum wage, contrary to the conventional wisdom. This point was illustrated using a simplified version of the standard optimal non-linear income tax of Mirrlees (1971).

The example we used had a number of limiting assumptions, most of them common to the optimal non-linear income tax literature. As in the latter, we assumed that ability was observable not just to individuals themselves but also to firms, whose wage offers precisely reflect individual abilities. Governments, on the other hand, knew only the distribution of abilities. Governments could, however, observe incomes. Thus, tax evasion is ruled out by assumption.³³ Two additional informational assumptions were used in our analysis. First, firms comply with the minimum wage legislation; that is, they do not illegally employ workers for less than the legislated minimum wage. Second, administrators know whether a transfer recipient has been offered a job. These are presumably limiting assumptions. Actual minimum wage regimes and transfer programs typically make it illegal to pay less than the minimum wage and require welfare recipients to accept job offers. But, as with tax evasion, there is undoubtedly illegal undercutting of the minimum wage. And, some welfare recipients undoubtedly turn down job offers without being detected, or perhaps do not take the effort to apply for them as required. The presence of such illegal or untruthful behaviour requires that some system of auditing or monitoring along with penalties be in place, as in the case of tax evasion. This will limit the extent to which the minimum wage can accomplish the objectives set out in this paper, just like tax evasion limits the ability of the government to redistribute using the tax-transfer system.³⁴ But it should necessarily render the arguments completely sterile.

Our model was slightly contrived in one further sense. To make our point in the clearest way, we assumed that it was optimal to have some set of low-ability individuals unemployed in the optimal tax-transfer system without a minimum wage. We chose to do

³³ It is known that the possibility of misreporting income can significantly constrain the ability of the government to redistribute. See Cremer and Gahvari (1996), Marhuenda and Ortuño-Ortín (1997) and Chandar and Wilde (1998).

³⁴ For an analysis of the implications for redistribution of imperfect monitoring of job applications and job offers, see Boadway and Cuff (1998). In their model, monitoring is needed to distinguish both between the voluntary and involuntary unemployed, and between the disabled and employables.

that by, first, assuming that there was bunching at the bottom because of a violation of the second-order conditions and, second, assuming that our revenue requirements were such that those bunched at the bottom would just supply zero labour. This is not as restrictive as it might seem. There are two other ways of ensuring that there is unemployment among those bunched at the bottom. First, as in Mirrlees (1971), this will occur if the ability level of the least skilled persons is close enough to zero. Second, incomes will be zero at the bottom if a constraint requiring incomes to be non-negative is binding at the bottom. We could have chosen to conduct our analysis in either of these situations, but decided not to in order to keep our analysis as tractable and intuitive as possible (given the limitations on clarity already implied by the non-linear income tax analysis).

More generally, there might either not be bunching at the bottom, or if there is, those in the bunched range may well have positive incomes. This would work to the disadvantage of the minimum wage, since imposing it would necessarily induce unemployment where none existed before. But, while unemployment would necessarily increase, and therefore output be foregone, social welfare could nonetheless rise with the minimum wage. The gain in social welfare because of the enhanced ability to redistribute could offset the cost associated with induced unemployment.

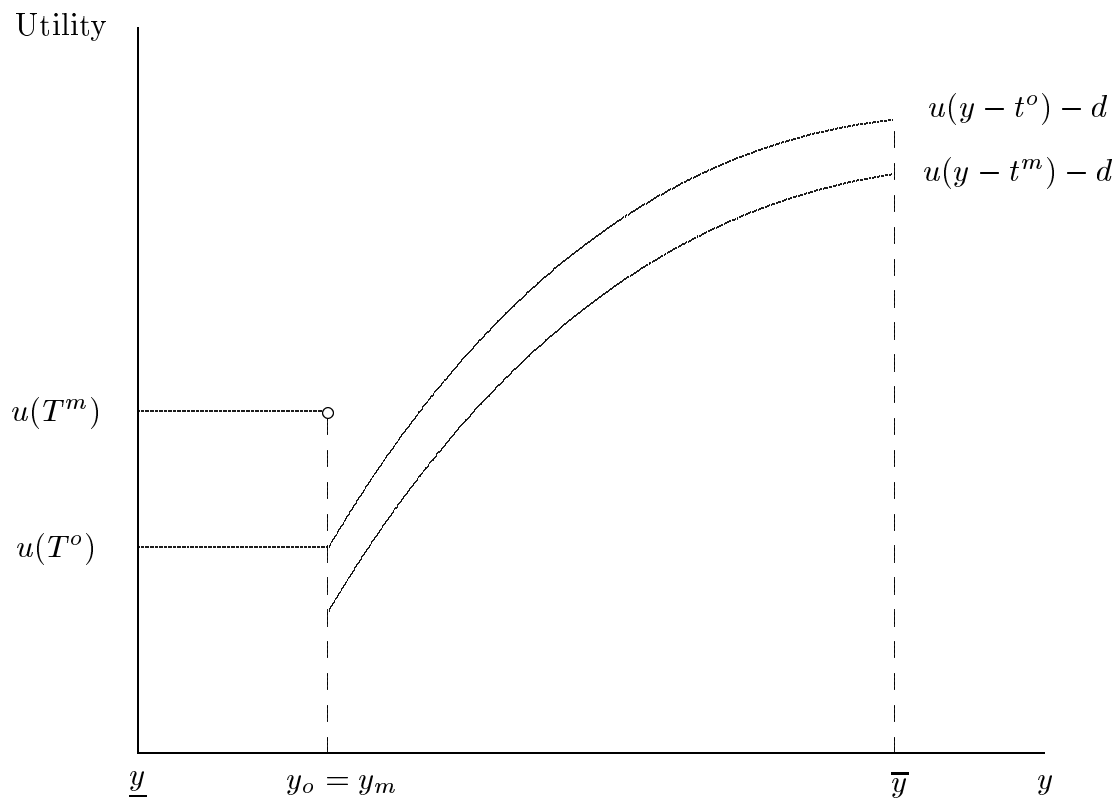


Figure 1

Fixed Labour Supply: Utility Paths with and without a Minimum Wage

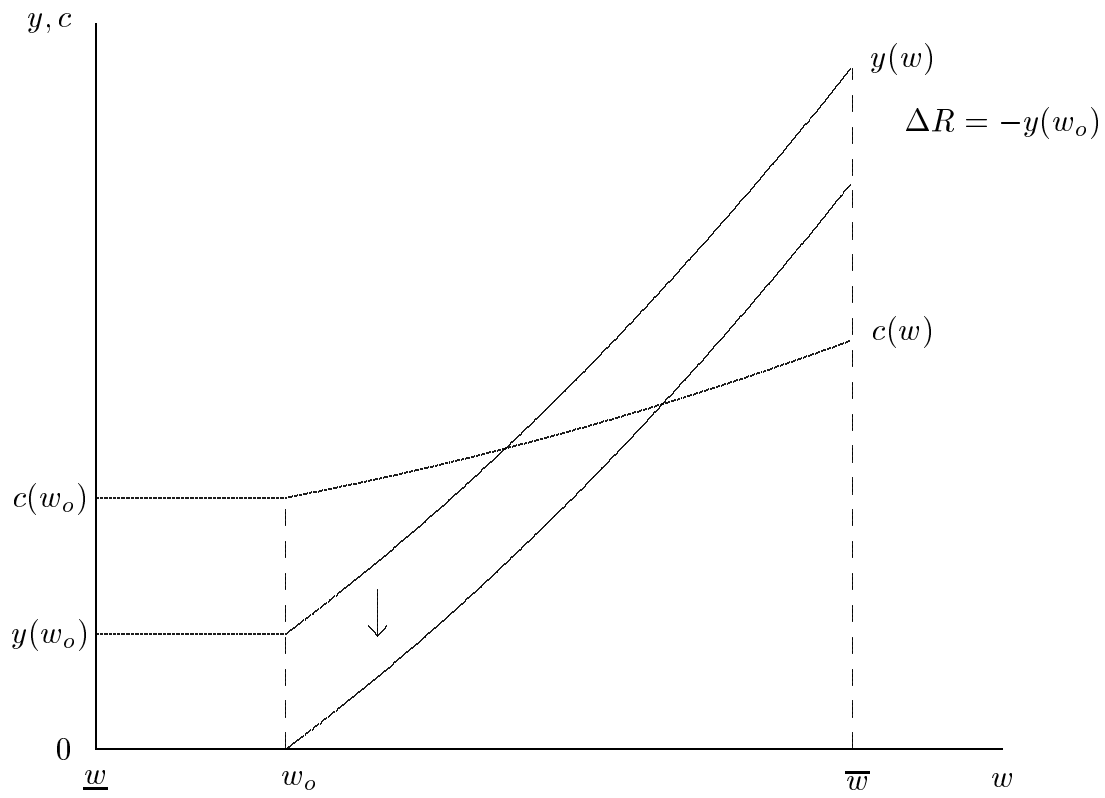


Figure 2

Variable Labour Supply: The Benchmark Case with Unemployment and No Minimum Wage

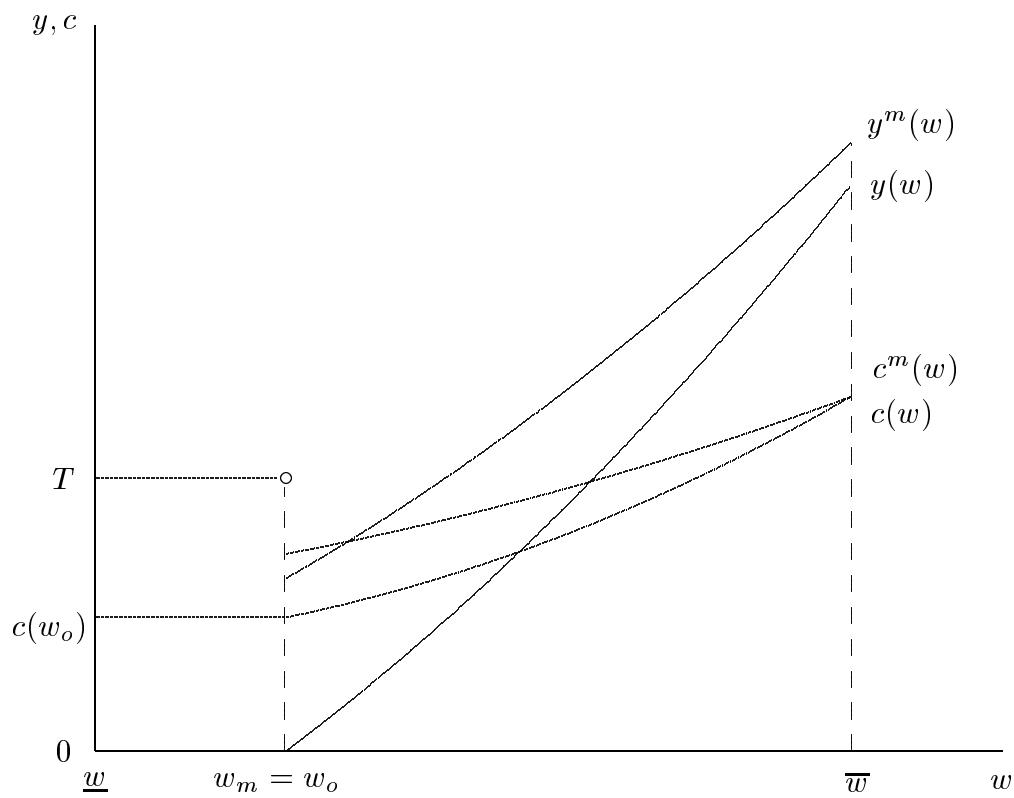


Figure 3

Variable Labour Supply with a Minimum Wage, $w_m = w_o$

References

- Boadway, R. and K. Cuff (1998), 'Monitoring Job Search as an Instrument for Targeting Transfers', paper presented at the 54th Congress of the International Institute of Public Finance, Cordoba, Argentina, August.
- Boadway, R. and M. Marchand (1995), 'The Use of Public Expenditures for Redistributive Purposes', *Oxford Economic Papers* **47**, 45-59.
- Boadway, R., M. Marchand, and M. Sato (1998), 'Subsidies Versus Public Provision of Private Goods as Instruments for Redistribution', *Scandinavian Journal of Economics* **100**, 545-64..
- Boadway, R., K. Cuff, and M. Marchand (1999), 'Optimal Income Taxation with Quasi-Linear Preferences Revisited', Queen's University, *mimeo*.
- Burkhauser, R., K. Couch, and D. Wittenburg (1998), 'A Reassessment of the New Economics of the Minimum Wage Literature Using Monthly Data from the SIPPS and CPS', *mimeo*.
- Cahuc, P. and P. Michel (1996), 'Minimum Wage Unemployment and Growth', *European Economic Review* **40**, 1463-82.
- Card, D. and A. Krueger (1995), *Myth and Measurements: The New Economics of the Minimum Wage* (Princeton, NJ: Princeton University Press).
- Chander, P. and L.L. Wilde (1998), 'A General Characterization of Optimal Income Tax Enforcement', *Review of Economic Studies* **65**, 1965-83.
- Cremer, H. and F. Gahvari (1996), 'Tax Evasion and Optimal General Income Tax', *Journal of Public Economics* **60**, 235-49.
- Ebert, U. (1992), 'A Reexamination of the Optimal Nonlinear Income Tax', *Journal of Public Economics* **49**, 47-73.
- Guesnerie, R. and K. Roberts (1984), 'Effective Policy Tools and Quantity Controls', *Econometrica* **52**, 59-86.

- Guesnerie, R. and K. Roberts (1987), 'Minimum Wage Legislation as a Second Best Policy', *European Economic Review* **31**, 490-98.
- Gunderson, M. and C.W. Riddell (1988), *Labour Market Economics, Second Edition* (Toronto: McGraw-Hill Ryerson Limited).
- Lollivier, S. and J.C. Rochet (1983), 'Bunching and Second-order Conditions: A Note on Optimal Tax Theory', *Journal of Economic Theory* **31**, 392-400.
- Marceau, N. and R. Boadway (1994), 'Minimum Wage Legislation and Unemployment Insurance', *Scandinavian Journal of Economics* **96**, 67-81.
- Marhuenda, F. and I. Ortuño-Ortín (1997), 'Tax Enforcement Problems', *Scandinavian Journal of Economics* **99**, 61-72.
- Mirrlees, J.A. (1971), 'An Exploration in the the Theory of Optimum Income Taxation', *Review of Economic Studies* **38**, 175-208.
- Mirrlees, J.A. (1974), 'Notes on Welfare Economics, Information, and Uncertainty', in M. Black, D. McFadden, and S. Wu (eds.), *Essays on Equilibrium Behavior Under Uncertainty* (Amsterdam: North-Holland), 243-58.
- Myles, G. D. (1995), *Public Economics* (Cambridge, UK: Cambridge University Press).
- OECD (Organization for Economic Cooperation and Development) (1998), *Employment Outlook* (Paris: OECD).
- Pigou, A.C. (1920), *The Economics of Welfare* (London, UK: MacMillan and Company).
- Rebitzer J. B. and L. J. Taylor (1995) 'The Consequences of Minimum Wage Laws: Some New Theoretical Ideas', *Journal of Public Economics* **56**, 245-255.
- Seade J.K. (1977), 'On the Shape of Optimal Tax Schedules', *Journal of Public Economics* **7**, 203-35.
- Stigler, G. (1946), 'The Economics of Minimum Wage Legislation', *American Economic Review* **36**, 358-65.

Swinnerton K. A. (1996), 'Minimum Wages in an Equilibrium Search Model with Diminishing Returns to Labor in Production', *Journal of Labor Economics* **14**, 340-355.

Weymark, J.A. (1986), 'A Reduced-Form Optimal Nonlinear Income Tax Problem', *Journal of Public Economics* **30**, 199-217.

