Working Papers

www.cesifo.org/wp

Is a Minimum Wage an Appropriate Instrument for Redistribution?

Aart Gerritsen
Bas Jacobs

CESifo Working Paper No. 4588
Category 1: Public Finance
January 2014

CESifo *Working Paper No. 4588*

Is a Minimum Wage an Appropriate Instrument for Redistribution?

Abstract

We analyze the redistributive (dis)advantages of a minimum wage over income taxation in competitive labor markets. A minimum wage causes more unemployment, but also leads to more skill formation as unemployment is concentrated on low-skilled workers. A simple condition based on three sufficient statistics shows that a minimum wage is desirable if the social welfare gains of more skill formation outweigh the social welfare losses of increased unemployment. Using a highly conservative calibration, a minimum wage decrease is shown to be part of a Pareto-improving policy reform for all countries under consideration, except possibly the United States.

JEL-Code: D600, H210, H240, J210, J240, J380.

Keywords: minimum wage, optimal redistribution, unemployment, education.

Aart Gerritsen
Department of Public Economics
Max Planck Institute for Tax Law and Public Finance
Munich / Germany
aart.gerritsen@tax.mpg.de

Bas Jacobs
Erasmus School of Economics
Erasmus University Rotterdam
Rotterdam / The Netherlands
bjacobs@ese.eur.nl

We are grateful to The Netherlands Organization for Scientific Research for financial support (NWO Vidi Grant No. 452-07-013).

1 Introduction

The public debate on increasing, decreasing, or implementing a minimum wage is not unlike the mythological hydra of Lerna. Not long after the debate is temporarily settled in one country it rears its head in another.[1] This regularity is suggestive of the lack of consensus on the desirability of a minimum wage. Proponents emphasize its positive effects on the earnings of low-skilled workers; opponents stress that a minimum wage tends to increase unemployment. As long as the debate is framed in terms of the trade-off between these two opposing welfare effects, this paper's title question is bound to remain contentious even among economists that are not principally opposed to redistribution. This is why the merits and demerits of a minimum wage should not be judged in isolation, but in contrast to the merits and demerits of obtaining a similar degree of redistribution in a direct way by means of the income-tax system. A minimum wage can only be desirable if it is more effective than income taxation in supporting the earnings of low-skilled workers.

The central purpose of this paper is to shed light on the question whether minimum wages are optimally employed alongside income taxes in order to redistribute income. To that end, we develop a model with perfectly competitive labor markets, in which firms demand high- and low-skilled labor. Individuals decide on hours worked, participation, and education, that is, on being high- or low-skilled. Individuals are heterogeneous in a single dimension: their disutility of work, which we refer to as their ability. An individual's ability determines whether he optimally decides to be voluntarily unemployed, a low-skilled worker, or a high-skilled worker. A minimum wage, if it binds, only does so for the low-skilled segment of the labor market, thereby creating involuntary unemployment among low-skilled workers. A minimum wage has general-equilibrium effects on the entire wage structure, which recent evidence suggests could be important (e.g., Teulings, 2000, 2003; Autor, Manning, and Smith, 2010).

Our first contribution is to determine whether and when a binding minimum wage is part of an optimal redistributive policy. For this, we consider a net-income-neutral (NIN) increase in the minimum wage. As the name suggests, the NIN minimum-wage reform raises the minimum wage, while keeping workers' net incomes constant by perfectly offsetting tax changes. Due to general-equilibrium effects minimum-wages affect the gross wage rates of all job types. As the job type must be verifiable to enforce a minimum wage, we assume that taxes can also be conditioned on job type. This enables the government to neutralize the effects of a minimum-wage increase on the net wages of

[1] In the same analogy, the modest ambition of the current study is to take up the role of Hercules.

both low- and high-skilled workers.[2] As net incomes remain constant, the reform has no direct effects on labor supply. As it raises firms' wage costs for low-skilled workers, it does reduce low-skilled labor demand. This reduced demand compresses the low-skilled labor market in two distinct ways. On the one hand, low-skilled workers are pushed 'downwards' into unemployment. On the other hand, the increased probability of unemployment induces low-skilled workers to invest in education and move 'upwards' into the high-skilled segment of the labor market.

The effects of a NIN minimum-wage increase are thus twofold: it raises both involuntary unemployment and education. The unemployment effect causes two separate welfare losses: laid-off workers who strictly prefer employment over unemployment suffer direct utility losses, and tax revenue declines if the government employs progressive taxation to redistribute from the employed to the unemployed. The education effect results in higher tax revenues and thus a welfare gain if high-skilled workers are taxed more heavily than low-skilled workers. A NIN minimum-wage increase is desirable if the welfare gains from more education outweigh the welfare losses from higher unemployment. This is true regardless of the policy mix in place. However, if this condition holds in the tax optimum without a minimum wage, a minimum wage must be part of the optimal policy mix and therefore be an appropriate instrument for redistribution. An alternative, but equivalent, way to state this is the following. If the optimal tax system is progressive and taxes education and participation on a net basis, the optimal tax system distorts both education and participation decisions downward. A NIN minimum-wage reform exacerbates the distortion on participation by increasing unemployment, but alleviates the distortion on education as low-skilled workers seek to avoid involuntary unemployment. The net welfare effect determines a minimum wage's desirability.

Our second contribution is to demonstrate that the net welfare effect of a minimum wage critically depends on the specific assumptions we make on exactly which low-skilled workers become unemployed – given heterogeneity among low-skilled workers. We propose a general *rationing schedule* that specifies the probability of unemployment at each ability level. We furthermore define the *unemployment incidence* at a specific ability level, as the relative degree to which higher unemployment is concentrated on this ability level. Therefore, a crucial difference with previous studies is that we do not need to make specific assumptions on how unemployment depends on the source of

[2] This is similar in spirit to Kaplow (2008), who analyzes policy reforms (e.g., larger public good provision, changes in indirect taxes) by adjusting the tax schedule to keep individuals' utilities fixed. We analyze a minimum wage reform while adjusting the tax system to keep the net incomes constant of those who retain their jobs. Naturally, the optimality of a minimum wage does not depend on the particular tax reform chosen.

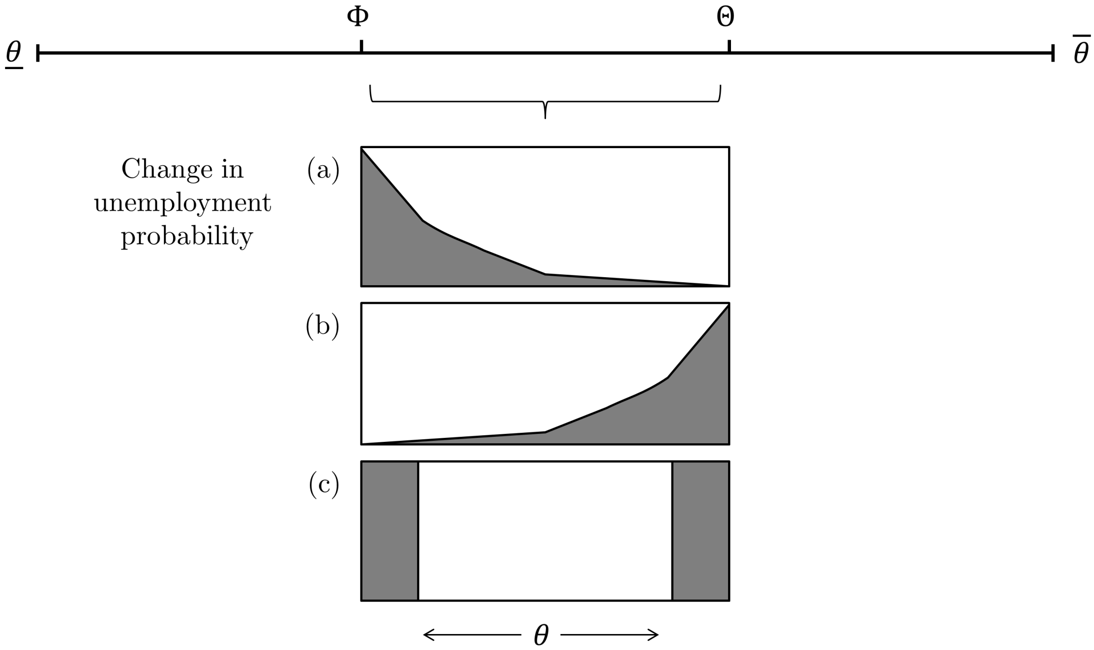

Figure 1: The unemployment incidence of a minimum wage increase: examples

heterogeneity.

To illustrate what we mean by unemployment incidence, consider Figure 1. The horizontal line represents the range of ability types, $\theta \in [\underline{\theta}, \overline{\theta}]$. In our model's equilibrium, there are two critical ability levels, Φ and Θ. For individuals with low ability, $\theta \in [\underline{\theta}, \Phi)$, the disutility of work is so large that it is optimal for them to be voluntarily unemployed. Individuals with intermediate ability, $\theta \in [\Phi, \Theta)$, supply labor as low-skilled workers. Individuals with high ability, $\theta \in [\Theta, \overline{\theta}]$, work as high-skilled workers. Now consider an increase in low-skilled unemployment due to a NIN minimum-wage increase. Panels (a)-(c) in Figure 1 depict, for three different types of unemployment incidence, the change in individual unemployment probabilities as a function of ability. In principle, there is an infinite number of possible assumptions on the exact unemployment incidence of a given increase in aggregate unemployment.

Earlier literature recognizes that the unemployment incidence is important to assess the welfare effects of a minimum wage as it determines the utility losses of laid-off workers (Lott, 1990; Palda, 2000; Lee and Saez, 2012). For example, Lee and Saez (2012) assume that the unemployment incidence is as depicted in panel (c) of Figure 1. They dub this case 'efficient rationing' as unemployment is concentrated on those individuals that are indifferent between low-skilled work and their outside option – be it non-participation or high-skilled employment. Panels (a) and (b) depict cases of 'inefficient rationing' as part of the unemployment incidence falls on individuals that are on neither the participation margin, nor the skill margin, and therefore strictly prefer low-skilled employment over their outside option. On a fundamental level, inefficient rationing occurs if a secondary market for jobs, in which workers can sell their jobs to the unemployed, is missing. We show that the unemployment incidence of the minimum wage is critical in two important ways: it determines not only the direct utility losses of the unemployed, but also the education response to a minimum-wage increase.

To see this, consider the first two panels. In panel (a), the lion share of the unemployment incidence falls on low-skilled workers who are close to the participation margin Φ and thus relatively indifferent between having a job or being non-participant. As a consequence, the education effect of higher unemployment will be relatively small. In panel (b), most of the unemployment incidence falls on low-skilled workers close to the skill margin Θ. Consequently, many low-skilled workers will decide to become high-skilled in order to avoid the increased probability of unemployment. For the same increase in aggregate unemployment, the case depicted in panel (b) leads to a larger increase in high-skilled employment than the case depicted in panel (a). A minimum wage is thus more likely to be desirable in case (b) than in case (a). More generally, the optimality of a minimum wage fundamentally depends on the incidence of unemployment. We prove that a minimum wage can always be 'made' optimal by making the appropriate assumptions on the unemployment incidence. Unfortunately, both theoretically and empirically, it is unclear how unemployment probabilities depend on workers' ability or disutility of work. Any specific assumption on the incidence of unemployment implies a certain degree of arbitrariness.

This brings us to the third contribution of this paper, which is to reinterpret the desirability condition of a NIN minimum-wage increase in terms of three sufficient statistics. These sufficient statistics allow us to forego the need to determine deeper model parameters, in particular the unemployment incidence. We assume that the low-skilled workers in our model correspond to actual workers that did not complete upper secondary education. The three sufficient statistics are then given by the social costs of low-skilled unemployment, the social gains of more upper-secondary education, and the effect of low-skilled unemployment on upper-secondary enrollment or graduation rates. The only element of the sufficient statistics that cannot easily be measured are the direct utility losses of laid-off workers. In our calibration we ignore these direct utility losses, thereby biasing results in favor of a minimum-wage increase. Data on net tax-revenue losses of unemployment and gains of education can be found for a large number of member countries of the Organisation for Economic Co-operation and Development (OECD). Moreover, there is a sizable empirical literature on the effect of unemployment on school enrollment. We make conservative assumptions when calibrating the desirability condition, further biasing our findings in favor of a minimum wage. Our results indicate that a NIN minimum-wage *decrease* would result in a Pareto improvement for almost all the countries under consideration. That is, it would increase government revenue and enable some involuntarily unemployed to find a job, while leaving no one worse off due to the net-income neutrality of the reform. Only for the United States, using a highly conservative calibration, we cannot reject the desirability of a minimum-wage increase outright.

The remainder of our paper is structured as follows. We discuss the relation to previous literature in Section 2. Section 3 introduces the theoretical model. The comparative statics of a NIN minimum-wage reform are developed in Section 4, while Section 5 provides the welfare analysis of the minimum-wage reform. Section 7 discusses the empirical application of the model and we conclude with some final thoughts.

2 Earlier literature

This paper contributes to a small strand of the literature, which analyzes minimum-wage legislation in models of competitive labor markets.[3] The classic references are Guesnerie and Roberts (1987) and Allen (1987), who adopt versions of the Stiglitz (1982) model in which workers are predetermined to be high-skilled or low-skilled. Apart from their skill type, individuals are identical so there is no need to specify a specific rationing schedule. The government sets a minimum wage which might bind for the low-skilled segment of the labor market. It also uses non-linear income taxes to optimally redistribute from high- to low-skilled workers, but cannot directly observe a worker's skill type. The goverment is thus restricted by an incentive-compatibility constraint stipulating that high-skilled workers prefer not to mimic low-skilled workers. Since this is the only constraint preventing full income redistribution, a minimum wage can never be optimal as it tightens the incentive-compatibility constraint by making it more attractive for a

[3]Note that there is also a larger literature on minimum wages in non-competitive labor markets. In most of that literature, the effects of a minimum wage are studied in isolation from income tax policy. See, for example, Gerritsen and Jacobs (2013) and Lee and Saez (2012) for overviews and extensive references. A notable exception is Hungerbühler and Lehmann (2009), who find a role for a minimum wage alongside optimal non-linear labor income taxes if workers' bargaining power is inefficiently low and the government has no means to directly control bargaining power. Our results should therefore be interpreted with caution, keeping in mind that labor-market frictions might have a separate impact on the desirability of a minimum wage.

high-skilled worker to mimic a low-skilled worker.

There are a number of drawbacks to this approach. First, workers cannot choose their skill type. For high-skilled workers it is impossible to obtain a low-skilled job; for low-skilled workers it is impossible to educate themselves to obtain a high-skilled job. A second drawback is that, in order to enforce a minimum wage, the government requires information on individual wage rates and, hence, on skill types. The very absence of such information, however, is assumed to be the reason for the government to resort to distortionary taxation in the first place.[4] This approach thus suffers from an *informational inconsistency* in the optimal redistribution problem.

These two drawbacks are avoided both by the current study and by Lee and Saez (2012). Lee and Saez, like us, assume individuals' decisions on participation and skill type are driven by their idiosyncratic disutility of work. Taxes are conditioned on skill type but the first-best allocation is unattainable because workers can decide on their skill type. As discussed above, Lee and Saez furthermore make the very specific assumption that unemployment is concentrated on those low-skilled workers with the lowest willingness to pay to remain a low-skilled worker. Under this assumption, workers that are indifferent between non-participation and low-skilled employment, or between low-skilled and high-skilled employment, are the first workers to be rationed, as illustrated by panel (c) of Figure 1. A minimum wage therefore enables the government to provide additional transfers to low-skilled workers without causing any distortion; anyone (be it a non-participant or a high-skilled worker) who tries to obtain a low-skilled job in response to these transfers faces certain unemployment, and thus no one tries. Consequently, a minimum wage is optimal if, in the optimum without a minimum wage, the government would like to redistribute more income towards low-skilled workers (away from high-skilled workers and non-participants), but is prevented from directly transferring more income towards the low-skilled because the distortions of doing so would become larger than the distributional gains.

The optimality condition of Lee and Saez crucially depends on their assumption on how the probability of unemployment is related to unobserved disutility of work. Since it is virtually impossible to empirically assess the plausibility of this assumption it remains unclear whether their results are of practical relevance. In this paper we attempt to provide a solution to this concern by developing a model in which individuals decide on participation and skill type, without imposing restrictions on how the probability

[4] Guesnerie and Roberts (1987, p.498) are perfectly aware of this, stating that the assumption that the government cannot observe wages but can enforce minimum wages "is a somewhat mixed observability assumption."

of unemployment depends on a low-skilled worker's disutility of work. We derive an optimality condition for a minimum wage that holds regardless of the incidence of unemployment. In Section 5 we provide a proof that this condition harbors the result of Lee and Saez as a special case. We furthermore show that the effect of unemployment on education can function as a sufficient statistic, foregoing the need to make a specific assumption on the unemployment incidence. Thus, in contrast to earlier literature, our model enables us to derive a condition for the desirability of a minimum wage, based on sufficient statistics that can be determined empirically, without depending on a predetermined skill distribution, informational inconsistencies, or very specific assumptions on the unemployment incidence of a minimum wage.

3 Model

This section describes in detail the behavior of individuals and firms. Individuals choose (i) whether to participate in the labor market as a high-skilled worker, participate as a low-skilled worker, or not to participate at all, and (ii) conditional on being high-skilled or low-skilled, how many labor hours to supply. Firms demand two types of labor, low-skilled and high-skilled, in competitive labor markets. In case of a binding minimum wage, low-skilled labor demand will be insufficient to match supply, causing involuntary unemployment among the low-skilled.

3.1 Preferences and budget constraints

Individuals differ in their ability θ, which has cumulative distribution $G(\theta)$, density $g(\theta)$, and support $\left[\underline{\theta}, \overline{\theta}\right]$, with $\overline{\theta} > \underline{\theta} \geq 0$. While in many studies of optimal redistribution ability determines a person's earning capacity, in our model it affects his disutility of work.[5] Specifically, take a worker with ability θ and skill type $i \in \{H, L\}$, where H stands for high-skilled and L for low-skilled work.[6] This worker's number of labor hours is denoted by l_θ^i. His disutility of work, denoted q_θ^i, is assumed to be a skill-specific

[5] We assumed heterogeneity in disutility of work, rather than earning capacity for two reasons. First, the ability to implement and enforce a minimum wage requires that the government has information on individual wages. If individuals are only heterogeneous with respect to their earnings ability, this information would enable the government to implement the first-best allocation. This is avoided by assuming heterogeneity with respect to disutility of work. Second, if workers are only heterogeneous with respect to their earning capacity, anyone with productivity below the minimum wage would simply become involuntarily unemployed. A minimum wage would, in that case, not raise any person's wage rate and as such not be an instrument for redistribution at all.

[6] Our results can readily be extended to a setting with more than two skill types. We briefly return to this point later.

function of labor hours and ability: $q^i_\theta \equiv q^i(l^i_\theta, \theta)$.

Disutility of work is increasing and convex in hours worked ($q^i_{\theta,l}, q^i_{\theta,ll} > 0$) and decreasing in ability ($q^i_{\theta,\theta} < 0$).[7] Disutility of work could simply represent the utility costs of sacrificing an hour of leisure, but could also encompass utility losses associated with educational effort. Its functional form therefore depends on whether the worker is high- or low-skilled. Specifically, disutility of a high-skilled worker is assumed to decrease with ability faster than disutility of a low-skilled worker, i.e., $q^H_{\theta,\theta} \leq q^L_{\theta,\theta}$. This ensures that individuals with high ability have a comparative advantage in high-skilled work, while individuals with low ability have a comparative advantage in low-skilled work.[8]

Utility is a twice differentiable function of consumption, c^i_θ, and disutility of work, $q^i(l^i_\theta, \theta)$. This function is assumed to be identical across individuals and given by:

$$(1) \qquad V(c^i_\theta, q^i(l^i_\theta, \theta)), \quad V_c, -V_q > 0, \quad V_{cc}, V_{qq} \leq 0, \quad i \in \{H, L\}, \quad \theta \in \left[\underline{\theta}, \overline{\theta}\right].$$

Utility is increasing in consumption, but at a non-increasing rate. Naturally, utility is decreasing in the disutility of work, and is so at a non-decreasing rate. Consequently, since disutility of work is increasing in labor hours and decreasing in ability, utility is decreasing in labor hours and increasing in ability: $V_q q^i_{\theta,l} < 0$ and $V_q q^i_{\theta,\theta} > 0$.

The household budget constraint stipulates that consumption cannot exceed the sum of after-tax labor income:

$$(2) \qquad c^i_\theta \leq \omega^i l^i_\theta + T^i, \quad i \in \{H, L\}, \quad \theta \in \left[\underline{\theta}, \overline{\theta}\right].$$

Here, ω^i is the *net* wage rate per hour worked for skill type i, i.e., the wage that is left after income taxes are paid. Skill-dependent government transfers are given by T^i. If we denote t^i as the income tax rate and w^i as the gross wage rate for skill type i, we can alternatively write net wages as:

$$(3) \qquad \omega^i \equiv (1 - t^i) w^i, \quad i \in \{H, L\}.$$

When the minimum wage binds, the low-skilled gross wage, w^L, equals the minimum wage. Due to its analytical convenience we employ net wages and the minimum wage as government instruments in the remainder of this paper. Taxes on low-skilled and

[7] For variables that depend on ability, the first subscript θ always denotes the ability level. Other subscripts denote partial derivatives. Hence, $q^i_{\theta,x} \equiv \partial q^i_\theta / \partial x$.

[8] For example, q^H_θ can include the utility costs of educational effort required to become high-skilled. The assumption that $q^H_{\theta,\theta} \leq q^L_{\theta,\theta}$ could then be reinterpreted as an assumption that the disutility of educational effort to become high-skilled decreases with ability.

high-skilled labor then follow residually from (3).

Implicit in our formulation are two assumptions. First, wages are identical for all workers with the same skill type, while they differ between workers with different skill types. Thus, workers of the same skill type are assumed to be perfect substitutes in production, whereas workers with different skill types are not.

Second, tax rates are identical for all workers with the same skill type, but differ between workers with different skill types. Allowing for fully non-linear income taxation – taxation that can be conditioned on total labor earnings, $w^i l^i_\theta$, as well as on skill type i – does not qualitatively alter our results. The reason is that a minimum wage is a skill-specific, linear policy instrument. Therefore, its effects on net labor earnings can be perfectly replicated by linear, skill-specific tax instruments. Consequently, judging the merits of a minimum wage relative to redistributive taxation does not require income taxes to be non-linear. We come back to this point in slightly more detail after deriving our main results.

So far, we discussed the utility function and budget constraints of the working population. When an individual is not employed – be it voluntarily or involuntarily so – he does not suffer any disutility of work, nor does he earn any labor income, but instead receives benefits equal to T^U. Hence, his utility is given by:

$$v^U \equiv V(T^U, 0). \tag{4}$$

3.2 Intensive labor supply

Individuals first decide to participate as a low-skilled or as a high-skilled worker or not to participate at all, then decide on how many hours of labor to supply. Using backward induction, we first solve for the optimal number of labor hours conditional on skill type. Every worker chooses his number of labor hours, l^i_θ, so as to maximize utility, (1), subject to the budget constraint, (2). Labor hours are implicitly given by equating the marginal rate of substitution of leisure for consumption with the net wage rate:

$$\frac{-V_q(c^i_\theta, q^i(l^i_\theta,\theta))q^i_l\left(l^i_\theta,\theta\right)}{V_c(c^i_\theta, q^i(l^i_\theta,\theta))} = \omega^i, \quad i \in \{H, L\}, \quad \theta \in \left[\underline{\theta}, \overline{\theta}\right]. \tag{5}$$

Jointly with the household budget constraint, this first-order condition determines an individual's intensive labor supply as a function of the net wage rate, ω^i, the government transfer, T^i, and his ability, θ:

$$l^i_\theta = l^i(\omega^i, T^i, \theta), \quad i \in \{H, L\}, \quad \theta \in \left[\underline{\theta}, \overline{\theta}\right]. \tag{6}$$

3.3 Rationing schedule and expected indirect utility

Substituting optimal labor hours and the budget constraint into the utility function yields skill-dependent indirect utility, v_θ^i, as a function of the net wage rate, ω^i, the transfer, T^i, and ability, θ:

(7)

$$v_\theta^i \equiv v^i(\omega^i, T^i, \theta) \equiv V(\omega^i l^i(\omega^i, T^i, \theta) + T^i, q^i(l^i(\omega^i, T^i, \theta), \theta)), \quad i \in \{H, L\}, \quad \theta \in \left[\underline{\theta}, \overline{\theta}\right].$$

Notice that the indirect utility function is *ex post*, in that it gives indirect utility for an individual with skill i, *conditional on obtaining a job*. However, in the case of a binding minimum wage, low-skilled workers face a positive probability of involuntary unemployment. When deciding to be low-skilled or not, an individual takes into account this probability of unemployment.[9] The *ex ante* expected indirect utility for a low-skilled individual depends on his probability of finding a job and his utility when unemployed, as well as his *ex post* indirect utility when obtaining a job. To formalize this, we first define the rationing schedule.

Definition 1 *The rationing schedule, $\{u_\theta\}$, assigns a probability of low-skilled unemployment to each level of ability $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$.*

We assume that this rationing schedule is differentiable with respect to θ. The relationship between the probability of unemployment and ability is, as we shall see, theoretically important, though empirically ambiguous. For that reason, we prefer to impose as few restrictions on the rationing schedule as possible. Instead, we determine the desirability of a minimum wage for any arbitrary rationing schedule.

Individuals decide to be either voluntarily unemployed, supply low-skilled labor, or supply high-skilled labor on the basis of the *expected* utility of the three options. Utility of being voluntarily unemployed is non-random and simply equals $V(T^U, 0)$. Similarly, as there is no unemployment among the high-skilled, utility of being high-skilled is also non-random and equals v_θ^H. Utility of an individual that decides to supply low-skilled labor, however, is random as he enjoys utility v_θ^L with probability $1 - u_\theta$, and utility $V(T^U, 0)$ with probability u_θ. Thus, the expected utility of becoming low-skilled equals:

$$(8) \quad v_\theta^{EL} \equiv v^{EL}(\omega^L, T^L, T^U, u_\theta, \theta) \equiv (1 - u_\theta) v^L(\omega^L, T^L, \theta) + u_\theta V(T^U, 0), \quad \theta \in [\underline{\theta}, \Phi).$$

[9]Thus, we assume individuals decide on their skill type before knowing with certainty whether they become employed or unemployed if low-skilled. Also, individuals cannot renege on their skill decision once unemployment is realized.

3.4 Extensive labor supply

We have assumed that individuals with low levels of θ have a comparative advantage in low-skilled work, whereas individuals with high levels of θ have a comparative advantage in high-skilled work. Therefore, in equilibrium, there are two critical ability levels, Φ and Θ, for which anyone with ability $\theta \in [\underline{\theta}, \Phi)$ becomes voluntarily unemployed, anyone with ability $\theta \in [\Phi, \Theta)$ wants to become a low-skilled worker, and anyone with ability $\theta \in \left[\Theta, \overline{\theta}\right]$ becomes a high-skilled worker. Figure 2 graphically illustrates this equilibrium. Notice that in the case of a binding minimum wage, individuals who decide to become low-skilled will either end up as low-skilled employed, or become involuntarily unemployed.

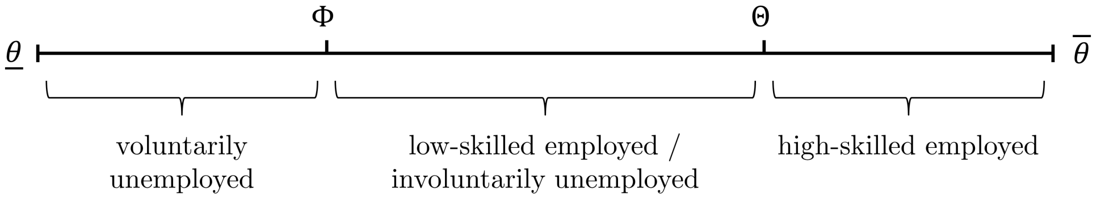

Figure 2: Graphical representation of an equilibrium

3.4.1 Participation decision

An individual with ability θ decides to be voluntarily unemployed if $v^U > v_\theta^{EL}$. Substituting for expected low-skilled utility from equation (8), this inequality can be written as $v^U > v_\theta^L$. The critical level of ability that separates non-participants from low-skilled workers, Φ, is therefore implicitly determined by:

$$V(T^U, 0) = v^L(\omega^L, T^L, \Phi). \tag{9}$$

Thus, the individual with ability Φ is indifferent between participating as a low-skilled worker and not participating at all.[10] Individuals with ability below Φ do not participate; those with ability above Φ do. As implied by equation (9), the critical level Φ can be

[10]For there to be any voluntarily unemployed individuals at all, a necessary condition is that individuals at the bottom of the ability distribution, i.e. those with ability $\underline{\theta}$, strictly prefer non-participation over participation: $V(T_U, 0) > v^L(\omega_L, T_L, \underline{\theta})$. Violation of this condition would imply everyone wants to participate. This case would have no bearing on our results and we disregard it in what follows. Furthermore, uniqueness of Φ is ensured by our assumption that low-skilled utility is strictly increasing in ability, i.e.: $v_{\theta,\theta}^L = V_q q_{\theta,\theta}^L > 0$ for all θ.

written as a function of the policy instruments ω^L, T^L and T^U:

$$\Phi = \Phi(\omega^L, T^L, T^U), \quad \Phi_{\omega^L}, \Phi_{T^L} < 0, \quad \Phi_{T^U} > 0. \tag{10}$$

Naturally, if the low-skilled net wage, ω^L, and transfer, T^L, decrease or unemployment benefits, T^U, increase, more individuals decide to become voluntarily unemployed and Φ increases as a consequence.

3.4.2 Skill decision

An individual with ability θ decides to become low-skilled rather than high-skilled if $v_\theta^{EL} > v_\theta^H$. The critical level of ability that separates the high-skilled from the low-skilled, Θ, is therefore implicitly determined by:

$$v^{EL}(\omega^L, T^L, T^U, u_\Theta, \Theta) = v^H(\omega^H, T^H, \Theta). \tag{11}$$

Thus, the individual with ability Θ is indifferent between being high- or low-skilled.[11] Individuals with ability below Θ (and above Φ) decide to be low-skilled; those with ability above Θ decide to be high-skilled.

Equation (11) implicitly determines the critical level Θ as a function of the government instruments ω^H, ω^L, T^H, T^L, and T^U, and the unemployment rate u_Θ:

$$\Theta = \Theta(\omega^H, \omega^L, T^H, T^L, T^U, u_\Theta), \quad \Theta_{\omega^H}, \Theta_{T^H}, \Theta_{u_\Theta} < 0, \quad \Theta_{\omega^L}, \Theta_{T^L}, \Theta_{T^U} > 0. \tag{12}$$

Notice that the higher is Θ, the lower is the number of high-skilled workers. Naturally, if expected low-skilled earnings rise or high-skilled earnings fall, less individuals decide to become high-skilled and Θ increases as a consequence. Therefore, the number of persons that are high-skilled is increasing in the high-skilled net wage rate, ω^H, and the high-skilled transfer, T^H. It is decreasing in the low-skilled net wage rate, ω^L, the low-skilled transfer, T^L, and unemployment benefits, T^U.

A crucial ingredient for our welfare analysis of minimum-wage legislation is the effect of unemployment on skill formation. The partial derivative of Θ with respect to u_Θ,

[11] A necessary condition for this to hold is that, at the critical level Θ high-skilled utility is increasing in ability at a faster rate than expected low-skilled utility: $v_{\Theta,\theta}^H > v_{\Theta,\theta}^{EL} + v_{\Theta,u_\theta}^{EL} u_{\Theta,\theta}$. A sufficient condition for Θ to be unique is that this holds for every level of ability: $v_{\vartheta,\theta}^H > v_{\vartheta,\theta}^{EL} + v_{\vartheta,u_\theta}^{EL} u_{\vartheta,\theta}$ for all ϑ. We assume this is indeed the case. Allowing for multiple equilibrium levels of Θ would not compromise any of the economic insights, but only increase the complexity of the analysis. We briefly return to this point later.

obtained by taking the derivative of equation (11), is given by:

$$\frac{\partial \Theta}{\partial u_\Theta} = -\frac{v^L_\Theta - v^U}{v^H_{\Theta,\theta} - v^{EL}_{\Theta,\theta} - v^{EL}_{\Theta,u_\Theta} u_{\Theta,\theta}} < 0. \tag{13}$$

Notice that the numerator gives the difference between low-skilled utility and unemployed utility for an individual with ability Θ, which is positive. It is assumed that high-skilled utility v^H_Θ is increasing in ability faster than expected low-skilled utility v^{EL}_Θ, such that the denominator is positive as well (see also footnote 11). Hence, the partial derivative of Θ with respect to u_Θ is negative. An increased probability of unemployment lowers the expected utility of being low-skilled, while it does not directly affect high-skilled utility. An increase in the unemployment rate u_Θ, *ceteris paribus*, therefore leads to more high-skilled workers and a lower level of Θ. Intuitively, when the low-skilled unemployment rate increases, some individuals escape the higher probability of low-skilled unemployment by becoming a high-skilled worker.

3.5 Aggregate labor supply

Given individuals' extensive labor supply decisions and their intensive labor supply decision, aggregate low-skilled labor supply is given by:

$$\int_\Phi^\Theta l^L_\theta \mathrm{d}G(\theta). \tag{14}$$

Notice that this aggregate labor supply is only notional, i.e., it includes labor hours of every low-skilled individual that would like to work, including the ones that are ultimately unable to find a job. Aggregate high-skilled labor supply is given by:

$$\int_\Theta^{\overline{\theta}} l^H_\theta \mathrm{d}G(\theta). \tag{15}$$

3.6 Firms

We assume workers of the same skill type are perfect substitutes in production, whereas high-skilled labor and low-skilled labor are imperfect substitutes in production. Denoting aggregate labor demand for high-skilled workers as L^H and for low-skilled workers as L^L, we can write the production function as:

$$Y \equiv F(L^H, L^L), \quad F_i > 0, \quad F_{ii} < 0, \quad F_{HL} > 0, \quad i \in \{H, L\}. \tag{16}$$

For notational convenience, we let F_i denote the partial derivative of F with respect to L^i. Marginal products of each labor type are positive but diminishing, and F is assumed to be homogeneous of degree one. A representative firm takes wages as given and decides on aggregate high- and low-skilled labor demand so as to maximize profits. This yields standard necessary (and sufficient) conditions for profit maximization:

$$F_i(L^H, L^L) = w^i, \quad i \in \{H, L\}. \tag{17}$$

Notice that labor demand is a function of gross wages. An increase in the gross wage of labor type i leads to lower demand for labor type i to realign marginal productivity and labor costs.

3.7 General equilibrium

The determination of equilibrium in a labor market with a binding minimum wage differs substantially from that in a labor market without a binding minimum wage. For high-skilled labor we assume the minimum wage is never binding. In that case, the high-skilled gross wage adjusts to equate labor supply and labor demand in the conventional way: higher demand leads to higher wages, higher supply to lower wages. In equilibrium, labor demand always equals labor supply:

$$L^H = \int_{\Theta}^{\overline{\theta}} l_\theta^H \mathrm{d}G(\theta). \tag{18}$$

If the low-skilled labor market is affected by a binding minimum wage, equilibrium can no longer be determined by wage-rate adjustments as the gross wage rate cannot be lowered. Instead, unemployment adjusts to ensure labor-market equilibrium: higher demand leads to lower unemployment, higher supply to higher unemployment. In equilibrium, labor demand equals *effective* labor supply, which is less than *notional* labor supply:

$$L^L = \int_{\Phi}^{\Theta} (1 - u_\theta) l_\theta^L \mathrm{d}G(\theta). \tag{19}$$

Notice that we have an infinity of unemployment probabilities $\{u_\theta\}$, but only one equilibrium condition. Thus, above equation simply provides the equilibrium condition for aggregate unemployment, but does not determine individual unemployment probabilities, $\{u_\theta\}$. Individual probabilities are determined by the rationing schedule. This goes directly to the core of our argument: theory points to higher unemployment due to higher minimum wages, but theoretically we know little about *which workers* become unemployed.

Substituting the labor-market equilibrium conditions into the first-order conditions of the firm, we get:

$$F_i\left(\int_{\Theta}^{\overline{\theta}} l_\theta^H \mathrm{d}G(\theta), \int_{\Phi}^{\Theta}(1-u_\theta) l_\theta^L \mathrm{d}G(\theta)\right) = w^i, \quad i \in \{H, L\}. \tag{20}$$

We can now define equilibrium in the private sector as follows.

Definition 2 *For given values of government instruments* $\{w^L, \omega^L, \omega^H, T^U, T^L, T^H\}$, *a competitive equilibrium in the private sector is defined as the allocation for which labor supply, participation and skill decisions* $\{l_\theta^L, l_\theta^H, \Phi, \Theta\}$, $\forall\theta \in [\underline{\theta}, \overline{\theta}]$, *the high-skilled wage rate* w^H, *and unemployment rates* $\{u_\theta\}$, $\forall\theta \in [\Phi, \Theta)$, *are such that the optimality conditions for labor supply, (6), participation, (10), skill formation, (12), and labor demand, (20), are satisfied.*

4 Equilibrium implications of a minimum-wage reform

4.1 A net-income-neutral minimum-wage reform

This section tracks the equilibrium implications of the decision to increase the minimum wage, while keeping net incomes of high- and low-skilled workers constant. That is, t^H and t^L endogenously adjust to changes in gross wages such that net wages ω^H and ω^L remain fixed. Transfers T^i remain constant as well. We refer to this policy as a *net-income-neutral (NIN) minimum-wage reform*, since it increases w^L while keeping $\{\omega^L, \omega^H, T^U, T^L, T^H\}$ constant.[12]

How do income taxes adjust under the NIN minimum-wage reform? Recall that $\omega^i \equiv (1-t^i)w^i$. Since $\mathrm{d}\omega^i = 0$, we must have that $\mathrm{d}t^i/(1-t^i) = \mathrm{d}w^i/w^i$. Furthermore, notice that linear homogeneity of the production function implies that profits are absent, such that $\mathrm{d}\left(F(L^H, L^L) - w^H L^H - w^L L^L\right) = 0$. Solving for the derivative and using the firm's first-order conditions, (17), yields $L^H \mathrm{d}w^H = -L^L \mathrm{d}w^L$. Intuitively, as the minimum wage increase reduces low-skilled employment, labor productivity and gross wages of high-skilled workers decline since high- and low-skilled workers are cooperant

[12] Comparative statics of a change in the minimum wage, for given tax rates, have straightforward interpretations but are mathematically tedious as we demonstrate in Gerritsen and Jacobs (2013). A NIN minimum-wage reform allows us to ignore the behavioral effects of changes in net wages.

factors of production, i.e., $F_{HL} > 0$. Thus, we can define a net-income-neutral minimum wage increase as follows.

Definition 3 *A net-income-neutral (NIN) minimum wage increase is a policy reform that raises the minimum wage, $\mathrm{d}w^L > 0$, while keeping net wage rates, ω^L, ω^H, and government transfers, T^U, T^L, T^H, constant. In order to keep net wages constant, income tax rates, t^L, t^H endogenously adjust, such that:*

$$\text{(21)} \qquad \frac{\mathrm{d}t^L}{1-t^L} = \frac{\mathrm{d}w^L}{w^L}, \quad \frac{\mathrm{d}t^H}{1-t^H} = \frac{\mathrm{d}w^H}{w^H} = -\frac{w^L L^L}{w^H L^H}\frac{\mathrm{d}w^L}{w^L}.$$

Under the NIN minimum-wage reform the low-skilled tax rate rises to offset the higher low-skilled gross wage. Furthermore, the high-skilled tax rate declines to compensate for the negative general-equilibrium effect on the high-skilled gross wage.

4.2 Comparative statics of the NIN minimum-wage reform

To understand the welfare effects of the NIN minimum-wage reform, we first derive the comparative statics of the reform for the model's key variables. To that end, it is helpful to introduce the concept of the unemployment incidence.

Definition 4 *The unemployment incidence at ability level θ, denoted by I_θ, gives the marginal increase in unemployment at that ability level, $g(\theta)\mathrm{d}u_\theta$, as a fraction of the total increase in rationing, $\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)$:*

$$\text{(22)} \qquad I_\theta \equiv \frac{g(\theta)\mathrm{d}u_\theta}{\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)} \in [0,\infty), \quad \theta \in [\Phi, \Theta).$$

What we refer to as the unemployment incidence, I_θ, is in fact a density function. I_θ measures the increase in unemployment at θ, $g(\theta)\mathrm{d}u_\theta$, relative to the increase in unemployment across all ability levels $\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)$. As with any density function, I_θ takes on a value between 0 and ∞ and the integral of the unemployment incidence equals one, i.e. $\int_\Phi^\Theta I_\theta \mathrm{d}\theta = 1$. If none of the incidence of increased unemployment is at ability level θ, then $I_\theta = 0$. If the incidence of increased unemployment is solely at ability level θ, then $I_\theta \to \infty$.

Lemma 1 *The comparative statics results of a NIN minimum-wage reform, given the equilibrium as described by Definition 2, and for given values of all other government* *instruments, are:*

$$\text{(23)} \qquad \mathrm{d}l_\theta^i = \mathrm{d}v^U = \mathrm{d}v_\theta^i = \mathrm{d}G(\Phi) = 0, \quad i \in \{H, L\}, \quad \theta \in [\underline{\theta}, \overline{\theta}],$$

$$\text{(24)} \qquad -\mathrm{d}G(\Theta) = -\frac{\partial \Theta}{\partial u_\Theta} I_\Theta \int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta),$$

$$\text{(25)} \qquad \int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta) = \left(\int_\Phi^\Theta \frac{l_\theta^L I_\theta}{L^L} \mathrm{d}\theta \right)^{-1} \left(\varepsilon_w^L \frac{\mathrm{d}w^L}{w^L} + \varepsilon_\Theta^L \mathrm{d}G(\Theta) \right).$$

Here, $\varepsilon_w^L \equiv \left(\frac{-F_{LL} L^L}{F_L} \right)^{-1} > 0$ *is the low-skilled labor-demand elasticity with respect to the low-skilled wage rate,* w^L*, and* $\varepsilon_\Theta^L \equiv \left(\frac{l_\Theta^H}{L^H} + \frac{(1-u_\Theta) l_\Theta^L}{L^L} \right) > 0$ *is the low-skilled labor-demand elasticity with respect to high-skilled labor supply,* $1 - G(\Theta)$*.*

Proof. From equations (4), (6), (7), and (10), follows that intensive labor supply, indirect utility, and the participation margin, do not depend on the minimum wage, which proves (23). Taking the derivatives of (11) and (20), and substituting for I_θ from (22), yields (24) and (25). ■

First, equations (23) demonstrate that a NIN minimum-wage reform is neutral with respect to intensive labor supply, indirect utility, and the number of voluntary unemployed. Intuitively, individuals' consumption and intensive labor supply decisions, and hence utility, depend on their net wage, not their gross wage. Provided that an individual remains (un)employed, his utility is unaffected by the NIN minimum-wage reform. Similarly, the critical level Φ equates low-skilled utility and unemployed utility and, therefore, also does not depend on the minimum wage.

Second, equation (24) gives the effect of the NIN minimum-wage reform on skill formation. The critical level Θ equates high-skilled utility with expected low-skilled utility. As such, it does not directly depend on the minimum wage, but does depend on the unemployment probability at the critical level, u_Θ. From equation (24) it follows that the change in Θ can be seen as the product of two terms. $\frac{\partial \Theta}{\partial u_\Theta}$ is the partial effect on Θ of a higher unemployment probability at ability level Θ, which is strictly negative and finite. Intuitively, when the chances of obtaining a low-skilled job diminish, individuals with ability Θ decide to become high-skilled instead. $I_\Theta \int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)$ gives the increase in the number of unemployed workers with ability Θ. Hence, if the minimum wage raises unemployment, and the unemployment incidence at ability level Θ is positive, the NIN minimum-wage reform leads to more high-skilled workers.

Third, equation (25) gives the effect of the NIN minimum-wage reform on unemployment. A minimum wage reduces low-skilled labor demand, thereby raising unemployment. This increase is seen to depend on three terms. First, it negatively depends on the labor-weighted unemployment incidence, $\int_\Phi^\Theta (l_\theta^L I_\theta / L^L)\mathrm{d}\theta$. If the unemployment incidence mainly falls on workers with high intensive labor supply, fewer workers become unemployed for a given aggregate reduction in labor demand. Second, $\varepsilon_w^L \frac{\mathrm{d}w^L}{w^L}$ captures the direct unemployment effect of higher minimum wages. For a given relative increase in minimum wages, $\mathrm{d}w^L/w^L$, the larger the labor-demand elasticity, ε_w^L, the larger the increase in low-skilled unemployment. Finally, $\varepsilon_\Theta^L \mathrm{d}G(\Theta)$ gives the unemployment effect of a change in skill formation. If a minimum wage leads to more high-skilled workers, such that $\mathrm{d}G(\Theta) < 0$, the increase in unemployment is smaller. Since low-skilled labor productivity is increasing in the number of high-skilled workers, there is less need to reduce low-skilled labor demand if skill formation increases.

The unemployment incidence at the skill margin, I_Θ, is *the* critical determinant of the comparative statics of a NIN minimum-wage increase. This is illustrated by the following Corollary.

Corollary 1 *The comparative statics of a NIN minimum-wage increase,* $\mathrm{d}w^L > 0$*, depend on* I_Θ *as follows:*

$$I_\Theta = 0: \quad -\mathrm{d}G(\Theta) = 0, \quad \int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta) > 0, \tag{26}$$

$$I_\Theta \to \infty: \quad -\mathrm{d}G(\Theta) > 0, \quad \int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta) = 0, \tag{27}$$

$$\forall\ I_\Theta \in (0,\infty): \quad \frac{-\mathrm{d}^2 G(\Theta)}{\mathrm{d}I_\Theta} > 0, \quad \frac{\int_\Phi^\Theta \mathrm{d}^2 u_\theta \mathrm{d}G(\theta)}{\mathrm{d}I_\Theta} < 0. \tag{28}$$

Proof. Solve equations (24) and (25) of Lemma 1 for $-\mathrm{d}G(\Theta)$ and $\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)$, and rearrange to obtain:

$$-\mathrm{d}G(\Theta) = -\frac{\varepsilon_w^L}{\int_\Phi^\Theta \frac{l_\theta^L I_\theta}{L^L}\mathrm{d}\theta + \varepsilon_\Theta^L \frac{-\partial\Theta}{\partial u_\Theta} I_\Theta} \frac{\mathrm{d}w^L}{w^L} \frac{\partial \Theta}{\partial u_\Theta} I_\Theta \geq 0, \tag{29}$$

$$\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta) = \frac{\varepsilon_w^L}{\int_\Phi^\Theta \frac{l_\theta^L I_\theta}{L^L}\mathrm{d}\theta + \varepsilon_\Theta^L \frac{-\partial\Theta}{\partial u_\Theta} I_\Theta} \frac{\mathrm{d}w^L}{w^L} \geq 0. \tag{30}$$

Substitute for $I_\Theta = 0$, and the limit of $I_\Theta \to \infty$ into these equations to obtain the first two lines of the Corollary. Take derivatives to obtain the third line. ■

Corollary 1 shows how changes in skill formation ($-\mathrm{d}G(\Theta)$) and total unemployment ($\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)$) depend on the unemployment incidence at the skill margin (I_Θ). Equation (26) indicates that if the incidence is 0, a NIN minimum-wage increase does not affect skill formation, but does lead to more unemployment. The intuition is straightforward: as no one at the skill margin is affected by the rise in unemployment, the minimum-wage increase does not provide incentives to become high-skilled.

Equation (27) indicates that the opposite occurs if the unemployment incidence is concentrated solely at ability level Θ. In that case, a NIN minimum-wage increase leads to more skill formation without affecting unemployment. Intuitively, if $I_\Theta \to \infty$, all individuals that would be affected by an increase in unemployment are located at the skill margin and decide to become high-skilled to escape low-skilled unemployment. Therefore, any reduction in low-skilled labor demand is completely offset by individuals moving from the low-skilled to the high-skilled sector, and no one becomes unemployed.

In general, equation (28) establishes that the increase in high-skilled employment is monotonically increasing in the unemployment incidence at the skill margin. For precisely that reason, the magnitude of the unemployment effect is monotonically decreasing in the unemployment incidence at the skill margin. Summing up, as I_Θ goes from zero to infinity, the change in skill formation monotonically goes from zero to some positive amount, and the change in unemployment monotonically goes from some positive amount to zero.

5 Welfare analysis

5.1 Social objectives and government budget constraint

We assume that social preferences are utilitarian such that social welfare, $\mathcal{W}$, is given by the unweighted sum of individual utilities:

$$(31)\quad \mathcal{W} \equiv \int_{\Theta}^{\overline{\theta}} v^H(\omega^H, T^H, \theta)\mathrm{d}G(\theta) + \int_{\Phi}^{\Theta} v^{EL}(\omega^L, T^L, T^U, u_\theta, \theta)\mathrm{d}G(\theta) + G(\Phi)V(T^U, 0).$$

It is relatively straightforward to allow for a social welfare function with social welfare weights that decline with utility. This does not affect our results.

Ability is private information, which precludes a first-best outcome as the government cannot condition its taxes on $\theta \in [\underline{\theta}, \overline{\theta}]$. Instead, the government must rely on distortionary taxes on verifiable labor income, or introduce a minimum wage, to pursue its redistributional goals. As a result, the trade-off between equity and efficiency emerges. As we have done throughout this paper, we assume that the government can condition tax rates and transfers on skill type $i \in \{H, L\}$. The government must observe wages, and thus skill levels, to implement and enforce a minimum wage.

The government obtains resources from income taxation, while it spends resources on transfers and some exogenous revenue requirement, R. Its budget $\mathcal{B}$ is thus given by:

(32)
$$\begin{aligned}\mathcal{B} \equiv &(w^H - \omega^H) \int_{\Theta}^{\overline{\theta}} l^H(\omega^H, T^H, \theta)\mathrm{d}G(\theta) + (w^L - \omega^L) \int_{\Phi}^{\Theta} (1 - u_\theta) l^L(\omega^L, T^L, \theta)\mathrm{d}G(\theta) \\ &- (1 - G(\Theta))T^H - \int_{\Phi}^{\Theta} (1 - u_\theta)\mathrm{d}G(\theta)T^L - \left(\int_{\Phi}^{\Theta} u_\theta \mathrm{d}G(\theta) + G(\Phi) \right) T^U - R.\end{aligned}$$

Notice that tax revenue per hour worked for workers of skill type i is given by $t^i w^i = w^i - \omega^i$. Thus, the first line gives total revenue from income taxation, whereas the second line gives total expenditures on transfers and the exogenous revenue requirement. The government is required to balance its budget, so that $\mathcal{B} = 0$.

5.2 Desirability of a minimum-wage increase

For any private equilibrium, a NIN minimum-wage increase, $\mathrm{d}w^L > 0$, is desirable if the change in social welfare is positive. Denoting the Lagrange multiplier associated with the government budget constraint by λ, the desirability condition for an increase in the minimum wage is given by:

$$\frac{\mathrm{d}\mathcal{W}}{\lambda} + \mathrm{d}\mathcal{B} > 0. \tag{33}$$

In order to facilitate the interpretation of the desirability condition of the minimum wage, we define *wedges* on skill formation and unemployment as Δ_Θ and Δ_u. The welfare gain of an increase in high-skilled employment is given by Δ_Θ. It is the difference between the social and private value of increased skill formation, which equals the increase in tax revenue when a low-skilled worker with ability Θ decides to become high-skilled:

$$\Delta_\Theta \equiv t^H w^H l^H_\Theta - T^H - (1 - u_\Theta)(t^L w^L l^L_\Theta - T^L) + u_\Theta T^U. \tag{34}$$

When a low-skilled individual with ability Θ becomes high-skilled, the government receives tax revenue $t^H w^H l^H_\Theta - T^H$, but foregoes expected tax revenue $(1 - u_\Theta)(t^L w^L l^L_\Theta - T^L)$ from employed low-skilled workers and $-u_\Theta T^U$ from unemployed low-skilled workers. For progressive tax systems, the tax wedge on skill formation tends to be positive, $\Delta_\Theta > 0$. In that case, more skill acquisition leads to higher tax revenue and, thus, higher social welfare.

Similarly, Δ_u is equal to the welfare loss of an additional involuntarily unemployed individual:

$$\Delta_u \equiv \int_{\Phi}^{\Theta} \left(\frac{v_\theta^L - v^U}{\lambda} + t^L w^L l_\theta^L - T^L + T^U \right) I_\theta \mathrm{d}\theta. \tag{35}$$

Unemployment affects both individuals' utility and tax revenue. Individuals that were previously employed and enjoyed utility v_θ^L become unemployed and enjoy utility $v^U \leq v_\theta^L$. Hence, welfare losses due to direct utility losses (in monetary equivalents) are given by the first term within brackets. Moreover, the government foregoes tax revenue, $t^L w^L l_\theta^L - T^L$, and has to pay additional unemployment benefits, T^U. These revenue effects are given by the remaining terms within brackets. The unemployment incidence, I_θ, determines the additional unemployment at ability level θ. Thus, the total wedge is given by the integral over utility and revenue costs, weighted by the unemployment incidence. If the government redistributes towards the unemployed, Δ_u tends to be positive as higher unemployment leads to both utility and revenue losses.

The following Proposition gives the main theoretical result of the paper: it provides the desirability condition of a NIN minimum-wage increase.

Proposition 1 *A NIN minimum-wage increase enhances social welfare if and only if the resulting marginal benefits of increased high-skilled employment outweigh the marginal costs of increased unemployment:*

$$-\Delta_\Theta \mathrm{d}G(\Theta) > \Delta_u \int_{\Phi}^{\Theta} \mathrm{d}u_\theta \mathrm{d}G(\theta). \tag{36}$$

Equivalently, a NIN minimum-wage increase enhances social welfare if and only if the marginal benefits of higher unemployment are larger than its marginal costs:

$$-\frac{\partial \Theta}{\partial u_\Theta} I_\Theta \Delta_\Theta > \Delta_u. \tag{37}$$

Proof. Take derivatives of (31) and (32) and substitute into (33). Rearrange and substitute (34) and (35) to obtain (36) (see the Appendix for the full derivation). Note from (12) that $\frac{\mathrm{d}\Theta}{\mathrm{d}u_\Theta} = \frac{\partial \Theta}{\partial u_\Theta}$ since $\left\{\omega^H, \omega^L, T^H, T^L, T^U\right\}$ remain constant. Substitute (24) and rearrange to obtain (37). ■

Lemma 1 established that the only comparative statics of a NIN minimum-wage increase consist of increases in involuntary unemployment and the number of high-skilled workers, while leaving intensive labor supply, participation and utility unaffected. In line with this finding, Proposition 1 establishes that a minimum-wage increase is desirable only if the benefits of having more high-skilled workers outweigh the costs of more unemployment. To see this, note that the left-hand side of inequality (36) gives the marginal benefits of increased skill formation while the right-hand side gives the marginal costs of higher unemployment.

Because a NIN minimum wage only affects welfare through unemployment, its desirability is equivalent to the desirability of higher unemployment. For this reason, we can rewrite the desirability condition as equation (37). Notice that the left-hand side gives the marginal benefits of unemployment. The benefits increase with the effect of aggregate unemployment on skill formation, given by $-\frac{\partial\Theta}{\partial u_\Theta}I_\Theta$. The incidence I_Θ determines the effect of an additional unemployed individual on u_Θ, whereas $\frac{\partial\Theta}{\partial u_\Theta}$ determines the effect of u_Θ on the number of high-skilled workers. The marginal benefits of unemployment are furthermore increasing with the marginal social gains of high-skilled employment, Δ_Θ. As the minimum wage leads to more high-skilled employment, such a policy is more likely to be desirable if the social benefits of high-skilled work, Δ_Θ, are high. The right-hand side of inequality (37) gives the marginal costs of unemployment, Δ_u. Since a minimum wage leads to higher unemployment, it is less likely to be desirable if the marginal social costs of higher unemployment, Δ_u, are large.[13]

Notice that if conditions (36) and (37) hold, a NIN minimum-wage increase raises social welfare, but it never constitutes a Pareto improvement if the unemployment incidence is inefficient. After all, the individuals that become involuntarily unemployed are worse off because of the reform. Conversely, if conditions (36) and (37) do not hold, a NIN minimum-wage decrease yields a genuine Pareto improvement if it leads to both higher government revenue and positive utility benefits for the involuntarily unemployed that are able to obtain a job due to the reform.[14]

Proposition 1 demonstrates the critical importance of the unemployment incidence I_Θ for the desirability of minimum wages. To illustrate this graphically, turn back to Figure 1. In the first panel, none of the unemployment incidence is on workers with ability Θ.

[13]Alternatively, the term $-\frac{\partial\Theta}{\partial u_\Theta}I_\Theta$ can be expressed in terms of an elasticity as $(H/L)\zeta_u$ which is the ratio of high-skilled individuals ($H \equiv 1 - G(\Theta)$) and low-skilled individuals ($L \equiv G(\Theta) - G(\Phi)$), multiplied with the semi-elasticity $\zeta_u \equiv \frac{-\mathrm{d}G(\Theta)/(1-G(\Theta))}{\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)/(G(\Theta)-G(\Phi))}$, which gives the relative change in the number of high-skilled workers with respect to a one-percent increase in all individual unemployment rates.

[14]Thus, the condition for a NIN minimum-wage decrease to be Pareto-improving is stronger than the simple negation of condition (37). Denoting the public revenue loss due to an additional involuntarily unemployed individual as $\Delta_\Phi \equiv t^L w^L l^L(\Phi) - T^L + T^U \leq \Delta_u$, this condition can be written as: $-\frac{\partial\Theta}{\partial u_\Theta}I_\Theta\Delta_\Theta \leq \Delta_\Phi$.

As a consequence, a NIN minimum-wage increase merely leads to more unemployment and is therefore unambiguously welfare decreasing. In that case, it would be desirable to reduce the minimum wage – in a net-income-neutral fashion by compensating tax changes. However, in the second panel of Figure 1, much of the incidence falls on workers with ability Θ. Consequently, a NIN minimum-wage increase has a large beneficial effect on high-skilled employment, relative to its adverse effect on unemployment. In that case, a minimum-wage increase might very well be desirable.

In the next section we calibrate the condition of Proposition 1 to get some idea of the desirability of a NIN minimum-wage increase (or decrease) in various OECD countries. Before doing so, however, we turn to the question of whether a binding minimum wage is part of the optimal policy mix.

5.3 A binding minimum wage as part of the policy optimum

Regardless of the initial equilibrium, Proposition 1 provides the condition under which a NIN minimum-wage increase leads to a welfare gain. However, we are especially interested in whether the minimum wage is also an appropriate instrument for redistribution, that is, whether it is part of the overall policy optimum. This must be the case if a NIN minimum-wage increase is desirable in the tax optimum without a minimum wage. In this subsection we therefore determine if condition (37) in Proposition 1 holds in the policy optimum without a binding minimum wage. For this, we rely on the following Lemma, which establishes that, in the absence of a minimum wage, the optimal wedges on unemployment (Δ_u) and skill formation (Δ_Θ) are positive and finite.

Lemma 2 *If the social marginal value of income is decreasing in ability, then the wedges on unemployment and skill formation, Δ_u and Δ_Θ, are positive and finite in the policy optimum without a minimum wage.*

Proof. See Appendix. ■

The intuition for a positive unemployment wedge ($\Delta_u > 0$) is as follows. Recall that the utility loss of a marginal increase in unemployment is non-negative, such that it suffices to show that a marginal increase in unemployment generates a net tax-revenue loss. If the government attaches a larger social marginal value of income to the unemployed than to the employed, it will, in the tax optimum, redistribute resources from the employed to the unemployed. In that case, an increase in unemployment, *ceteris paribus*, causes revenue losses, implying that $\Delta_u > 0$. The intuition for $\Delta_\Theta > 0$ follows a similar logic. If the government attaches a lower social marginal value of income to high-skilled workers than to low-skilled workers, it will redistribute resources away from the high-skilled towards the low-skilled. In that case, an increase in high-skilled employment, *ceteris paribus*, generates a revenue gain, such that $\Delta_\Theta > 0$.

Lemma 2 allows us to formulate how the optimality of a minimum wage depends on the incidence of involuntary unemployment.

Corollary 2 *Given the policy optimum without a minimum wage, there exists a value of the unemployment incidence $I^*_\Theta = \frac{\Delta_u}{\Delta_\Theta}\left(-\frac{\partial\Theta}{\partial u_\Theta}\right)^{-1}$ for which the introduction of a binding minimum wage has no effect on social welfare. For any value of $I_\Theta > I^*_\Theta$, a binding minimum wage is part of the policy optimum. For any value of $I_\Theta < I^*_\Theta$, a (marginally) binding minimum wage is not part of the policy optimum.*

Proof. Lemma 2 establishes that Δ_u and Δ_Θ are positive and finite in the policy optimum without a minimum wage. Since there is no unemployment in this optimum, Δ_u and Δ_Θ are necessarily independent of I_Θ. Equation (13) establishes that $-\partial\Theta/\partial u_\Theta$ is positive and finite and independent of I_Θ. Together with the desirability condition (37) of Proposition 1, this implies that a marginally binding minimum wage has no welfare effect for $I_\Theta = I^*_\Theta$, is welfare increasing for all $I_\Theta > I^*_\Theta$, and is welfare decreasing for all $I_\Theta < I^*_\Theta$. ■

From Corollary 1 we know that if the unemployment incidence I_Θ goes to infinity, a minimum wage does not lead to unemployment at all, but only to more high-skilled workers. In that case, a minimum wage is strictly welfare enhancing and optimally applied alongside taxes and transfers. If I_Θ goes to zero, a minimum wage only leads to higher unemployment and does not affect the number of high-skilled workers, such that a marginally binding minimum wage is strictly welfare decreasing and not a part of the optimal policy mix.[15] Thus, a binding minimum wage is an appropriate instrument for redistribution if the unemployment incidence at Θ is large enough. That is, if the increased probability of unemployment falls, to a large degree, on low-skilled workers that are indifferent between being low-skilled or high-skilled. In that case, a binding minimum wage leads to a large increase of high-skilled employment, relative to the increase in unemployment. The social benefits of this increase in high-skilled employment then outweigh the social costs of higher unemployment.

[15]Notice that a *marginal* increase in the minimum wage is not welfare enhancing if $I_\Theta < I^*_\Theta$. This does not necessarily imply that larger increases of the minimum wage are not welfare enhancing, as that would depend on the total inframarginal unemployment incidence. This is the reason we employ the term 'marginally binding minimum wage.'

5.4 Robustness

In deriving our results we made a number of assumptions that warrant a brief discussion. The results appear robust to relaxing these assumptions.

Number of skill types – We could extend the model by allowing for more than two skill types in production, in which case only the lowest skill type would be subject to a binding minimum wage. Our main results would be unaffected, provided that the government is able to tax each skill type separately. In particular, a NIN minimum-wage reform, which neutralizes the general-equilibrium effects of a minimum wage on the net returns of the additional factor inputs by appropriate tax adjustments, would result in a similar desirability condition for a minimum wage. With constant net factor returns, factor supply is unaffected, except to the extent that increased unemployment causes some low-skilled workers to supply labor as a different skill type. Again, the desirability of a minimum wage would be determined by the net balance of the costs of unemployment and the benefits of low-skilled workers deciding to become another skill type.

Uniqueness of Θ – Throughout the analysis, we assumed that the equilibrium skill margin, Θ, is unique. Uniqueness is implied by the assumption that the utility difference between high-skilled and low-skilled work is increasing in ability. This facilitates the analysis and graphical representation. However, even when Θ is not unique, the economic insights remain the same. Suppose that there are three critical ability levels at which individuals are indifferent between being low-skilled or high-skilled, Θ_a, Θ_b, and Θ_c, such that individuals with ability $\theta \in [\Phi, \Theta_a) \cup [\Theta_b, \Theta_c)$ become low-skilled, while individuals with ability $\theta \in (\Theta_a, \Theta_b] \cup \left[\Theta_c, \overline{\theta}\right]$ become high-skilled. The comparative statics of a NIN minimum-wage increase would again consist of higher unemployment and more skill formation. The extent to which it leads to more skill formation, however, now depends on the unemployment incidence at all three critical levels. Similarly, the desirability condition would feature the unemployment incidence and wedges on skill formation at all three critical levels. However, the intuition would remain the same: a NIN minimum-wage increase is only desirable if the resulting social benefits from skill formation outweigh the social costs of unemployment.

Social welfare function – Our results do not depend on the presumed utilitarian social welfare function, and would remain valid for any concave social welfare function. Naturally, a different social welfare function would affect the optimal second-best allocations, and therefore the optimal wedges on unemployment and skill formation (Δ_u and Δ_Θ). They would, however, still be positive and finite such that our results remain unaffected.

Non-linear taxation – Allowing for fully non-linear taxation, i.e., taxation conditional on both skill type and the level of labor earnings, would simply have brought more mathematical complexity without generating additional insights. Recall that the relative changes in tax rates to keep the net incomes of high-skilled and low-skilled workers constant should satisfy $\mathrm{d}t^i/(1-t^i) = \mathrm{d}w^i/w^i$ for $i \in \{H, L\}$. The NIN minimum-wage reform thus requires skill-specific, proportional tax changes. Hence, linear tax instruments are sufficient, since a minimum wage is also a linear policy instrument. Of course, any NIN minimum-wage reform under a linear tax system can be perfectly replicated using a non-linear tax system. The same welfare analysis carries over, leading to the same desirability condition for a NIN minimum-wage increase, except that the wedges (Δ_u and Δ_Θ) are determined by the non-linear tax schedule. With social marginal utility of income decreasing in ability, the government would still like to redistribute income from high-skilled to low-skilled workers and from low-skilled workers to non-participants. This implies, again, positive values for Δ_u and Δ_Θ in the optimum, confirming our results.

5.5 Comparison with Lee and Saez (2012)

Our findings can be seen to harbor the results of Lee and Saez (2012) as a special case. They analyze the desirability of a minimum wage under the specific assumption that the low-skilled workers who face unemployment are those that have the lowest willingness to pay to remain low-skilled employed. In terms of our model this implies that only individuals with ability Φ or Θ are rationed by a minimum wage, as depicted in the third panel of Figure 1. Under this assumption, dubbed 'efficient rationing' by Lee and Saez, a minimum wage is shown to be part of an optimal policy mix if, in the tax optimum without a minimum wage, the marginal social value of income of low-skilled workers exceeds the marginal value of public funds. In the Appendix we demonstrate that, in the special case of 'efficient rationing', this condition is equivalent to our desirability condition, (37), evaluated at the optimum without a minimum wage.

Intuitively, if unemployment only hits those workers with the lowest net benefits of remaining employed, the effect of a minimum-wage increase on the allocation of jobs is identical to the effect of a lower low-skilled transfer T^L. A lower transfer leads workers with the lowest willingness to pay to remain employed to decide on their outside option – be it unemployment or high-skilled employment. The only difference then, between a NIN minimum-wage increase and a decrease in the transfer T^L, is that the lower transfer leads to a transfer of resources from low-skilled workers to the government. If, in the optimum without a minimum wage, the net social value of such a transfer from low-skilled workers to the government is negative, then an unambiguous welfare gain can be made by a higher low-skilled transfer, combined with a NIN minimum-wage increase which leaves the original allocation of jobs unaltered. Such a reform would not affect the allocation of jobs, but would redistribute resources from the government to low-skilled workers.

Naturally, this result only holds if the incidence of unemployment is efficient – i.e., if the effect on the allocation of jobs of involuntary unemployment is identical to the effect of a lower low-skilled transfer. For any other rationing schedule the result of Lee and Saez breaks down. The plausibility of efficient rationing in the specific sense of Lee and Saez, however, could be criticized on both theoretical and empirical grounds. Theoretically, there is little to say in favor of any specific assumption on the incidence of unemployment, simply because it is not clear why and how the labor market would discriminate between workers which are identical in all respects but their disutility of work.[16] Moreover, there is no reliable empirical evidence on the relationship between unemployment and disutility of work.[17] We try to circumvent these problems in the next section.

6 Minimum-wage reform: an empirical application

6.1 A sufficient-statistics approach

We have seen, in Proposition 1 and Corollary 2, that the optimality of a minimum wage, or even the desirability of a minimum-wage increase, depends crucially on the unemployment incidence, I_Θ. A cynic could argue that a minimum wage can always be 'made' optimal by making the appropriate *ad hoc* assumptions on this unobservable incidence. Note, however, that the unemployment incidence enters the desirability condition, (37),

[16] Furthermore, we have theoretical difficulties with the alleged 'efficiency' of the rationing scheme of Lee and Saez. In their model, as in ours, unemployment realizations are made after individuals decide on their skill type. But, once individuals have decided on their skill type, rationing individuals with ability Θ is in fact *most* inefficient as their utility surplus over unemployment is the largest of all low-skilled workers. Once unemployment materializes and individuals can no longer renege on their skill decision, the most efficient rationing would be concentrated solely at ability Φ, rather than at both Φ and Θ.

[17] The only more or less direct evidence of the (in)efficiency of lay-offs due to a minimum wage, and thus indirectly of the relationship between lay-offs and ability, is given by Luttmer (2007). He measures the change in the average (proxy of the) reservation wage of low-skilled workers after an increase in the minimum wage. For two out of four proxies, he finds a statistically significant drop in reservation wages. This could be interpreted as evidence that workers with the highest reservation wages, and thus highest disutility of work, are rationed first. In a sensitivity analysis, however, he finds significant *increases* in reservation wages for the other two proxies, suggesting that rationing is inefficient. There is, however, plenty of evidence on misallocation due to price controls in markets for rental housing, gasoline, and natural gas. See, again, Luttmer (2007) for references.

solely because it partly determines the effect of higher unemployment on skill formation. If we have a measure of this unemployment-education effect, we can use this measure as a *sufficient statistic*, avoiding the need to calibrate I_Θ. In this section we illustrate how such a sufficient-statistics approach might help us to bring the desirability condition to the data.

To calibrate the effect of unemployment on high-skilled employment for a number of countries, we make the assumption that low-skilled workers in our model are individuals that did not complete upper-secondary education. This is a strong assumption: dropping out of secondary school does not condemn one to working for a minimum wage, and upper-secondary education is hardly a guarantee for a job at a higher wage rate. Nevertheless, schooling seems to be an important force driving both labor earnings and employment opportunities (e.g., Nickell, 1979; Card, 1999). Moreover, there is a sizable empirical literature on the effect of the low-skilled unemployment rate on enrollment rates for upper-secondary education. Denoting the enrollment rate as e, and the low-skilled unemployment rate as $\bar{u} \equiv \int_\Phi^\Theta u_\theta \mathrm{d}G(\theta)/(G(\Theta) - G(\Phi))$, such studies generally attempt to measure

$$\eta \equiv \frac{\mathrm{d}e}{\mathrm{d}\bar{u}}, \tag{38}$$

by regressing the enrollment rate on the unemployment rate (and various control variables) for a panel of regions.[18] We can use these estimates to avoid making specific assumptions on the unemployment incidence I_Θ. For this, we relate $\mathrm{d}e$ to the change in the number of high-skilled individuals, $-\mathrm{d}G(\Theta)$, and $\mathrm{d}\bar{u}$ to the (weighted) change in unemployment rates $\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)$.

First, denote the cohort size of people that are eligible for graduation from upper-secondary education by S. Assume that drop-out rates are negligible.[19] In that case, an increase in the enrollment rate of $\mathrm{d}e$ increases the number of high-skilled individuals by $S\mathrm{d}e$. Hence, we can write

$$\mathrm{d}e = \frac{-\mathrm{d}G(\Theta)}{S}. \tag{39}$$

Second, denote the number of low-skilled individuals as $L \equiv G(\Theta) - G(\Phi)$, and note

[18] Note that a drop in enrollment rates does not necessarily imply a drop in upper-secondary educational attainment levels, as people who drop out of high-school may decide to enter again later. Card and Lemieux (2001), however, analyze trends in both enrollment rates and educational attainment and conclude that dropping out of high school is, by and large, a once-for-all decision.

[19] By not allowing for drop out in upper-secondary education we overstate the effects of minimum wages on high-skilled employment, so that our results are biased in favor of minimum wages.

that $\mathrm{d}\bar{u}$ can be written as:[20]

$$\mathrm{d}\bar{u} = \frac{1}{L}\left(\int_{\Phi}^{\Theta} \mathrm{d}u_\theta \mathrm{d}G(\theta) + (u_\Theta - \bar{u})\mathrm{d}G(\Theta)\right). \tag{40}$$

The change in the average low-skilled unemployment rate does not directly correspond to the weighted sum of changes in unemployment probabilities, $\int_{\Phi}^{\Theta} \mathrm{d}u_\theta \mathrm{d}G(\theta)$, but also incorporates the effect of a change in high-skilled employment on the average unemployment rate, $(u_\Theta - \bar{u})\mathrm{d}G(\Theta)$. This implies that we need to make a specific assumption on $(u_\Theta - \bar{u})$ to write the desirability condition in terms of η. The smaller is $(u_\Theta - \bar{u})$, the more likely it is that a NIN minimum-wage increase is welfare enhancing. Therefore, we make the highly conservative assumption that $u_\Theta - \bar{u} = -0.5$, which is based on the theoretical minimum for the unemployment rate at the skill margin ($u_\Theta = 0$), and a fifty percent average unemployment rate among the low-skilled workers ($\bar{u} = 0.5$), which can be regarded as an empirical upper bound. The following Lemma establishes how η relates to the desirability condition for a NIN minimum-wage increase.

Lemma 3 *Assume that low-skilled workers are workers without upper-secondary education. Furthermore, assume that drop-out rates are negligible. Finally, assume that $u_\Theta - \bar{u} = -0.5$. Then we can rewrite the desirability condition for a NIN minimum-wage increase as:*

$$\eta > \eta^* \equiv \frac{L}{S}\left(\frac{\Delta_u/\Delta_\Theta}{1 + \frac{1}{2}\Delta_u/\Delta_\Theta}\right). \tag{41}$$

Proof. Substitute for $\int_{\Phi}^{\Theta} \mathrm{d}u_\theta \mathrm{d}G(\theta)$ in the desirability condition (36), by using (40). Substitute for $-\mathrm{d}G(\Theta)$ by using (39). Finally, substitute $u_\Theta - \bar{u} = -0.5$ and rearrange to obtain (41). ■

For given levels of L and S, which are readily available for almost any country, the desirability of a minimum-wage increase is thus seen to depend on three sufficient statistics: the effect of the unemployment rate on school enrollment rates, η, the welfare gain of schooling, Δ_Θ, and the welfare loss of unemployment, Δ_u.[21] Since upper-secondary

[20] We substituted for $\mathrm{d}\Phi = 0$, as a NIN minimum wage reform does not affect the participation margin.

[21] Our formulation captures the marginal impact of the increase in skilled employment through an increase in the enrollment rate in upper-secondary education. If the increase in the enrollment rate would be permanently higher, the number of high-skilled workers would steadily increase, whereas the number of unemployed workers would correspondingly fall over time. As the number of low-skilled workers diminishes, the marginal cost of more low-skilled unemployment would thus decrease, making

education is typically completed around the age of 18, S is taken to be the size of the 18-year-old population cohort, which is reported for a number of OECD countries in the first column of Table 1.[22] The second column gives L, the size of the labor force that completed at most primary education. Below we discuss the calibration of the remaining sufficient statistics to determine if a NIN minimum-wage increase could be welfare enhancing.

6.2 The welfare gain of schooling

Following our assumption on skill formation, Δ_Θ measures the public revenue gain from one additional person with an upper secondary educational degree. For a number of countries these revenue gains are provided by OECD (2011a, pp. 172-73) and are reported in column 3 of Table 1. The OECD considered revenue gains from higher income taxes and employees' social-security contributions, lower transfers, and higher labor utilization, and the revenue losses from direct costs of financing education and the foregone taxes on earnings associated with education. Gains and losses are calculated over the entire life cycle and discounted at a three percent annual real interest rate to obtain the public net present value of an additional high-skilled worker: Δ_Θ.

6.3 The welfare loss of unemployment

As discussed in the previous section, Δ_u consists of both utility losses and public revenue losses associated with unemployment. Unfortunately, we have no empirical approximation for the direct utility losses and thus focus solely on the revenue losses. This implies that our empirical approximation of Δ_u is a potentially severe underestimation of the total welfare costs of unemployment. It also implies that if we find that $\eta < \eta^*$, a NIN minimum-wage decrease is not only welfare enhancing, but also constitutes a Pareto improvement. Such a reform would then lead to higher government revenue, in addition to higher utility for those unemployed individuals that manage to find a job thanks to the lower minimum wage (also see footnote 14).

Statistics on the revenue losses from low-skilled unemployment are extracted from OECD (2011b, p. 56). The OECD reports the participation tax rate of an individual moving from short-term unemployment to full-time work at 50 percent of the average

minimum wages less harmful over time or maybe even desirable. However, our desirability condition does not permit the analysis of such non-marginal changes; it should be evaluated at the new allocation using updated values for e and L.

[22] The sample is restricted only by the availability of data on Δ_Θ and Δ_u.

wage. These values take into account the losses from lower income taxes and social-security contributions, and higher social, housing, family, and unemployment benefits, together with the gains from lower in-work tax benefits, if applicable. Multiplying these participation tax rates with the average minimum wage income[23], also obtained from the OECD, we calculate a value for Δ_u as shown in the fourth column of Table 1.[24]

6.4 The effect of unemployment on schooling

A recent study by Clark (2011) estimates the impact of the youth unemployment rate among workers aged 18 and 19 on the enrollment rate for 16-year-olds for a sample of English regions between 1975 and 2005. He finds that a one percentage-point increase in the unemployment rate leads to a 0.32 percentage-point increase in the enrollment rate for boys, while for girls the increase amounts to 0.45 percentage point. He also provides a detailed survey of earlier estimates for the United Kingdom and concludes that his estimate is at least twice as large as those found in previous studies. In another recent study on the UK, Tumino and Taylor (2013) find an effect similar to that of Clark, namely an increase in enrollment of 0.48 percentage point.[25]

While most studies on the impact of unemployment on school enrollment focus on the UK, a few studies analyze the relationship for the United States, Spain, and Denmark. For the US, Card and Lemieux (2001) use variations over states and years to estimate the effect of unemployment on enrollment rates, and find that a one percentage-point increase in the unemployment rate raises school enrollment rates of 17-year-olds by 0.40 percentage point. They also determine the effect of the unemployment rate in the state of birth at age 17 on educational attainment, and find that a one percentage-point increase in the unemployment rate leads to a 0.17 percentage-point increase in the share of high-school graduates.[26] In a study on African-American students, Kane (1994) finds

[23]Calculated as the minimum wage, relative to the average wage, multiplied by average wage income. For countries without a minimum wage we assume that the low-skilled workers that would become unemployed due to the minimum wage would earn 25 percent of the average wage. This percentage is our sample's lower bound of the minimum wage in terms of the average wage.

[24]One might wonder why the life-time value of the increase in tax revenues is taken for Δ_Θ, whereas Δ_u is a one-year cost of higher unemployment. If we would analyze a multiperiod life-cycle model with human capital formation, the marginal cost of a one-year increase in schooling would be the one-year expected forgone net earnings as a low-skilled worker, whereas the marginal benefits would consist of the discounted value of the increase in future net earnings. The corresponding wedge Δ_u would then be equal to the net cost of a one-year unemployment spell, whereas Δ_Θ would capture the discounted value of the net taxes on all future earnings increases.

[25]When splitting their sample in home-owner and non-home-owner families, they find an even larger effect on the enrollment rates of children from non home-owner families.

[26]This estimate is likely to suffer from attenuation bias because of interstate migration. After all, the unemployment rate in the state of birth is not likely to affect the schooling decision of a person

the effect to be as large as 0.6.

The disadvantage of the US studies is that data availability confines them to using the prime-age unemployment rate, which is arguably a worse proxy for low-skilled unemployment than the youth unemployment rate used in UK studies. Similar to the UK studies, Petrongolo and San Segundo (2002) analyze the impact of youth unemployment on school enrollment in Spain and find that a one percentage-point increase in the unemployment rate leads to an increase in the enrollment rate for boys of 0.44.

The Appendix gives a further overview of the estimates we discussed. Summing up, the impact of a one percentage-point increase in the unemployment rate on the enrollment rate in upper-secondary education is found to be in a range of 0.1 to 0.6, while all estimates except for one are well below 0.6.

6.5 The desirability of a minimum-wage increase

Column 5 in Table 1 provides values of η^*, the right-hand side of the desirability condition (41). A NIN minimum-wage increase is only desirable if the effect of the low-skilled unemployment rate on enrollment rates exceeds this critical value, such that $\eta > \eta^*$. If, on the other hand, the effect of unemployment on enrollment is smaller than the value in column 5, a NIN minimum-wage decrease leads to a Pareto improvement. For all countries, values in column 5 range from 0.4 for the United States to 10.3 for Spain.[27]

It is useful to consider the two extreme cases in some more detail. In the United States, a NIN minimum-wage increase is only desirable if a one percentage-point increase in the unemployment rate leads to a higher enrollment rate of at least 0.4 percentage point. At the other extreme, a minimum-wage increase in Spain only enhances welfare if a percentage-point increase in the unemployment rate leads to a higher enrollment rate of at least 10.3 percentage point. The reasons for these differences between the United States and Spain are readily observable from Table 1. For the United States, we see that the public benefits of more workers with secondary education, (Δ_Θ), are relatively large. On top of that, the size of the labor force with only primary education (L) is relatively small, such that a percentage-point increase in the low-skilled unemployment rate is less costly. Spain, on the other hand, shows a relatively small public return to secondary education and a relatively large unskilled population, raising the costs of an increase in

who moved to another state. On the basis of interstate migration data, Card and Lemieux suspect this attenuation bias to be in the order of 10-25 percent.

[27] As can be seen from the bottom row of Table 1, in France net tax revenues from a person completing upper-secondary education actually decline. Hence, regardless of the value of η, a NIN minimum-wage decrease leads to a Pareto improvement in France as lower unemployment and lower education both lead to higher public revenue.

Table 1: Calibrating the desirability condition

Country	S (1)	L (2)	Δ_Θ (3)	Δ_u (4)	η^* (5)	Minimum wage (6)
United States	4245	14993	60	8	0.4	Y
Czech Republic	132	344	20	5	0.6	Y
Germany	969	6496	65	8	0.8	N
Hungary	126	585	33	6	0.8	Y
Austria	100	754	65	8	0.9	N
United Kingdom	813	6592	95	11	0.9	Y
Poland	548	1558	9	6	1.3	Y
Sweden	133	913	30	8	1.7	N
Norway	65	537	34	10	2.2	N
Canada	450	2537	25	13	2.3	Y
Denmark	68	766	45	11	2.5	N
Italy	606	9403	37	7	2.5	N
Slovenia	23	145	23	12	2.6	Y
Finland	67	452	16	8	2.7	N
Australia	299	3177	30	10	3.0	Y
Ireland	55	465	33	15	3.1	Y
Portugal	118	3761	43	8	5.3	Y
Spain	465	10213	15	9	10.3	Y
France	833	7327	-6	14	$\Delta_\Theta < 0$	Y

All values 2009 or latest. L and S are measured in thousands; Δ_Θ and Δ_u are measured at 2009 prices, in thousands of PPP equivalent USD. Δ_Θ is an average of male and female values using shares in age-18 cohorts as weights. Δ_u is the unweighted average of revenue losses from an additional unemployed minimum- wage earning single, single parent with 2 children, one-earner married with 2 children, and two-earner married with 2 children with a spouse earning 67 percent of the average wage. All data are available in a separate spreadsheet, available upon request from the authors.
Source: OECD (2011a, pp. 172-73), OECD (2011b, p. 56), Eurostat and national statistical offices.

the unemployment rate.

On the basis of an empirical calibration of the desirability condition (41), we conclude that the effect of unemployment on enrollment should exceed 0.4 to be able to build a case for a higher minimum wage for the United States. That is, a percentage-point increase in the unemployment rate should, *ceteris paribus*, lead to at least a 0.4 percentage-point increase in enrollment rates. As this is well within the range of empirical estimates we found, we cannot reject that a NIN minimum-wage increase might be beneficial for the United States. However, given our very conservative assumptions when calibrating the desirability condition – most importantly the fact that we ignore direct utility losses of unemployment – we do not consider the case for a NIN minimum-wage increase in the United States very strong. For all other countries, a percentage-point increase in the unemployment rate should lead to at least a 0.6 percentage-point increase in enrollment rates. Since this is the upper bound of the empirical estimates, we conclude that the case for a NIN minimum-wage increase is weak for those countries. Instead, a *decrease* of the minimum wage, along with compensating tax changes to keep net wages constant – a NIN minimum-wage decrease – would lead to a Pareto improvement.[28]

7 Conclusion

Minimum-wage legislation distinguishes itself from redistributive income taxation by raising employers' labor costs, thereby reducing low-skilled employment. As some low-skilled workers will seek to avoid an increased probability of unemployment by acquiring more skills, a minimum wage leads to both more unemployment and more education. We show that the degree to which a minimum wage leads to additional education rather than unemployment is crucially governed by the unemployment incidence of the minimum wage. If the incidence falls mainly on those low-skilled workers who are relatively inclined towards high-skilled work, the education effect is large relative to the unemployment effect. If the incidence falls mainly on low-skilled workers that are inclined towards non-participation, a minimum wage mainly leads to higher unemployment.

The welfare consequences of a minimum wage are therefore theoretically ambiguous. On the one hand, it leads to lower social welfare as the newly unemployed suffer utility losses and pay less taxes. On the other hand, it leads to social welfare gains as high-skilled workers tend to pay more taxes than low-skilled workers. A minimum wage is optimally employed alongside income taxation if and only if the gains from more

[28]Naturally, such a decrease of the minimum wage is only possible in countries that have a minimum wage. The final column of Table 1 indicates in which countries this is the case.

education outstrip the losses from higher unemployment. This is the case only if the incidence of rationing falls to a large enough extent on low-skilled workers inclined towards high-skilled work, such that a given increase in the unemployment rate leads to enough additional education.

We used data on the net revenue gains and losses of education and unemployment to calibrate the desirability condition of a net-income-neutral minimum-wage increase. Naturally, the results from this exercise should be interpreted with caution. While at several junctions we have deliberately biased our empirical calibration in favor of a minimum wage, we did not take into account the effect that labor-market frictions might have on the desirability of a minimum wage. That being said, we find that for a one percentage-point increase of the low-skilled unemployment rate, we need an increase of school enrollment rates of around half a percentage-point for a minimum-wage increase to be desirable in the United States. This is well within empirical estimates of this effect and, accordingly, we cannot reject the hypothesis that a net-income-neutral minimum wage increase might be welfare enhancing in the United States. For any other country under consideration, the required unemployment-education effect should be well above empirical estimates of this effect. For these countries, the policy recommendation is to decrease the minimum wage, while adjusting income taxes to offset effects on net wages. Our model predicts that such a reform leads to a Pareto improvement: it leads to (i) additional government revenue due to lower unemployment, exceeding the loss in government revenue due to lower education, and (ii) higher utility for those formerly unemployed individuals that can find a job due to the lower minimum wage.

Appendix

A1. Proof of Proposition 1

The change in welfare, $\mathrm{d}\mathcal{W}$, in response to a NIN minimum-wage increase, $\mathrm{d}w^L > 0$, is given by taking the derivative of welfare $\mathcal{W}$, and substituting for $\mathrm{d}\omega^i = \mathrm{d}T^i = \mathrm{d}\Phi = 0$:

$$\begin{aligned} \mathrm{d}\mathcal{W} = &-(v^H(\omega^H, T^H, \Theta) - v^{EL}(\omega^L, T^L, T^U, u_\Theta, \Theta))\mathrm{d}G(\Theta) \\ &+ \int_\Phi^\Theta v_u^{EL}(\omega^L, T^L, T^U, u_\theta, \theta)\mathrm{d}u_\theta \mathrm{d}G(\theta). \end{aligned} \tag{42}$$

The individual's first-order condition (11) for the skill choice implies that the first line is zero. Furthermore, use equation (8) to establish that: $v_u^{EL}(\omega^L, T^L, T^U, u_\theta, \theta) =$

$-v^L(\omega^L, T^L, \theta) + V(T^U, 0)$. Substituting this result into equation (42), we get:

$$\text{(43)} \qquad \mathrm{d}\mathcal{W} \equiv -\int_{\Phi}^{\Theta} (v^U - v_\theta^L)\mathrm{d}u_\theta \mathrm{d}G(\theta).$$

The change in the government budget, $\mathrm{d}\mathcal{B}$, is obtained by taking derivatives of $\mathcal{B}$, and substituting for $\mathrm{d}\omega^i = \mathrm{d}T^i = \mathrm{d}\Phi = 0$. Rearranging then yields:

$$\begin{aligned}\text{(44)} \quad \mathrm{d}\mathcal{B} = & \int_{\Theta}^{\bar{\theta}} l_\theta^H \mathrm{d}G(\theta)\mathrm{d}w^H + \int_{\Phi}^{\Theta} (1-u_\theta) l_\theta^L \mathrm{d}G(\theta)\mathrm{d}w^L \\ & - ((w^H - \omega^H) l_\Theta^H - T^H - (1-u_\Theta)\left((w^L - \omega^L) l_\Theta^L - T^L\right) + u_\Theta T^U)\mathrm{d}G(\Theta) \\ & - \int_{\Phi}^{\Theta} ((w^L - \omega^L) l_\theta^L - T^L + T^U)\mathrm{d}u_\theta \mathrm{d}G(\theta).\end{aligned}$$

From equation (21) we know that:

$$\text{(45)} \qquad \mathrm{d}w^H = -\frac{L^L}{L^H}\mathrm{d}w^L = -\frac{\int_{\Phi}^{\Theta}(1-u_\theta) l_\theta^L \mathrm{d}G(\theta)}{\int_{\Theta}^{\bar{\theta}} l_\theta^H \mathrm{d}G(\theta)}\mathrm{d}w^L.$$

This eliminates the first two terms in equation (44). Furthermore, substitute for $(w^i - \omega^i) = t^i w^i$ to obtain:

$$\begin{aligned}\text{(46)} \qquad \mathrm{d}\mathcal{B} = & -(t^H w^H l_\Theta^H - T^H - (1-u_\Theta)(t^L w^L l_\Theta^L - T^L) + u_\Theta T^U)\mathrm{d}G(\Theta) \\ & - \int_{\Phi}^{\Theta} (t^L w^L l_\theta^L - T^L + T^U)\mathrm{d}u_\theta \mathrm{d}G(\theta).\end{aligned}$$

Substituting equations (43) and (46) into condition (33), we get:

$$\begin{aligned}\text{(47)} \qquad & -\int_{\Phi}^{\Theta} \left(\frac{v^U - v_\theta^L}{\lambda} + t^L w^L l_\theta^L - T^L + T^U\right)\mathrm{d}u_\theta \mathrm{d}G(\theta) \\ & -(t^H w^H l_\Theta^H - T^H - (1-u_\Theta)(t^L w^L l_\Theta^L - T^L) + u_\Theta T^U)\mathrm{d}G(\Theta) > 0.\end{aligned}$$

Using the definition of the unemployment incidence in equation (22) to substitute for $g(\theta)\mathrm{d}u_\theta = I_\theta \int_{\Phi}^{\Theta} \mathrm{d}u_\theta \mathrm{d}G(\theta)$ yields:

$$\begin{aligned}\text{(48)} \qquad & -\int_{\Phi}^{\Theta} \left(\frac{v^U - v_\theta^L}{\lambda} + t^L w^L l_\theta^L - T^L + T^U\right) I_\theta \mathrm{d}\theta \int_{\Phi}^{\Theta} \mathrm{d}u_\theta \mathrm{d}G(\theta) \\ & -(t^H w^H l_\Theta^H - T^H - (1-u_\Theta)(t^L w^L l_\Theta^L - T^L) + u_\Theta T^U)\mathrm{d}G(\Theta) > 0.\end{aligned}$$

Finally, substituting the wedges Δ_Θ and Δ_u from equations (34) and (35) we obtain the final result of equation (36).

A2. Proof of Lemma 2

Without a minimum wage and, hence, without involuntary unemployment, social welfare is given by:

$$\text{(49)}\qquad \mathcal{W} \equiv \int_{\Theta}^{\overline{\theta}} v^H(\omega^H, T^H, \theta)\mathrm{d}G(\theta) + \int_{\Phi}^{\Theta} v^L(\omega^L, T^L, \theta)\mathrm{d}G(\theta) + G(\Phi)V(T^U, 0).$$

The government's budget constraint is given by $\mathcal{B}$:

$$\text{(50)}\qquad \mathcal{B} \equiv \int_{\Theta}^{\overline{\theta}} (w^H - \omega^H)l^H(\omega^H, T^H, \theta)\mathrm{d}G(\theta) + \int_{\Phi}^{\Theta} (w^L - \omega^L)l^L(\omega^L, T^L, \theta)\mathrm{d}G(\theta)$$
$$- (1 - G(\Theta))T^H - (G(\Theta) - G(\Phi))T^L - G(\Phi)T^U - R = 0.$$

Defining λ as the Lagrange multiplier of the budget constraint, we can set up the following maximization problem for the government:

$$\text{(51)}\qquad \max_{\{T^H, T^L, T^U, \omega^H, \omega^L\}} \mathcal{L} = \mathcal{W} + \lambda\mathcal{B},$$

which is subject to Θ, Φ, w^H, and w^L as determined by the individuals' and firms' first-order conditions in equations (6), (10), (12), and (20). Notice again that we chose net wages ω^i, rather than tax rates t^i, as the government's control variables.

The first-order conditions of this maximization problem are obtained by equating the partial derivatives of the Lagrangian to zero. These derivatives are given by:

(52)
$$\frac{\partial \mathcal{L}}{\partial \omega^H} = \int_{\Theta}^{\overline{\theta}} \left(V_{\theta,c}^H - \lambda \left(1 - \Delta_H \frac{\omega^H}{l_\theta^H} \frac{\partial l_\theta^H}{\partial \omega^H} \right) \right) l_\theta^H \mathrm{d}G(\theta) - \lambda \Delta_\Theta \frac{\mathrm{d}G(\Theta)}{\mathrm{d}\omega^H} - \lambda \Delta_\Phi \frac{\mathrm{d}G(\Phi)}{\mathrm{d}\omega^H},$$

(53)
$$\frac{\partial \mathcal{L}}{\partial T^H} = \int_{\Theta}^{\overline{\theta}} \left(V_{\theta,c}^H - \lambda \left(1 - \Delta_H \omega^H \frac{\partial l_\theta^H}{\partial T^H} \right) \right) \mathrm{d}G(\theta) - \lambda \Delta_\Theta \frac{\mathrm{d}G(\Theta)}{\mathrm{d}T^H} - \lambda \Delta_\Phi \frac{\mathrm{d}G(\Phi)}{\mathrm{d}T^H},$$

(54)
$$\frac{\partial \mathcal{L}}{\partial \omega^L} = \int_{\Phi}^{\Theta} \left(V_{\theta,c}^L - \lambda \left(1 - \Delta_L \frac{\omega^L}{l_\theta^L} \frac{\partial l_\theta^L}{\partial \omega^L} \right) \right) l_\theta^L \mathrm{d}G(\theta) - \lambda \Delta_\Theta \frac{\mathrm{d}G(\Theta)}{\mathrm{d}\omega^L} - \lambda \Delta_\Phi \frac{\mathrm{d}G(\Phi)}{\mathrm{d}\omega^L},$$

(55)
$$\frac{\partial \mathcal{L}}{\partial T^L} = \int_{\Phi}^{\Theta} \left(V_{\theta,c}^L - \lambda \left(1 - \Delta_L \omega^L \frac{\partial l_\theta^L}{\partial T^L} \right) \right) \mathrm{d}G(\theta) - \lambda \Delta_\Theta \frac{\mathrm{d}G(\Theta)}{\mathrm{d}T^L} - \lambda \Delta_\Phi \frac{\mathrm{d}G(\Phi)}{\mathrm{d}T^L},$$

(56)
$$\frac{\partial \mathcal{L}}{\partial T^U} = G(\Phi)(V_c^U - \lambda) - \lambda \Delta_\Theta \frac{\mathrm{d}G(\Theta)}{\mathrm{d}T^U} - \lambda \Delta_\Phi \frac{\mathrm{d}G(\Phi)}{\mathrm{d}T^U}.$$

We took a number of steps to arrive at these expressions. First, observe that changes in Θ and Φ have no direct effect on individuals' utility, but only affect welfare indirectly through their effect on the government budget. This is because individuals at Θ (Φ) are indifferent between low-skilled work and high-skilled work (voluntary unemployment). Second, we applied Roy's identity to rewrite derivatives of indirect utility in terms of direct utility, i.e., $v_{\theta,\omega^i}^i = l_\theta^i v_{\theta,T^i}^i$ and $v_{\theta,T^i}^i = V_{\theta,c}^i$, where $V_{\theta,c}^i$ gives the marginal utility of consumption of an individual with skill type i and ability θ. Third, from the firm's first-order conditions we substituted for $L^H \mathrm{d}w^H = -L^L \mathrm{d}w^L$, which implies that the effect on the government budget of any increase in one of the two gross wage rates is exactly offset by a decrease in the other gross wage rate. This is the reason $\mathrm{d}w^H$ and $\mathrm{d}w^L$ do not show up in the expressions. Fourth, and final, we defined wedges as follows: $\Delta_H \equiv \frac{t^H}{1-t^H}$ and $\Delta_L \equiv \frac{t^L}{1-t^L}$ are the wedges on high- and low-skilled net labor earnings; $\Delta_\Phi \equiv t^L w^L l^L(\Phi) - T^L + T^U$ is the wedge on the participation margin; $\Delta_\Theta \equiv t^H w^H l^H(\Theta) - T^H - t^L w^L l^L(\Phi) + T^L$ is the wedge on the skill margin. In other words, Δ_i, $i \in \{L^H, L^L, \Phi, \Theta\}$ denote the gains in government revenue from higher levels of $\omega^H l_\theta^H$ and $\omega^L l_\theta^L$, and lower levels of $G(\Phi)$ and $G(\Theta)$. Notice that Δ_Θ corresponds to (34) in the absence of a minimum wage and unemployment.

We can now rewrite the first-order conditions for each control variable. To facilitate the derivations, we adopt a number of additional notational conventions. We define uncompensated net-wage elasticities of intensive labor supply as $\varepsilon_{\theta,\omega^i}^{i,u} \equiv \frac{\partial l_\theta^i}{\partial \omega^i} \frac{\omega^i}{l_\theta^i}$, and income elasticities as $\varepsilon^i_{\theta,T^i} \equiv -\omega^i \frac{\partial l^i_\theta}{\partial T^i} > 0$. Applying the Slutsky equation we can write the compensated net-wage elasticity of labor supply as $\varepsilon^{i,c}_{\theta,\omega^i} = \varepsilon^{i,u}_{\theta,\omega^i} + \varepsilon^i_{\theta,T^i}$. We define the income-weighted average compensated elasticities of labor supply as:

$$\overline{\varepsilon^L_{\omega^L}} \equiv \int_\Phi^\Theta \frac{w^L l^L_\theta}{w^L L^L} \varepsilon^{L,c}_{\theta,\omega^L} \mathrm{d}G(\theta) > 0, \quad \overline{\varepsilon^H_{\omega^H}} \equiv \int_\Theta^{\overline{\theta}} \frac{w^H l^H_\theta}{w^H L^H} \varepsilon^{H,c}_{\theta,\omega^H} \mathrm{d}G(\theta) > 0, \tag{57}$$

where the signs follow from the fact that compensated wage elasticities of labor supply are always positive.

The semi-elasticities of participation and skill formation are defined as follows:

$$\eta_{T^U} \equiv \frac{\mathrm{d}G(\Phi)}{G(\Phi)\mathrm{d}T^U}, \quad \eta_{T^L} \equiv \frac{-\mathrm{d}G(\Phi)}{(G(\Theta) - G(\Phi))\,\mathrm{d}T^L}, \quad \eta_{\omega^L} \equiv \frac{-\mathrm{d}G(\Phi)}{L^L \mathrm{d}\omega^L}, \tag{58}$$

$$\zeta_{T^L} \equiv \frac{\mathrm{d}G(\Theta)}{(G(\Theta) - G(\Phi))\,\mathrm{d}T^L}, \; \zeta_{T^H} \equiv \frac{-\mathrm{d}G(\Theta)}{(1 - G(\Theta))\,\mathrm{d}T^H}, \; \zeta_{\omega^L} \equiv \frac{\mathrm{d}G(\Theta)}{L^L \mathrm{d}\omega^L}, \; \zeta_{\omega^H} \equiv \frac{-\mathrm{d}G(\Theta)}{L^H \mathrm{d}\omega^H}. \tag{59}$$

The term η_j measures the change in $G(\Phi)$ due to a marginal change of $j \in \{T^U, T^L, \omega^L\}$, and ζ_k measures the change in $G(\Theta)$ due to a marginal change of $k \in \{T^L, \omega^L, T^H, \omega^H\}$. All semi-elasticities are defined to be positive.

Following Diamond (1975), we define γ^i_θ as the social marginal value of income for an individual with ability $\theta \in [\underline{\theta}, \overline{\theta}]$ and skill level $i \in \{H, L\}$. This term consists of the private marginal utility of income, minus the social value of the loss in tax revenue due to the income effect on labor hours. Normalizing in terms of resources, by dividing by λ, this yields:

$$\gamma^U \equiv \frac{V^U_c}{\lambda}, \quad \gamma^i_\theta \equiv \frac{V^i_{\theta,c}}{\lambda} + \Delta_i \omega^i \frac{\partial l^i_\theta}{\partial T^i}, \quad \theta \in [\underline{\theta}, \overline{\theta}], \quad i \in \{H, L\}. \tag{60}$$

The average values for the social marginal value of income for high- and low-skilled workers are given by:

$$\overline{\gamma^L} \equiv \frac{\int_\Phi^\Theta \gamma^L_\theta \mathrm{d}G(\theta)}{G(\Theta) - G(\Phi)}, \quad \overline{\gamma^H} \equiv \frac{\int_\Theta^{\overline{\theta}} \gamma^H_\theta \mathrm{d}G(\theta)}{1 - G(\Theta)}. \tag{61}$$

Finally, we define the distributional characteristics of the income tax bases as ξ^i (cf.

Atkinson and Stiglitz, 1980, p.388):

$$\xi^L \equiv 1 - \frac{(G(\Theta) - G(\Phi)) \int_\Phi^\Theta w^L l_\theta^L \gamma_\theta^L \mathrm{d}G(\theta)}{\int_\Phi^\Theta w^L l_\theta^L \mathrm{d}G(\theta) \int_\Phi^\Theta \gamma_\theta^L \mathrm{d}G(\theta)} = -\frac{\mathrm{cov}\left[w^L l_\theta^L, \gamma_\theta^L\right]}{\overline{w^L l^L \gamma^L}} \tag{62}$$

$$\xi^H \equiv 1 - \frac{(1 - G(\Theta)) \int_\Theta^{\overline{\theta}} w^H l_\theta^H \gamma_\theta^H \mathrm{d}G(\theta)}{\int_\Theta^{\overline{\theta}} w^H l_\theta^H \mathrm{d}G(\theta) \int_\Theta^{\overline{\theta}} \gamma_\theta^H \mathrm{d}G(\theta)} = -\frac{\mathrm{cov}\left[w^H l_\theta^H, \gamma_\theta^H\right]}{\overline{w^H l^H \gamma^H}} \tag{63}$$

That is, ξ^i gives the negative of the normalized covariance between gross labor income and the social marginal value of income for skill type i. The larger is the term ξ^i, the more the social welfare weight γ_θ^i is decreasing with income $w^i l_\theta^i$, making the labor-income tax a more attractive instrument to redistribute income. Due to the normalization, the distributional term takes on a value between zero and one: $\xi^i \in [0, 1]$.

This, then, allows us to rewrite the first-order conditions for T^U, T^L, ω^L, T^H, and ω^H, respectively, as follows:

$$\gamma^U = 1 + \Delta_\Phi \eta_{T^U}, \tag{64}$$

$$\overline{\gamma^L} = 1 - \Delta_\Phi \eta_{T^L} + \Delta_\Theta \zeta_{T^L}, \tag{65}$$

$$(1 - \xi^L)\overline{\gamma^L} = 1 - \Delta_L \overline{\varepsilon_{\omega^L}^L} - \Delta_\Phi \eta_{\omega^L} + \Delta_\Theta \zeta_{\omega^L}, \tag{66}$$

$$\overline{\gamma^H} = 1 - \Delta_\Theta \zeta_{T^H}, \tag{67}$$

$$(1 - \xi^H)\overline{\gamma^H} = 1 - \Delta_H \overline{\varepsilon_{\omega^H}^H} - \Delta_\Theta \zeta_{\omega^H}. \tag{68}$$

Notice that if social welfare weights are decreasing in ability θ, we can write $\gamma^U > \overline{\gamma^L} > \overline{\gamma^H}$. This allows us to combine the first, second and fourth first-order conditions to find:

$$\Delta_\Phi \left(\eta_{T^U} + \eta_{T^L}\right) > \Delta_\Theta \zeta_{T^L}, \tag{69}$$

$$\Delta_\Phi \eta_{T^U} > -\Delta_\Theta \zeta_{T^H}, \tag{70}$$

$$\Delta_\Theta \left(\zeta_{T^L} + \zeta_{T^H}\right) > \Delta_\Phi \eta_{T^L}. \tag{71}$$

From the first line we establish that if $\Delta_\Phi < 0$, then $\Delta_\Theta < 0$. However, from the second line we see that if $\Delta_\Phi < 0$, then $\Delta_\Theta > 0$. This is a contradiction. Thus, in the optimum we must have $\Delta_\Phi > 0$. Notice, from the definition of Δ_u in equation (35), that if $\Delta_\Phi > 0$, we must necessarily have that $\Delta_u > 0$. From the third line, it immediately follows that $\Delta_\Theta > 0$. Combining the remaining first-order conditions, it is relatively straightforward to show that the optimal wedges on high- and low-skilled labor earnings are positive as well: $\Delta_H > 0$ and $\Delta_L > 0$. Furthermore, since individuals' marginal utility of income is strictly positive, we can readily deduce from the first-order conditions that all wedges are finite. This proves Lemma 2.

A3. Efficient rationing: Lee and Saez (2012)

We derive the desirability condition for the minimum wage when rationing is 'efficient', i.e., when the involuntary unemployed are those that have the smallest benefits of being low-skilled employed. In this case, the rationing schedule is no longer continuous. Instead, rationing is concentrated on the extremes of the low-skilled ability distribution, Φ and Θ. Individuals with ability Φ or Θ are indifferent between low-skilled employment and their outside option – voluntary unemployment for Φ, high-skilled employment for Θ. In response to a NIN minimum-wage reform, $\mathrm{d}w^L$, the change in welfare is therefore nil: $\mathrm{d}\mathcal{W} = 0$. Hence, the desirability condition of a NIN minimum-wage reform simplifies to $\mathrm{d}\mathcal{B} > 0$.

In the case of efficient rationing, the minimum-wage induced increase in high-skilled employment, relative to the increase in unemployment, is identical to the increase in high-skilled employment, relative to the increase in voluntary unemployment, due to a decrease in the low-skilled transfer T^L. In response to a NIN minimum-wage reform, $\mathrm{d}w^L$, this yields:

$$\frac{-\mathrm{d}G(\Theta)/\mathrm{d}w^L}{\int_\Phi^\Theta \mathrm{d}u_\theta \mathrm{d}G(\theta)/\mathrm{d}w^L} = \frac{-\mathrm{d}G(\Theta)/\mathrm{d}T^L}{\mathrm{d}G(\Phi)/\mathrm{d}T^L}. \tag{72}$$

Since only individuals with ability Φ or Θ are rationed, there is no direct utility loss of unemployment and the loss of tax revenue is solely determined by the earnings of a worker with ability Φ. Hence, we can write Δ_u as:

$$\Delta_u = \Delta_\Phi \equiv t^L w^L l_\Phi^L - T^L + T^U, \tag{73}$$

where Δ_Φ is the wedge on participation. Substituting above two equations, together with the elasticities as defined in (58) and (59), into the desirability condition for a NIN minimum-wage increase, (36), we get:

$$\frac{\zeta_{T^L}}{\eta_{T^L}} > \frac{\Delta_\Phi}{\Delta_\Theta}. \tag{74}$$

The values for the wedges on participation and skill formation, Δ_Φ and Δ_Θ, in the tax optimum without a minimum wage, are given by the first-order condition (65).

Substituting this condition into above desirability condition yields:

$$\overline{\gamma^L} > 1. \tag{75}$$

Here, $\overline{\gamma^L}$ is the low-skilled average marginal social value of income as defined in the proof of Lemma 2. The term $\overline{\gamma^L}$ exactly corresponds with g_1 in Lee and Saez (2012). Hence, we can see that our desirability condition is equivalent to their Proposition 2. A binding minimum wage is optimal if and only if the social marginal welfare weight of low-skilled workers ($\overline{\gamma^L}$ or g_1) exceeds unity. In the case of efficient rationing, the only difference between a NIN minimum-wage increase and a low-skilled transfer decrease, is that the minimum-wage increase does not redistribute away from the low-skilled. Hence, if redistribution away from the low-skilled is socially costly, a binding minimum wage is part of the policy optimum.

A4. The effect of unemployment on school enrollment

Table A1 gives an overview of empirical studies on the impact of unemployment on school enrollment, as discussed in the paper. Earlier UK evidence is surveyed in a similar overview by Clark (2011), in his Table 1. The first column indicates the study of interest; the second column indicates the country of analysis; the third column indicates the time span of the analysis; the fourth column indicates whether the schooling variable refers to enrollment rates (E), or high-school graduation rates (G), whether it refers to boys (b), girls (g), or both (bg), and the age group under consideration; the fifth column indicates to which age-group the unemployment variable refers; the final column gives the estimate of η.

Table A1: Empirical estimates of η

Study	Country	Time	Schooling	Unemployment	η
Clark (2011)	UK	1975-2005	E, b, 16 y/o	18-19 y/o	0.32
			E, g, 16 y/o	18-19 y/o	0.45
Tumino and Taylor (2013)	UK	1991-2009	E, bg, 16 y/o	16-21 y/o	0.48
Kane (1994)	US	1973-1988	G, bg, 18-19 y/o	Total	0.60
Card and Lemieux (2001)	US	1968-	E, bg, 15-16 y/o	25-54 y/o	0.14
		1996	E, bg, 17 y/o	25-54 y/o	0.40
		1954-	G, b	25-54 y/o	0.17
		1964	G, g	25-54 y/o	0.18
Petrongolo and San Segundo (2002)	ES	1991	E, b, 16-17 y/o	16-24 y/o	0.44

The column on schooling indicates whether the dependent variable was the enrollment rate (E) or the high-school graduation rate (G), whether it concerned boys (b), girls (g), or both (bg), and the age group to which the schooling variable refers. Note that Kane (1994) uses the graduation-rate of blacks. Estimates of Clark (2011) are found in his Tables 2 and 3 on pages 533-534. Estimates for Tumino and Taylor (2013) are found in their Table 3 on page 22. Estimates for Kane (1994) are found in his text, page 890. Estimates for Card and Lemieux (2001) are found in their table 9.4 on page 467 and table 9.6 on page 471. Estimates for Petrongolo and San Segundo (2002) are found in their text, page 364.

References

Allen, Stephen P. 1987. "Taxes, redistribution, and the minimum wage: A theoretical analysis." *Quarterly Journal of Economics* 102 (3):477–489.

Atkinson, Anthony B. and Joseph E. Stiglitz. 1980. *Lectures on Public Economics*. New York/Singapore: McGraw-Hill.

Autor, David H, Alan Manning, and Christopher L. Smith. 2010. "The contribution of the minimum wage to U.S. wage inequality over three decades: a reassessment." NBER Working Paper 16533.

Card, David. 1999. "The causal effect of education on earnings." In *Handbook of Labor Economics*, vol. 3, edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier, 1801–1863.

Card, David and Thomas Lemieux. 2001. "Dropout and enrollment trends in the postwar period: What went wrong in the 1970s?" In *Risky behavior among youths: An economic analysis*, edited by Jonathan Gruber. University of Chicago Press, 439–482.

Clark, Damon. 2011. "Do Recessions Keep Students in School? The Impact of Youth Unemployment on Enrolment in Post-compulsory Education in England." *Economica* 78 (311):523–545.

Diamond, Peter A. 1975. "A many-person Ramsey tax rule." *Journal of Public Economics* 4 (4):335–342.

Gerritsen, Aart and Bas Jacobs. 2013. "Minimum wages and taxation in competitive labor markets with endogenous skill formation." Mimeo, Erasmus University Rotterdam.

Guesnerie, Roger and Kevin Roberts. 1987. "Minimum wage legislation as a second best policy." *European Economic Review* 31 (1-2):490–498.

Hungerbühler, Mathias and Etienne Lehmann. 2009. "On the optimality of a minimum wage: New insights from optimal tax theory." *Journal of Public Economics* 93 (3):464–481.

Kane, Thomas J. 1994. "College entry by blacks since 1970: The role of college costs, family background, and the returns to education." *Journal of Political Economy* 102 (5):878–911.

Kaplow, Louis. 2008. *The Theory of Taxation and Public Economics*. Princeton University Press.

Lee, David and Emmanuel Saez. 2012. "Optimal minimum wage policy in competitive labor markets." *Journal of Public Economics* 96 (9–10):739–749.

Lott, John R. 1990. "Nontransferable rents and an unrecognized social cost of minimum wage laws." *Journal of Labor Research* 11 (4):453–460.

Luttmer, Erzo F. P. 2007. "Does the Minimum Wage Cause Inefficient Rationing?" *BE Journal of Economic Analysis & Policy* 7 (1). (Contributions, Article 49).

Nickell, Stephen. 1979. "Education and lifetime patterns of unemployment." *Journal of Political Economy* 87 (5):S117–S131.

OECD. 2011a. "Education at a Glance 2011: OECD Indicators." OECD Publishing, Paris.

———. 2011b. "Taxation and Employment." OECD Tax Policy Studies No. 21, OECD Publishing, Paris.

Palda, Filip. 2000. "Some deadweight losses from the minimum wage: The cases of full and partial compliance." *Labour Economics* 7 (6):751–783.

Petrongolo, Barbara and Marıa J. San Segundo. 2002. "Staying-on at school at 16: The impact of labor market conditions in Spain." *Economics of Education Review* 21 (4):353–365.

Stiglitz, Joseph E. 1982. "Self-selection and Pareto efficient taxation." *Journal of Public Economics* 17 (2):213–240.

Teulings, Coen N. 2000. "Aggregation bias in elasticities of substitution and the minimum wage paradox." *International Economic Review* 41 (2):359–398.

———. 2003. "The contribution of minimum wages to increasing wage inequality." *The Economic Journal* 113 (490):801–833.

Tumino, Alberto and Mark Taylor. 2013. "The impact of local labour market conditions on school leaving decisions." Mimeo.