Minimum wages and taxation in competitive labor markets with endogenous skill formation

Aart Gerritsen and Bas Jacobs*

September 2013

The effect of minimum wages on skill formation is ambiguous, since a minimum wage (i) reduces skill formation by lowering the skill premium, and (ii) boosts skill formation by raising low-skilled unemployment. We demonstrate that, under proportional income taxation, the latter effect dominates if the elasticity of substitution between high-skilled and low-skilled labor is larger than one, which seems to be the empirically plausible case. We further demonstrate how minimum wages interact with the government's tax instruments. In the case of proportional income taxes, minimum wages may be used to correct for the distributional imperfections of the tax system. If income taxes can be conditioned on skill type, taxation can yield the same distributional consequences as a minimum wage. A minimum wage reform then differentiates itself from a distributionally equivalent tax reform by generating both higher unemployment and higher skill formation. A binding minimum wage is optimal if the welfare benefits of the latter effect outweigh the welfare costs of the former.

Keywords: Minimum wage, skill formation, optimal taxation

JEL codes: J38; J24; H30

*Gerritsen: Erasmus University Rotterdam and Tinbergen Institute, PO Box 1738, 3000 DR Rotterdam, The Netherlands, agerritsen@ese.eur.nl. Internet: https://sites.google.com/site/aartgerritsen/. Jacobs: Erasmus University Rotterdam, Tinbergen Institute, Netspar, and CESifo, PO Box 1738, 3000 DR Rotterdam, The Netherlands, bjacobs@ese.eur.nl. Internet: http://people.few.eur.nl/bjacobs/. We are grateful to The Netherlands Organization for Scientific Research for financial support (NWO Vidi Grant No. 452-07-013).

The desirability of a minimum wage has been fiercely debated by both policy makers and academics. Proponents emphasize that a minimum wage leads to a higher income for low-income employees. Opponents mainly stress that it leads to higher unemployment rates as workers with productivity levels below the minimum wage find themselves unable to secure a job. As of yet, this debate has not been settled. Minimum wages were an important topic in the 2009 and 2013 federal elections in Germany, one of nine countries within the Organisation for Economic Co-operation and Development (OECD) without a statutory minimum wage (Immervoll, 2013), as well as in the American and French presidential elections of 2012. As noted by Cahuc and Laroque (2013), the OECD itself changed its appraisal of a minimum wage at least twice in the 1990s. The empirical literature on the effects of a minimum wage likewise seems to lack consensus. Some surveys report employment effects of a minimum wage to be absent or even positive (e.g., Card and Krueger, 1995), while in a more recent survey, Neumark and Wascher (2006) argue that the vast majority of the evidence points to a negative employment effect, albeit not always statistically significant.

The evidence on the effect of minimum wages on human capital investments, whether it concerns adolescent education or on-the-job training, seems to be even more ambiguous. As we argue below, it is *a priori* unclear how minimum wages affect human capital investments. On the one hand, a higher minimum wage drives down the skill premium, thereby undermining incentives to develop skills. On the other hand, if mainly low-skilled wages are affected by the minimum wage, it will lead to higher unemployment among the low-skilled, thereby providing more incentives to develop skills so as to avoid involuntary unemployment.

The purpose of this paper is to contribute to our understanding of the economic effects of minimum wages in two ways. From a positive perspective we determine how minimum wages affect the incentives to acquire skills, and identify conditions under which higher minimum wages lead to more skill formation. From a normative perspective we aim to contribute to the debate on the desirability of minimum wages by analyzing whether minimum wages are part of an optimal redistributive policy when skill formation is endogenous - and how this depends on the effect of minimum wages on skill formation. Importantly, we allow for income taxes as an alternative instrument to redistribute income.

We develop a general equilibrium model with perfectly competitive labor markets. Firms demand both low-skilled and high-skilled labor. Individuals are heterogeneous with respect to their disutility of work, and optimally decide first to become either low-skilled or high-skilled, second how much labor hours to supply. Individuals with little disutility of work have both an absolute and comparative advantage of working in high-skilled jobs, and thus end up becoming high-skilled, whereas high-disutility individuals become low-skilled. Minimum wages are binding for the low-skilled market segment, causing involuntary unemployment among the low-skilled only. As such, a minimum wage simultaneously destimulates skill formation, by boosting low-skilled wages, and stimulates skill formation through higher unemployment. The government maximizes a social welfare function that features redistributive concerns. Due to informational constraints individualized lump-sum taxes are ruled out, such that the government needs to resort to distortionary income taxation and minimum wages to achieve its redistributive goals.

We demonstrate that the net effect of a minimum wage on skill formation critically depends on the substitutability of high-skilled and low-skilled labor in the production function. Intuitively, if substitutability is high, a given increase in the minimum wage will cause firms to strongly substitute away from low-skilled labor, leading to a large increase in unemployment. If the substitutability is high enough, the increase in unemployment will outweigh the increase in the skill premium, and skill formation will rise. More specifically we show that in the absence of skill-dependent taxes and transfers, a minimum wage leads to more skill formation if the elasticity of substitution is larger than one, which seems to be the empirically plausible case.

The welfare effects of a minimum wage are studied in three different cases which are progressively more complex in the government's instrument set. First, we determine the desirability of a minimum wage in the absence of income taxation, second in the presence of skill-independent taxation, and third in the presence of skill-dependent taxation. In the absence of taxation, the social welfare gains of a minimum wage are a higher degree of income equality between low-skilled and high-skilled workers, and a higher degree of income equality among high-skilled workers as a higher minimum wage leads to lower high-skilled wages through general equilibrium effects. The social welfare losses of a minimum wage are given by higher inequality among low-skilled workers and the utility losses of laid off workers. Whether the gains outweigh the losses, and thus whether a minimum wage is optimally implemented, is ambiguous and crucially depends on initial inequalities, social redistributive preferences, and the minimum-wage elasticity of unemployment.

When the government sets taxes that are not conditioned on skill type, a minimum wage tends to lead to additional welfare losses as increased unemployment erodes the income tax base and therefore reduces tax revenue. In this case, taxes cannot be targeted well to deal with both inequality within skill groups and between high- and low-skilled workers. There might therefore be a role for the minimum wage in its capacity to redistribute income between skill groups if, for a given amount of redistribution, the welfare costs associated with a higher minimum wage (utility and tax revenue losses from higher unemployment) are sufficiently smaller than the welfare costs associated with higher income taxes (tax revenue losses from lower intensive labor supply). Minimum wages can in that case be seen to correct for the distributional imperfection of taxes that cannot be conditioned on skill type.

When income taxes and transfers can be conditioned on skill type, the government can use its tax instruments to emulate the distributional consequences of a minimum wage. That is, decreasing taxes on low-skilled income and increasing taxes on high-skilled income results in a higher net income for the low-skilled and lower net income for the high-skilled, just as a minimum wage would. This moreover leads to the same degree of distortion on the intensive labor supply margin. A minimum wage thus only differs from a distributionally equivalent tax-rate adjustment by causing higher unemployment and, as a direct result of this, more skill formation. Higher unemployment leads to utility losses and an erosion of the tax base. Higher skill formation, on the other hand, constitutes a welfare gain through higher tax revenues, provided that taxes are set progressively. A minimum wage is desirable if the benefits of higher skill formation outstrip the costs of higher unemployment.

In most of our paper, we assume that that every low-skilled worker has an equal probability of becoming unemployed. The uncomfortable fact is that we do not really know in what way employment is decreased due to a minimum wage. We therefore also study a separate case in which unemployment is 'efficient,' implying that hours of work, rather than jobs, are rationed. In that case, there is no first-order utility loss associated with the unemployment caused by marginally binding minimum wage. This ensures that a minimum wage is always optimal in the absence of taxation. However, in the presence of skill-independent taxes and transfers, the optimality of a minimum wage is still ambiguous. In the presence of skill-dependent taxes and transfers, a minimum wage is redundant as it can be perfectly mimicked by taxation.

Our paper is structured as follows. Section 1 is devoted to a discussion of relevant literature. Section 2 introduces the theoretical model, the comparative statics of which are derived in Section 3. In Sections 4 and 5 we discuss the welfare effects of a minimum wage in the presence of skill-independent and skill-dependent tax instruments, respectively. In Section 6 we study the case of efficient rationing. We close the paper with some concluding remarks.

1 Related literature

1.1 Theory

There are roughly two approaches to studying the implications of a minimum wage. One strand of the literature takes certain market imperfections in the labor market as given and determines how a minimum wage affects efficiency, employment, and/or social welfare. A popular assumption is that employers have a degree of monopsony power over wages, leading to inefficiently low wages. The classical argument is due to Robinson (1933). Indeed, it is straightforward to show that, in a partial equilibrium setting with a monopsony in the labor market, wages and employment are set inefficiently low. A binding minimum wage might in that case be employed to ensure an efficient outcome. However, confronting this argument with a more realistic context significantly complicates the analysis. As Stigler (1946) argues, the optimum wage will vary with occupation, among firms and, often rapidly, through time. Therefore, "[a] uniform national minimum wage, infrequently changed, is wholly unsuited to these diversities of conditions". More recent studies bring further nuance to the discussion. For example, Manning (2003), focussing on employment, considers a general equilibrium model with heterogeneous firms and concludes that a minimum wage might have opposing employment effects for different firms, leaving the aggregate employment effect ambiguous. Bhaskar and To (1999) consider monopsonistic competition with exit and entry of firms, firm-specific job types and heterogeneous preferences for job types, and reach a similar conclusion. While a minimum wage increases employment per firm, it also forces some firms to exit the market, leaving aggregate employment and welfare outcomes ambiguous. Still, as Cahuc and Laroque (2013) show, with a sufficiently rich set of tax instruments the government can always reach the second-best competitive allocation without any need to resort to minimum wages.

The minimum wage is also studied in frameworks combining monopsony power with other market imperfections. For example, Rebitzer and Taylor (1995) study a model in which firms imperfectly monitor their employees and therefore set efficiency wages to motivate them not to shirk. If higher labor supply leads to costlier monitoring, they show that a minimum wage will increase employment over the short term, with ambiguous results over the long term. Cahuc, Saint-Martin and Zylberberg (2001) introduce a model where high- and low-skilled wages are bargained over between employers and unions that represent high-skilled workers. They show that a higher minimum wage might reduce the unions' bargaining power over the high-skilled wage, potentially leading to more employment for both low-skilled and high-skilled workers through general equilibrium effects. Flinn (2006) analyzes a matching model of the labor market and argue that if workers' bargaining power is too low for the Hosios (1990) efficiency condition to hold, a minimum wage might function as a crude measure to push labor market outcomes towards efficiency. In a similar vain, Hungerbühler and Lehmann (2009) find a binding minimum wage might be part of an optimal redistributive policy as an indirect way to increase workers' bargaining power. This holds even if government makes use of non-linear taxation to achieve its redistributive goals.

The second strand of the literature, which is closer in spirit to the present study, applies an optimal taxation framework to competitive labor markets and heterogeneous workers with either continuous skill types as in Mirrlees (1971) or, more often, two skill types as in Stiglitz (1982). In the latter tradition, Allen (1987) and Guesnerie and Roberts (1987) show that a minimum wage might be optimal as part of a redistributive policy if government is confined to linear taxation only. However, if non-linear taxation is available, a minimum wage will never be optimal because it will make it more attractive for high-skilled workers to imitate the low-skilled, thereby tightening the incentive compatibility constraint. This approach has been criticized by Lee and Saez (2012) on informational grounds. They argue that a government needs to be able to discern high-skilled from low-skilled workers in order to enforce a minimum wage, thereby making incentive compatibility constraints irrelevant. Marceau and Boadway (1994) and Boadway and Cuff (2001), who are liable to the same criticism, combine a minimum wage with unemployment insurance and find that a minimum wage might still be optimal in combination with a non-linear tax schedule. Assuming that individuals can only apply for unemployment benefits if they are unable to find a job, Boadway and Cuff (2001) show that a minimum wage provides information about the bottom of the skill distribution, which can not be obtained by merely using taxes.

Almost any study takes skill levels of individuals as exogenously given. Two exceptions with endogenous skill formation on the extensive margin are Saint-Paul (1996) and Cahuc and Michel (1996). In Saint-Paul (1996), as in our model, an increase in low-skilled unemployment causes more individuals to become high-skilled. As he assumes perfect substitutability between high- and low-skilled labor, higher low-skilled unemployment might thereby lead to lower labor productivity and to even higher levels of unemployment in the case of real wage rigidity. The implementation of a binding minimum wage might thereby induce increasing returns to education and soaring low-skilled unemployment rates of up to a hundred percent. As we assume imperfect substitutability, such an extreme result is not attainable in our model. Cahuc and Michel (1996) develop an overlapping generations model with a high-skilled and a low-skilled production sector. Furthermore, high-skilled production exhibits positive externalities and hence serves as a catalyst of endogenous growth. They show that if a minimum wage increases human capital formation, this can lead to higher growth.[1] Our model exhibits similar extensive skill-formation as in Cahuc and Michel, though we analyze the effects of a minimum wage in an optimal taxation setting without externalities.

[1]Naturally, as in the case of monopsonistic labor markets, this begs the question why the externalities are not internalized by appropriately set taxes.

Lee and Saez (2012) and Gerritsen and Jacobs (2013) are particularly closely related to the current study as they analyze the optimality of a minimum wage alongside optimal taxes and transfers in models with two skill types and competitive labor markets. Lee and Saez study the case in which rationing is efficient, such that new entrants are unable to find a job in a rationed low-skilled labor market. In that case, a binding minimum wage might be optimal to implement as it effectively mutes the distortionary effects of a transfer towards low-skilled workers. In Gerritsen and Jacobs (2013), we derive a general optimality condition for a binding minimum wage that hold for any arbitrary rationing schedule, including but not restricted to efficient rationing. Calibration of this condition shows that a minimum wage decrease yields a Pareto-improvement in all countries under consideration, except possibly the United States. The current study distinguishes itself from Lee and Saez (2012) mainly by its focus on uniform rationing – i.e., a common probability of unemployment for every low-skilled worker – and from Gerritsen and Jacobs (2013) by its focus on the skill formation and social welfare effects of a binding minimum wage under varying tax regimes.

1.2 Informational inconsistency

Following the seminal analysis of Mirrlees (1971), modern literature on public finance builds on the assumption that the fundamental source of heterogeneity across individuals is their wage rates, or earning ability, and that this is private information. Government can only observe labor earnings, which is the product of the wage rate and the total number of hours worked. Since tax policy can only be conditioned on observables, and therefore not on earnings ability, a first-best outcome cannot be attained. However, to be able to implement a minimum wage, government must observe individual wage rates. This leads to the problem of *informational inconsistency*: in the Mirrlees (1971) framework it contradicts the assumption that wage rates are private information and are thus not verifiable by the government. Indeed, information on individual wage rates theoretically enables government to reach any desired redistribution without efficiency losses by implementing individualized lump-sum taxes and transfers. Consequently, studies that use the Mirrlees framework for the analysis of minimum wages usually make an – often implicit – *ad hoc* assumption that information required for the implementation of minimum wages cannot be used for taxes and transfers. Guesnerie and Roberts (1987, p.498), somewhat euphemistically, remark that "this is a somewhat mixed observability assumption."

The informational inconsistency appears when the source of heterogeneity can be defined in terms of observable variables. In the standard Mirrlees exercise, for example, exogenous earning ability can be defined as $n \equiv \frac{z}{l}$, z labor income, and l the number of hours worked. First best is not attainable because only z, and not l, is observable. If a minimum wage is introduced in this framework as, for example, in Boadway and Cuff (2001), first best is attainable since the wage rate must be observable and, obviously, the exogenous ability can be defined in terms of the wage rate $n \equiv w$.

The exact same inconsistency occurs when introducing minimum wage legislation in a model based on Stiglitz (1982), see for example Allen (1987), Marceau and Boadway (1994), Cahuc and Michel (1996), Aronsson and Koskela (2008), and Danziger and Danziger (2010). They all assume, contrary to Mirrlees, that workers with different wage rates are imperfect substitutes. Still, because in these models wage rates are generally exogenously given and the sole source of heterogeneity, we can again write $n \equiv w$, which implies first best is attainable once government can observe wage rates. This result suggests that to be informationally consistent, we need to direct attention away from models in which the source of heterogeneity can be defined by the wage rate.

One way to do this is to introduce a labor-effort decision alongside an hours-of-work decision. Denoting labor effort as e, we can define the wage rate $w \equiv en$, or alternatively, earnings ability as $n \equiv \frac{w}{e}$. As long as government cannot observe effort, exogenous ability cannot be defined by observables only, and first best is not attainable. This approach is taken by Deltas (2007).

An alternative approach is taken by Lee and Saez (2012). The model of Lee and Saez includes multiple job types and individuals supply one unit of labor if employed ($l = 1$). Thus, earnings are given by $z \equiv w$, which is verifiable by the government so that it can enforce a minimum wage and set income-tax policy. To avoid a first-best outcome without violating informational consistency, individuals are assumed to be heterogeneous with respect to their costs of participation in a particular job, which are unrelated to earnings and thus unobserved. These costs of participation, θ, cannot simply be defined in terms of observables, making the first-best allocation infeasible. In the model below, we adopt, like Lee and Saez, disutility of work as the fundamental source of heterogeneity across individuals, safeguarding us from informational inconsistency.[2]

1.3 Empirics

The effect of minimum wages on employment has recently been surveyed by Neumark and Wascher (2006). Most studies find that minimum wages reduce employment, although the estimates are not always significant. When it comes to the effect of minimum wages on skill formation, empirical results are much scarcer and more ambiguous. A number of potential effects of higher minimum wages on skill formation are recognized. When minimum wages lead to a compression of wages, the net return of human capital investments will drop, leading to lower skill formation. However, if employment opportunities decline for low-skilled jobs, skill formation might be boosted in order to avoid unemployment. These arguments hold for investments in education and, perhaps to a lesser degree, for employee-financed on-the-job training. However, for on-the-job training additional arguments play a role. On the one hand, employees might finance their training by accepting a lower hourly wage rate, the possibility of which is diminished by a higher minimum wage (e.g., Rosen, 1972). On the other hand, if training is firm-sponsored and labor markets are not perfectly competitive, a minimum wage might decrease the rents on low-skilled labor, leading to more investment in on-the-job training such that firms can obtain higher rents (e.g., Acemoglu and Pischke, 1999).

The studies that try to capture the effect of minimum wages on skill formation can be divided in those that explain school enrollment and related variables and those that explain on-the-job training. The evidence on either of the two human capital variables is scarce and ambiguous. Moreover, most studies are unsuited to evaluate the distinct effects of minimum wages on skill formation – i.e., through a compressed wage structure and through higher unemployment – and analyze

[2]Nevertheless, in Section 4 we do study the social welfare effects of minimum wages in the case that the government does not condition its tax instruments on skill type, and thus does not fully use the information at hand.

which effect dominates. Empirical studies of school enrollment are often flawed in this respect because they usually control for unemployment, such that estimates only show the direct minimum wage effect through the wage structure.[3] Studies of on-the-job training are often confounded because minimum wages can impact the training decision in many different ways as it is usually a joint decision of employer and employee, each with their own incentives. Data seem to be too scarce to adequately take account of the different incentives.[4] Hence, amongst empirical ambiguity, we hope to contribute to our understanding of minimum wage legislation by theoretically identifying under what circumstances a minimum wage leads to more or less skill formation.

2 Model

We assume a unit mass of individuals and two job-types: high-skilled jobs and low-skilled jobs. We assume that wages on the high-skilled labor market are perfectly flexible to assure there is no unemployment among high-skilled workers. Government might, however, impose a binding minimum wage on the market for

[3]Studies that find that higher minimum wages lead to less schooling if controlling for the unemployment rate, include Cunningham (1981); Neumark and Wascher (1995*a*,*b*, 2003); Landon (1997); Chaplin, Turner and Pape (2003); Montmarquette, Viennot-Briot and Dagenais (2007); Pacheco and Cruickshank (2007). Interestingly, Cunningham only controls for the white unemployment rate and finds the schooling effect of minimum wages reversed for black youths. Similarly, Pacheco and Cruickshank find that the negative effect of higher minimum wages on enrollment rates is significant at a level of 1 percent when controlling for the unemployment rate, but only significant at a level of 10 percent if not. A number of studies do not find a significant effect of minimum wages on education, even when controlling for unemployment, see Ragan (1977); Ehrenberg and Marcus (1982); Card (1992); Crofton, Anderson and Rawe (2009). Only Mattila (1981) finds a positive effect of minimum wages on education, although this might be caused by the fact that she controls for the unemployment rate among people aged 35-44, which might be fairly irrelevant for students deciding whether to enroll for school. These findings suggest the importance of distinguishing the distinct effects of a minimum wage on skill formation. We express our hopes that future empirical research will give due attention to minimum wage effects stemming from a compressed wage structure and minimum wage effects stemming from higher unemployment.

[4]This is apparent in the contradictory findings. Negative effects of minimum wages on training are found by Hashimoto (1982); Schiller (1994); Neumark and Wascher (2001). Positive effects are found by Arulampalam, Booth and Bryan (2004); Dustmann and Schönberg (2009). Insignificant, or non-robust findings are presented by Mincer and Leighton (1981); Grossberg and Sicilian (1999); Acemoglu and Pischke (1999); Fairris and Pedace (2004).

low-skilled labor. Unemployment will therefore be concentrated on the group of low-skilled workers. Thus, workers can either be unemployed low-skilled, employed low-skilled, or employed high-skilled workers. The fractions of each are denoted by N^U, N^L, and N^H, respectively ($N^U + N^L + N^H = 1$). For short, we will denote the unemployed low-skilled as the unemployed. Similarly, the employed low-skilled workers are referred to as low-skilled workers. Type-specific variables are indexed with superscripts U, L, and H.

We assume that workers are heterogeneous with respect to their ability, θ. Rather than making the assumption, common in the optimal tax literature, that θ reflects the productivity per hour worked, we instead assume that θ reflects the utility cost per hour worked.[5] A higher θ implies that utility costs per hour worked are lower. This assumption is similar to Lee and Saez (2012), who also assume that more able workers have lower costs of participation, rather than higher labor productivity. Moreover, we assume that individuals with a higher ability enjoy a comparative advantage of performing high-skilled work, whereas individuals with a low level of ability enjoy a comparative advantage for low-skilled work. θ has support $[\underline{\theta}, \overline{\theta}]$ and follows a cumulative distribution function $G(\theta)$ with corresponding density function $g(\theta)$. We assume that $\underline{\theta} > 0$, which in our model implies that in the absence of unemployment insurance, individuals prefer being employed over being unemployed.

Individuals decide on the number or working hours and on whether to invest in human capital. The number of working hours is chosen to maximize utility, which is increasing in income and decreasing in the number of hours worked. Since more able high-skilled individuals have a lower cost of work, they supply more labor for a given wage rate. Hours worked is denoted by l^L for low-skilled workers and by l^H for high-skilled workers. Human capital investment is made on the extensive margin, i.e., it is a discrete decision to become a skilled worker. The skilled wage is denoted by w^H, and the unskilled wage rate is denoted by w^L. If the individual

[5]Due to this assumption we manage to avoid the informational inconsistency to which we alluded in Section 1. It more over ensures that hourly labor earnings are constant within groups and that a non-negligible share of the population earns the minimum wage, which conforms with reality but cannot be the case if θ would reflect labor productivity. That is, had θ reflected exogenously given marginal productivity, any person with θ above the minimum wage would be hired and the mass of workers earning the minimum wage would be zero.

invests in human capital, he earns $w^H l^H$, if not, he earns $w^L l^L$.

2.1 Individual optimization

Utility is denoted by V and is assumed to be separable and quasi-linear in consumption and working hours. Moreover, it exhibits a constant labor supply elasticity on the intensive margin, ε, which is assumed to be equal for both low and high-skilled workers.

Initially, we assume that tax instruments are not differentiated according to skill type. Moreover, we restrict attention to linear instruments throughout the paper. Hence, labor income is taxed at a common rate, t. In addition, tax revenue is rebated as a non-individualized lump-sum transfer, T. Later we explore the robustness of our results by allowing for skill-dependent instruments. Thus, with skill-independent tax policy, utility when unemployed, low skilled, and high skilled are given by:

$$V^U \equiv T, \tag{1}$$

$$V_\theta^L \equiv T + (1-t)w^L l^L - \frac{1}{\theta^\beta}\frac{(l^L)^{1+1/\varepsilon}}{1+1/\varepsilon}, \tag{2}$$

$$V_\theta^H \equiv T + (1-t)w^H l^H - \frac{1}{\theta}\frac{(l^H)^{1+1/\varepsilon}}{1+1/\varepsilon}. \tag{3}$$

Variables that depend on ability are denoted by a subscript θ. Note that there is no disutility of labor for unemployed workers, since they do not work. For employed workers, the marginal costs of labor supply are inversely related to ability, θ. As we assume that $\beta \in (0,1)$, individuals with a higher ability have a comparative advantage in doing high-skilled work. The higher is ability, the lower are the costs of labor effort in high-skilled jobs relative to the costs of labor effort in low-skilled jobs. This comparative advantage is stronger if β is lower. Since marginal utility of consumption is constant, households are risk-neutral with respect to the probability of becoming unemployed.[6] Each worker first decides to invest in human

[6] Allowing for risk-aversion would strengthen our result that a minimum wage leads to higher human capital accumulation (if the substitution elasticity is larger than one), see below. In that case, unemployment does not only raise skill formation by lowering expected utility of being low skilled, but also by increasing the variance in low skilled earnings. However, if unemployment is

capital or not, and then, given the skill level, they optimally supply labor. We solve this optimization problem backwards.

Optimal labor supply for high- and low-skilled employed workers is given by:

$$l_\theta^H = (\theta(1-t)w^H)^\varepsilon, \tag{4}$$

$$l_\theta^L = (\theta^\beta(1-t)w^L)^\varepsilon. \tag{5}$$

Labor supply is an increasing function of the gross wage rate, decreasing in the tax rate and increasing with ability, θ. There are no income effects in labor supply, which facilitates the analysis considerably. Substituting these expressions into the utility functions for low-skilled and high-skilled workers yields the following indirect utility functions:

$$V^U = T, \tag{6}$$

$$V_\theta^L = T + \frac{\theta^{\beta\varepsilon}((1-t)w^L)^{1+\varepsilon}}{1+\varepsilon}, \tag{7}$$

$$V_\theta^H = T + \frac{\theta^{\varepsilon}((1-t)w^H)^{1+\varepsilon}}{1+\varepsilon}. \tag{8}$$

Individuals decide to invest in human capital if and only if their ability, θ, is such that their utility from being high-skilled is larger than or equal to the expected utility of being low-skilled.[7] We assume rationing is uniform so that the probability of being unemployed is equal for every low-skilled individual and does not depend on θ.[8] Hence, individuals decide to become high-skilled if $V_\theta^H \geq uV^U + (1-u)V_\theta^L$,

not concentrated on specific individuals, but instead spread uniformly across low-skilled workers, risk aversion would disappear, since the variance in low-skilled earnings would nil.

[7]Alternatively, we could speak of self-selection or sorting into skill levels. Our model is thus equivalent to an occupational-choice model with a high-skilled (high-wage) occupation and a low-skilled (low-wage) occupation.

[8]Our assumption of uniform rationing along the extensive margin, i.e., by laying off workers, is not innocuous. Rationing could as well be dependent on the ability level, θ, or could occur along the intensive margin by restricting hours. In Gerritsen and Jacobs (2013), we analyze the welfare consequences of a binding minimum wage in a more general setup in which individual unemployment rates may or may not depend on θ. Had rationing occurred along the intensive margin, it would be more efficient than rationing along the extensive margin. In an unrationed situation the marginal gain of an extra hour of work equals the marginal cost, such that a marginal level of intensive rationing does not have any welfare cost. We discuss the implications of such efficient rationing in a later section.

where $u = N^U/(N^L + N^U)$ is the unemployment rate amongst the low-skilled, defined as the share of the low-skilled population that is unemployed. Thus, we obtain a cut-off ability, Θ, for the individual who is indifferent between becoming skilled or staying unskilled:

$$\Theta = (1-u)^{\frac{1}{\varepsilon(1-\beta)}} \left(\frac{w^H}{w^L}\right)^{-\frac{1+\varepsilon}{\varepsilon(1-\beta)}}. \tag{9}$$

The cut-off ability decreases with the skill premium, $\frac{w^H}{w^L}$, and with the unemployment rate, u. A larger skill premium increases the benefits of being high-skilled, thereby leading to a decrease of the cut-off level of ability, Θ. Similarly, a higher unemployment rate increases the relative benefits of being skilled, since high-skilled workers are not affected by unemployment. Thus, a larger unemployment rate decreases the cut-off level of ability. The minimum wage therefore has an ambiguous effect on skill formation. On the one hand, it lowers the skill premium. On the other hand, it raises unemployment among low-skilled workers. Note that the tax system does not affect skill formation, since high-skilled and low-skilled workers face the same tax rate. Individuals respond more elastically to wage differentials and unemployment rates if the elasticity of labor supply, ε, decreases, or if comparative advantage is weaker, i.e. β higher. Intuitively, a low ε and a high β make individuals more similar across skill types. As individuals are more similar, small changes in relative earnings translate into large changes in Θ.

For later reference, we note that

$$V_\Theta^H = V_\Theta^L - u\left(V_\Theta^L - V^U\right) \tag{10}$$

Hence, if unemployment is strictly positive, utility for the marginal high-skilled worker is below the utility of the marginal employed low-skilled worker: $V_\Theta^H < V_\Theta^L$. The reason is that the marginal high-skilled worker avoids low-skilled unemployment.

By denoting total high-skilled labor supply and total low-skilled labor supply by $L^{H,S}$ and $L^{L,S}$, the cut-off level, Θ, implies the following values for aggregate labor supply and group-sizes for high-skilled and low-skilled workers:

$$L^{H,S} \equiv \int_{\Theta}^{\overline{\theta}} l_\theta^H \mathrm{d}G(\theta), \tag{11}$$

$$L^{L,S} \equiv \int_{\underline{\theta}}^{\Theta} l_\theta^L \mathrm{d}G(\theta), \tag{12}$$

$$N^H \equiv 1 - G(\Theta), \tag{13}$$

$$N^L \equiv G(\Theta) - N^U. \tag{14}$$

Note that $L^{L,S}$ is the *notional* aggregate low-skilled labor supply. In the presence of unemployment, not all low-skilled workers notionally supplying labor find employment.

2.2 Firms

There is a representative, competitive, profit-maximizing firm which produces output, Y, by employing aggregate high-skilled labor, L^H, and low-skilled labor, L^L, as factors of production. The price of output is normalized to unity. The firm operates a neoclassical constant-returns-to-scale production technology, which satisfies the Inada conditions:

$$\begin{gathered} Y = F(L^H, L^L), \quad F_H, F_L > 0, \quad F_{HH}, F_{LL} < 0, \quad F_{HL} > 0, \\ \lim_{H\to\infty} F_H = \lim_{L\to\infty} F_L = 0, \quad \lim_{H\to 0} F_H = \lim_{L\to 0} F_L = \infty. \end{gathered} \tag{15}$$

The subscripts H and L of the production function denote partial derivatives with respect to L^H and L^L. The marginal products of labor are positive, but diminishing for each type of labor. Both inputs are essential and cooperant factors of production.

Firms demand labor, taking wage rates as given. The labor market is perfectly competitive and frictionless. The first-order conditions for profit maximization imply that the marginal labor productivities equal the wage rates of each type of worker:

$$F_H(L^H, L^L) = w^H, \tag{16}$$

$$F_L(L^H, L^L) = w^L. \tag{17}$$

These conditions, together with homogeneity of the production function, implicitly define the equilibrium factor ratio, L^H/L^L, as a function of the minimum wage, w^L. The Inada-conditions, joint with the cut-off ability level, Θ, in equation (9), imply that in equilibrium the numbers of high- and low-skilled individuals are both strictly positive, i.e., $\underline{\theta} < \Theta < \overline{\theta}$.

2.3 Labor market equilibrium

Labor market equilibrium conditions for high-skilled and low-skilled workers are given by:

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

$$L^L = L^{L,S} - \int_{\underline{\theta}}^{\Theta} u l_\theta^L \mathrm{d}G(\theta) = (1-u) \int_{\underline{\theta}}^{\Theta} l_\theta^L \mathrm{d}G(\theta). \tag{19}$$

High-skilled labor demand should equal high-skilled labor supply, since the high-skilled wage adjusts to clear the labor market. Low-skilled labor demand equals low-skilled labor supply, minus the potential working hours of the unemployed. The latter equality follows from the assumption of uniform rationing, i.e., independence of u from θ.

3 Comparative statics

We derive comparative statics to determine how unemployment and labor supply respond to a change in one of the policy variables. We do so by loglinearizing the model to find the (semi-)elasticities of the endogenous variables with respect to the policy variables: w^L, t, and T. These elasticities are an important ingredient of the government's optimization problem that we solve later. Equilibrium is described by equation (9) for Θ, the labor demand equations (16) and (17), and the two labor market clearing conditions (18) and (19).

We denote a relative change in variable x by $\tilde{x} \equiv \mathrm{d}\ln x = \mathrm{d}x/x$. Exceptions are variables that are already expressed in percentage terms: $\tilde{t} \equiv \mathrm{d}t/(1-t)$, $\tilde{u} \equiv \mathrm{d}u/(1-u)$, and $\tilde{N}^i \equiv \mathrm{d}N^i$, $i \in \{H, L, U\}$. As the latter variables are already expressed in percentage terms it is more convenient to write the elasticities of these variables as semi-elasticities. Loglinearization of the cut-off ability level, the first-order conditions for the firm, and the labor-market equilibrium conditions yields:

$$\tilde{\Theta} = \frac{1+\varepsilon}{(1-\beta)\varepsilon}(\tilde{w}^L - \tilde{w}^H) - \frac{1}{(1-\beta)\varepsilon}\tilde{u}, \tag{20}$$

$$\tilde{w}^H = \frac{1-\alpha}{\sigma}(\tilde{L}^L - \tilde{L}^H), \tag{21}$$

$$\tilde{w}^L = \frac{\alpha}{\sigma}(\tilde{L}^H - \tilde{L}^L), \tag{22}$$

$$\tilde{L}^H = -\frac{l^H_\Theta \Theta g(\Theta)}{L^H}\tilde{\Theta} + \varepsilon(\tilde{w}^H - \tilde{t}), \tag{23}$$

$$\tilde{L}^L = \frac{(1-u) l^L_\Theta \Theta g(\Theta)}{L^L}\tilde{\Theta} - \tilde{u} + \varepsilon(\tilde{w}^L - \tilde{t}), \tag{24}$$

where $\alpha \equiv F_H L^H / Y$ denotes the share of skilled labor earnings in total income and $\sigma \equiv -\mathrm{d}\ln(L^H/L^L)/\mathrm{d}\ln(F_H/F_L) = F_H F_L/(F_{HL} Y)$ is the elasticity of substitution between skilled and unskilled workers in production.

Combining these equations, and substituting for $\Theta g(\Theta)\tilde{\Theta} = -\tilde{N}^H$, yields a system of two equations: one for households and one for firms, relating changes in high-skilled employment and the low-skilled unemployment rate to changes in the minimum wage:

$$\frac{(1-\beta)\varepsilon}{\Theta g(\Theta)}\tilde{N}^H = -\left(\frac{1+\varepsilon}{\alpha}\right)\tilde{w}^L + \tilde{u}, \tag{25}$$

$$\tilde{u} = \left(\frac{\sigma+\varepsilon}{\alpha}\right)\tilde{w}^L - \left(\frac{l^H_\Theta}{L^H} + \frac{(1-u) l^L_\Theta}{L^L}\right)\tilde{N}^H. \tag{26}$$

The first equation shows that, given the unemployment rate, a higher minimum wage reduces the number of high-skilled, because of a fall in the skill premium. Higher unemployment results in more skill formation, since individuals would like to avoid unemployment, which is concentrated among the unskilled. The second equation shows that, for a given number of high-skilled workers, a rise in the minimum wage increases the unemployment rate. As the minimum wage rises, firms start laying off low-skilled workers. A higher number of high-skilled workers increases the return to low-skilled labor and thus decreases the unemployment rate for a given minimum wage.

The equilibrium conditions can be solved for the changes in the number of high-skilled workers and the unemployment rate to find:

$$\frac{\tilde{N}^H}{\tilde{w}^L} = \frac{\sigma - 1}{\alpha\eta}, \tag{27}$$

$$\frac{\tilde{u}}{\tilde{w}^L} = \frac{\sigma + \varepsilon - (\sigma - 1)\kappa}{\alpha} > 0, \tag{28}$$

where $\eta \equiv \frac{(1-\beta)\varepsilon}{\Theta g(\Theta)} + \frac{(1-u)l^L_\Theta}{L^L} + \frac{l^H_\Theta}{L^H} > 0$ and $\kappa = \left(\frac{(1-u)l^L_\Theta}{L^L} + \frac{l^H_\Theta}{L^H}\right)\eta^{-1} \in (0,1)$. The sign of (28) follows from $\sigma + \varepsilon - (\sigma - 1)\kappa = (1-\kappa)\sigma + \varepsilon + \kappa > 0$.

As we can see from the first equation, there is a knife-edge condition that determines whether a rise in the minimum wage increases or decreases the amount of skill formation in the economy. If $\sigma < 1$, an increase in the minimum wage leads to less high-skilled workers. Intuitively, if high-skilled workers and low-skilled workers are poor substitutes, firms are less willing to substitute low-skilled workers for high-skilled workers. Therefore, employment of low-skilled workers does not fall enough compared to the drop in the skill premium to induce individuals to invest more in human capital. For $\sigma > 1$ the converse is true, and unemployment of low-skilled workers increases so much that individuals invest more in human capital, even though the skill premium has decreased. For $\sigma = 1$, an increase in the minimum wage has no effect on the share of high-skilled workers, since the effects of a lower skill premium exactly offsets the effect of a higher unemployment rate.

This result is particularly sensitive to a number of simplifying assumptions we made. First, the assumption of uniform rationing ensures that unemployment affects skill formation. Had rationing been more efficient and had low-skilled workers with ability Θ had a higher chance of obtaining a job than other low-skilled workers, unemployment would have had a smaller effect on skill formation. Second, the assumption of quasi-linear utility, or risk-neutrality, affects the effect of unemployment on skill formation. Had individuals been more risk-averse, they would have been more averse to the possibility of unemployment, and a minimum wage would have had a more positive effect on skill formation. Finally, because in this simple setup low-skilled workers and high-skilled workers face the same tax rates, t, and transfer, T, these tax instruments do not affect the effect of a minimum wage on skill formation. This changes once we allow for skill-specific tax rates and transfers, as we show below.

What are plausible values for the substitution elasticity is an empirical question. Estimates of the substitution elasticity between high- and low-skilled workers are typically found to be larger than one. Katz and David (1999) find that a common estimate for σ is around 1.4, although much higher estimates are also reported. Hence, in the simple setup of our model, empirically plausible values for the substitution elasticity imply that the introduction of a minimum wage would typically lead to more skill formation. This finding is similar to the finding that in response to a minimum wage increase total wage income of the affected group declines if labor demand elasticities for minimum wage workers exceed unity (e.g., Dolado, Felgueroso and Jimeno, 2000; Freeman, 1996).

From the second equation follows that a minimum wage unambiguously increases the unemployment rate amongst the low-skilled. The first two terms in equation (28), σ and ε, represent labor demand and intensive labor supply responses to a higher minimum wage. An increase in the minimum wage leads to lower labor demand and higher intensive labor supply, both contributing to an increase in unemployment. The third term $-(\sigma - 1)\kappa$ represents the human capital response. If $\sigma > 1$, the increase of the minimum wage leads to more skill formation, which renders this term negative, so that the unemployment effect diminishes. Intuitively, if more low-skilled workers transfer to the skilled sector, less of them need to be laid off.[9] Assuming $\sigma > 1$, studies that do not take into account

[9] There is a large empirical literature dealing with the effect of a higher minimum wage on total employment. Although most of the evidence seems to point to negative employment effects (e.g., Neumark and Wascher, 2006), some present evidence of positive or non-negative employment effects (e.g., Card and Krueger, 1995). In our model, it can be shown to be theoretically possible that a minimum wage increases high-skilled employment by so much that the number of unemployed, N^U, decreases, even though the low-skilled unemployment rate, u, increases. However, calibration points out that such a positive employment effect would only happen under extreme

human capital responses to a minimum wage might therefore underestimate the desirability of a minimum wage.

Proposition 1 *The minimum wage reduces (increases) the fraction of skilled workers (N^H) if the elasticity of substitution between low-skilled and high-skilled workers (σ) is smaller (larger) than 1. If $\sigma = 1$, a change in the minimum wage has no effect on the number of high-skilled workers. A higher minimum wage (w^L) increases the low-skilled unemployment rate (u). A higher minimum wage boosts the unemployment rate more if σ is higher and if the elasticity of low-skilled labor supply (ε) is higher.*

4 Optimal skill-independent policy

4.1 Government's objective

The government maximizes social welfare by optimally deciding on the minimum wage, the income tax rate, and the non-individualized lump-sum transfer. We rule out individualized lump-sum taxes and transfers. Consequently, the government has to resort to distortionary policy instruments to implement its redistributive goals. All individuals receive a lump-sum transfer T and, if employed, are taxed at a rate t of labor earnings. The informational requirement to implement this linear tax system is for the government to observe aggregate labor earnings. We assume for now that the government is unable to distinguish high-skilled workers from low-skilled workers for tax purposes, which implies that we do not allow for group specific lump-sum taxes and transfers, such as education subsidies.

Social welfare, $\mathcal{W}$, is a weighted sum of utilities:

$$(29) \qquad \mathcal{W} = N^U \Psi(V^U) + (1-u)\int_{\underline{\theta}}^{\Theta} \Psi(V_\theta^L) \mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} \Psi(V_\theta^H) \mathrm{d}G(\theta),$$

where $\Psi'(\cdot)$ is a concave function of utility, with $\Psi'(\cdot) > 0$ and $\Psi'(\cdot) \leq 0$. Since utility is assumed to be quasi-linear in income, any social desire for redistribution enters through concavity of $\Psi(\cdot)$. Thus, if the government is utilitarian ($\Psi(V_\theta^i) =$

parameter values with σ exceeding 10 or the unemployment rate, u, exceeding 80 percent.

V_θ^i, $\Psi'(V_\theta^i) = 1$), the social objective exhibits no preference for redistribution. On the other extreme, if the government is Rawlsian, it only values the utility of the least well off (V^U in the presence of unemployment, $V_{\underline{\theta}}^L$ otherwise).

The government budget constraint states that government expenditures on the lump-sum transfer, T, and an exogenously given expenditure requirement, E, equal total tax revenue from labor earnings:

$$T + E = t(L^L w^L + L^H w^H). \tag{30}$$

By defining λ as the shadow price for the budget constraint, we obtain government's first-order conditions for the minimum wage, the transfer, and the tax rate. These are given by:

$$\begin{aligned}(1-u)(1-t)\int_{\underline{\theta}}^{\Theta} l_\theta^L \Psi'(V_\theta^L)\mathrm{d}G(\theta) + (1-t)\int_{\Theta}^{\overline{\theta}} l_\theta^H \Psi'(V_\theta^H)\mathrm{d}G(\theta)\frac{\mathrm{d}w^H}{\mathrm{d}w^L} \\ -\left((1-u)\int_{\underline{\theta}}^{\Theta}(\Psi(V_\theta^L) - \Psi(V^U))\mathrm{d}G(\theta) + \lambda t w^L L^L\right)\frac{1}{1-u}\frac{\mathrm{d}u}{\mathrm{d}w^L} \\ +(\Psi(V_\Theta^H) - (1-u)\Psi(V_\Theta^L) - u\Psi(V^U))\frac{\mathrm{d}N^H}{\mathrm{d}w^L} = 0.\end{aligned} \tag{31}$$

$$N^U\Psi'(V^U) + (1-u)\int_{\underline{\theta}}^{\Theta} \Psi'(V_\theta^L)\mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} \Psi'(V_\theta^H)\mathrm{d}G(\theta) - \lambda = 0. \tag{32}$$

$$\begin{aligned}-(1-u)w^L\int_{\underline{\theta}}^{\Theta} l_\theta^L \Psi'(V_\theta^L)\mathrm{d}G(\theta) - w^H\int_{\Theta}^{\overline{\theta}} l_\theta^H \Psi'(V_\theta^H)\mathrm{d}G(\theta) \\ +\lambda(w^L L^L + w^H L^H)\left(1 - \varepsilon\frac{t}{1-t}\right) = 0.\end{aligned} \tag{33}$$

4.2 Optimal minimum wages absent taxes and transfers

To highlight the main mechanisms at work, we first determine whether the introduction of a minimum wage above the market clearing wage for low-skilled labor is desirable in the absence of taxation. We thus set $T = t = E = 0$ in order to abstract from taxation, and $u = 0$ to determine the desirability of a minimum wage in an initial equilibrium without unemployment. Note that the utility of the marginal low-skilled worker in this case exactly equals that of the marginal high-skilled worker.[10] Hence, we have $\Psi(V_\Theta^H) = \Psi(V_\Theta^L)$. To interpret equation (31), we follow Feldstein (1972) by introducing the redistributive characteristics of low-skilled and high-skilled labor income:

$$(34) \qquad 0 \leq \xi^L \equiv 1 - \frac{\frac{1}{G(\Theta)} \int_{\underline{\theta}}^{\Theta} \Psi'(V_\theta^L) w^L l_\theta^L \mathrm{d}G(\theta)}{\frac{1}{G(\Theta)} \int_{\underline{\theta}}^{\Theta} \Psi'(V_\theta^L) \mathrm{d}G(\theta) \frac{1}{G(\Theta)} \int_{\underline{\theta}}^{\Theta} w^L l_\theta^L \mathrm{d}G(\theta)} \leq 1,$$

$$(35) \qquad 0 \leq \xi^H \equiv 1 - \frac{\frac{1}{1-G(\Theta)} \int_{\Theta}^{\overline{\theta}} \Psi'(V_\theta^H) w^H l_\theta^H \mathrm{d}G(\theta)}{\frac{1}{1-G(\Theta)} \int_{\Theta}^{\overline{\theta}} \Psi'(V_\theta^H) \mathrm{d}G(\theta) \frac{1}{1-G(\Theta)} \int_{\Theta}^{\overline{\theta}} w^H l_\theta^H \mathrm{d}G(\theta)} \leq 1.$$

ξ^i, $i = \{L, H\}$, is the negative normalized covariance between the social welfare weights and labor earnings. It measures the marginal social welfare gain expressed in monetary units as a fraction of labor income from redistributing one unit of income through lowering the wage rate in skill-group i. The redistributive characteristic is positive for governments that value redistribution from rich to poor, as in that case social welfare weights are decreasing with income. For governments that do not value redistribution ($\Psi'(\cdot) = 1$), ξ^i equals zero. Similarly, if there is no income inequality in either group, the redistributive characteristic is also zero. The redistributive characteristic increases with stronger social redistributive preferences and with larger pre-tax income inequality.

By rearranging the first-order condition for the minimum wage, equation (31), and substituting in the redistributive characteristics and the elasticity of unemployment, we find that it is desirable to implement a minimum wage if the following condition holds:

$$(36) \quad \overline{\Psi'(V^L)} - \overline{\Psi'(V^H)} - \xi^L \overline{\Psi'(V^L)} + \xi^H \overline{\Psi'(V^H)} > \left(\frac{\overline{\Psi(V^L)} - \Psi(V^U)}{w^L L^L} N^L \right) \frac{\tilde{u}}{\tilde{w}^L},$$

where $\overline{\Psi'(V^L)} \equiv \int_{\underline{\theta}}^{\Theta} \Psi'(V_\theta^L) dG(\theta)/G(\Theta)$ and $\overline{\Psi'(V^H)} \equiv \int_{\underline{\theta}}^{\Theta} \Psi'(V_\theta^H) dG(\theta)/(1 - G(\Theta))$ are the averages of the marginal social welfare of income of skilled and

[10] We henceforth refer to a low-skilled worker with ability Θ as 'the marginal low-skilled worker,' and to the high-skilled worker with ability Θ as 'the marginal high-skilled worker.'

low-skilled workers. $\overline{\Psi(V^L)} = \int_{\underline{\theta}}^{\Theta} \Psi(V_\theta^L) dG(\theta)/G(\Theta)$ is the average social welfare of low-skilled workers.

In the absence of unemployment and interactions with the tax system, a minimum wage has two first-order effects on social welfare. The left-hand side of inequality (36) shows the distributional benefits of a higher minimum wage, whereas the right-hand side shows the efficiency costs of a higher minimum wage. A higher minimum wage affects both inequality between the groups of low-skilled and high-skilled workers and inequality within the groups of low-skilled and high-skilled workers. The minimum wage reduces inequality between high-skilled and low-skilled workers through general equilibrium effects on the wage structure. By raising the minimum wage, low-skilled employment declines, leading to lower high-skilled productivity and wages. This reduction in inequality is given by the first two terms of above condition. For a government with redistributive preferences, this general-equilibrium effect raises social welfare as the average social marginal utility of a low-skilled worker is larger than that of a high-skilled worker: $\overline{\Psi'(V_L)} > \overline{\Psi'(V_H)}$.

Besides between-group inequality, the minimum wage also affects inequality within the groups of high-skilled and low-skilled workers. Since high-skilled individuals with high ability make relatively many working hours, they suffer more from a decline in their wage rate than high-skilled individuals with a lower ability. This reduction in inequality within the group of high-skilled workers contributes to the desirability of the minimum wage as long as $\xi^H > 0$. However, by raising the low-skilled wage rate per hour worked, a higher minimum wage also increases inequality in low-skilled labor earnings if $\xi^L > 0$. Therefore, the minimum wage is less attractive for redistributive reasons if it generates much inequality among low-skilled workers, or if the government is strongly averse to inequality within the group of low-skilled workers.

A minimum wage has, overall, favorable distributional gains, since $(1-\xi^L)\overline{\Psi'(V^L)} = \int_{\underline{\theta}}^{\Theta} \frac{\Psi'(V_\theta^L)}{\lambda} \frac{(1-u)l_\theta^L}{L^L} \mathrm{d}G(\theta) > \int_{\Theta}^{\overline{\theta}} \frac{\Psi'(V_\theta^H)}{\lambda} \frac{l_\theta^H}{L^H} \mathrm{d}G(\theta) = (1-\xi^H)\overline{\Psi'(V^H)}$. The second and third terms give weighted averages of the marginal social welfare of income for low- and high-skilled workers. The inequality follows from the fact that, in the case of $u = 0$, the marginal social welfare of each low-skilled worker is higher than the marginal social welfare of any high-skilled worker. However, if unemployment is positive ($u > 0$) this inequality does not necessarily hold. For example, if $\xi^L >> \xi^H$, and with positive unemployment, the minimum wage might cause net distributional costs rather than benefits. Intuitively, in that case the increase in inequality *within* the group of low-skilled workers is not off-set by a reduction in inequality *between* low-skilled and high-skilled workers and *within* the group of high-skilled workers.

The second first-order effect of a minimum wage on social welfare is given by the right-hand side of (36) and originates from involuntary unemployment. An increase in the low-skilled wage rate above the market-clearing level causes firms to lay off low-skilled workers. $(\overline{\Psi(V^L)} - \Psi(V^U))/w^L L^L$ measures the welfare loss due to larger unemployment in terms of total low-skilled income. For every laid off low-skilled worker society looses on average $\overline{\Psi(V^L)}$ of social welfare and gains $\Psi(V^U)$. Since the unemployed have lower utility levels than the employed, larger unemployment results in lower social welfare. Notice that any social welfare effect of the minimum wage on human capital investment is a second-order effect when there is no unemployment, since the utility of a marginal high-skilled worker equals the utility of a marginal low-skilled worker: $V^L_\Theta = V^H_\Theta$.

The desirability of a minimum wage crucially depends on the elasticity of the unemployment rate with respect to the minimum wage, $\tilde{u}/\tilde{w}^L$. In particular, as becomes clear from the discussion of equation (28), a minimum wage raises unemployment more if the elasticity of substitution between skilled and unskilled workers, σ, is larger, the labor supply elasticity of low-skilled workers, ε, is larger, and, assuming $\sigma > 1$, if the human capital response, $(\sigma - 1)\kappa$, is smaller.

That unemployment results in a first-order welfare loss is an important deviation from Lee and Saez (2012) who assume efficient rationing. In Lee and Saez, the marginal laid-off worker has zero surplus from working, and is thus indifferent between working and being unemployed. Consequently, starting from a situation without unemployment, the social welfare loss of larger unemployment is only a second-order effect. In our model this does not hold, since every low-skilled worker prefers working over being unemployed. Contrary to Lee and Saez, individuals in our model incur disutility of work on the intensive margin, so that lay-offs are always inefficient. Later on, we briefly turn to the case of efficient rationing.

As we can see from equation (36), the desirability of a minimum wage depends on the redistributive preferences of the government. If it does not value redistribution – i.e., in the case of a utilitarian social welfare function – $\overline{\Psi'(V^L)} = \overline{\Psi'(V^H)} = 1$ and $\xi^L = \xi^H = 0$, and the left-hand side of the inequality vanishes. Therefore, the government would not want to introduce a distortionary minimum wage as it produces no distributional benefits. If the government has Rawlsian preferences, the social welfare function without unemployment is given by $\mathcal{W} = V^L_{\underline{\theta}}$ and with unemployment is given by $\mathcal{W} = V^U$. In that case, a minimum wage is always welfare decreasing, since government only cares for the utility of the least well off. On both extremes of the spectrum of redistributive preferences – without any redistributive and with maximum redistributive preferences – a minimum wage is not desirable. However, for intermediate cases of redistributive preferences, this is not necessarily the case.

Proposition 2 *Starting from an undistorted initial equilibrium, the introduction of a minimum wage has ambiguous welfare effects for any redistributive, non-Rawlsian social welfare function. A minimum wage is more likely to be socially desirable if the elasticity of substitution between high-skilled and low-skilled workers (σ) is small, if the labor supply elasticity (ε) is small, if the welfare differential between the low-skilled employed and the unemployed $\overline{\Psi(V^L)} - \Psi(V^U)$ is small, and if the general equilibrium effects on wages yield large distributional gains, $(1-\xi^L)\overline{\Psi'(V^L)} - (1-\xi^H)\overline{\Psi'(V^H)}$. Distributional gains are higher with larger inequality between skilled and unskilled workers, $\overline{\Psi'(V^L)} >> \overline{\Psi'(V^H)}$, and with larger inequality within the group of high-skilled workers compared to low-skilled workers, $\xi^H >> \xi^L$. A minimum wage is never optimal if the social welfare function is Rawlsian or when it exhibits no preference for redistribution.*

To find the optimal minimum wage in the absence of taxes and transfers, we rewrite the first-order condition for the minimum wage (31) to obtain:

$$\begin{aligned}(37)\quad \overline{\Psi'(V^L)} - \overline{\Psi'(V^H)} - \xi^L\overline{\Psi'(V^L)} + \xi^H\overline{\Psi'(V^H)} = &\left(\frac{\overline{\Psi(V^L)} - \Psi(V^U)}{w^L L^L} N^L\right)\frac{\tilde{u}}{\tilde{w}^L} \\ &- \left(\frac{\Psi(V^H_\Theta) - (1-u)\Psi(V^L_\Theta) - u\Psi(V^U)}{w^L L^L}\right)\frac{\tilde{N}^H}{\tilde{w}^L}.\end{aligned}$$

Notice that this is only an optimality condition provided that the desirability condition (36) holds. Provided that a binding minimum wage is indeed welfare increasing, the optimal minimum wage is set in such a way that the marginal redistributive gains due to lower income inequality between low-skilled and high-skilled workers (left-hand side) equals the marginal welfare losses of raising involuntary unemployment (right-hand side, first term) minus the marginal welfare gain or loss associated with the change in skill formation (right-hand side, second term). The first two terms are discussed above, the last one is new.

The second term on the right-hand side represents a positive externality from skill formation. If government has redistributive preferences, and thus if $\Psi(\cdot)$ is strictly concave, we can establish that $\Psi(V_\Theta^H) - (1-u)\Psi(V_\Theta^L) - u\Psi(V^U) > 0$ if $u > 0$, since we know from equation (10) that $V_\Theta^H - (1-u)V_\Theta^L - uV^U = 0$. Intuitively, becoming high skilled can be seen as an insurance against unemployment. Due to concave social preferences, the government attaches a higher cost to the risk of becoming unemployed than the risk-neutral individuals themselves. Therefore, in the presence of involuntary unemployment, skill formation's social value exceeds its private value. A binding minimum wage, resulting in a positive unemployment rate, thus leads to an externality on skill formation. Clearly, there is no externality in the absence of unemployment as it vanishes for $u = 0$ or when government has no desire to redistribute income.[11] A higher minimum wage changes human capital formation if $\tilde{N}^H/\tilde{w}^L \neq 0$, and thus, as we have seen above, if $\sigma \neq 1$. As discussed, if $\sigma > 1$, a higher minimum wage leads to more high-skilled workers. Due to the positive externality associated with skill formation, a higher minimum wage generates an additional welfare gain. If $\sigma < 1$ a higher minimum wage leads to less skill formation, exacerbating the inefficiently low degree of skill formation.

[11] By assuming quasi-linear utility functions we abstracted from risk aversion at the individual level. Risk aversion would reduce the positive externality from human capital investment, since individuals hedge against against labor market risk by investing more in human capital (see also Jacobs, Schindler and Yang, 2012). However, the positive externality will not disappear as long as the social welfare function is a (strictly) concave transformation of the individuals' private utility functions.

4.3 Optimal minimum wages, taxes and transfers

4.3.1 Optimal transfer

The first order condition for the transfer, equation (32), can be rewritten to find that the social marginal benefits of a higher transfer, T, should equal its social marginal costs:

$$\frac{\overline{\Psi'(\cdot)}}{\lambda} \equiv N^U \frac{\Psi'(V^U)}{\lambda} + (1-u)\int_{\underline{\theta}}^{\Theta} \frac{\Psi'(V_\theta^L)}{\lambda} \mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} \frac{\Psi'(V_\theta^H)}{\lambda} \mathrm{d}G(\theta) = 1, \tag{38}$$

where $\overline{\Psi'(\cdot)}$ denotes the average marginal social welfare of income. The first three terms on the left-hand side give the increase in social welfare (expressed in monetary units) of the unemployed, low-skilled employed and high-skilled employed due to a marginally higher lump-sum transfer. This equals the transfer's resource costs on the right-hand side, equaling 1.

4.3.2 Optimal tax rate

In order to derive the optimal tax rate, we again follow Feldstein (1972) by introducing the redistributive characteristic of total labor income:

$$\xi \equiv 1 - \frac{(1-u)\int_{\underline{\theta}}^{\Theta} \Psi'(V_\theta^L) w^L l_\theta^L \mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} \Psi'(V_\theta^H) w^H l_\theta^H \mathrm{d}G(\theta)}{(w^L L^L + w^H L^H)\left(N^U \Psi'(V^U) + (1-u)\int_{\underline{\theta}}^{\Theta} \Psi'(V_\theta^L)\mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} \Psi'(V_\theta^H)\mathrm{d}G(\theta)\right)} \geq 0 \tag{39}$$

The interpretation is identical to the distributional characteristics of low-skilled and high-skilled earnings introduced before. It is the negative of the normalized covariance between the social welfare weights and labor earnings across the entire population.[12] We can establish a direct link between the distributional characteristics of skilled and unskilled labor income and the distributional characteristic for aggregate labor income:

$$(1-\xi) \equiv (1-\alpha)\frac{\overline{\Psi'(V^L)}}{\overline{\Psi'(\cdot)}}(1-\xi^L) + \alpha\frac{\overline{\Psi'(V^H)}}{\overline{\Psi'(\cdot)}}(1-\xi^H). \tag{40}$$

[12] Note that labor earnings of the unemployed are zero, and do not feature in the numerator.

Consequently, one minus the distributional characteristic of aggregate labor is a weighted sum of one minus the distributional characteristics of skilled and unskilled labor – where the income shares α and $1-\alpha$ have been used as weights and a correction has been made for the differences in the average marginal welfare of income. We can further substitute for $\overline{\Psi'(\cdot)}$ and rearrange to disentangle ξ in a between-groups and a within-groups part:

$$\xi = N^U \frac{\Psi'(V^U) - \overline{\Psi'(V^L)}}{\overline{\Psi'(\cdot)}} + (\alpha - N^H)\frac{\overline{\Psi'(V^L)} - \overline{\Psi'(V^H)}}{\overline{\Psi'(\cdot)}} + (1-\alpha)\frac{\overline{\Psi'(V^L)}}{\overline{\Psi'(\cdot)}}\xi^L + \alpha\frac{\overline{\Psi'(V^H)}}{\overline{\Psi'(\cdot)}}\xi^H. \tag{41}$$

The first two terms represent between-group inequality. It is high for large social welfare differences between the unemployed and the employed low-skilled, $\Psi'(V^U) - \overline{\Psi'(V^L)}$, and between low-skilled and high-skilled workers, $\overline{\Psi'(V^L)} - \overline{\Psi'(V^H)}$. Furthermore, it increases with the number of unemployed N^U and the difference between the high-skilled share of income and the high-skilled share of population $(\alpha - N^H)$. The last two terms are a weighted sum of the within-group welfare inequality for the low-skilled and the high-skilled, ξ^L and ξ^H. For a government without redistributive preferences $\xi = 0$, for Rawlsian governments $\xi = 1$.

Labor income taxation distorts labor supply. The tax rate drives a wedge between the private and social benefits of work, leading to a substitution effect from consumption to leisure. The marginal social welfare costs associated with this distortion are increasing in the elasticity of labor supply and in the tax rate and are given by $\frac{t}{1-t}\varepsilon$. The optimal tax rate is set such that the marginal redistributive gains of the tax rate equal its marginal efficiency costs. Rearranging the first order condition for t, equation (33), thus yields:

$$\frac{t}{1-t} = \frac{\xi}{\varepsilon}. \tag{42}$$

From this equation we derive the familiar result of optimal tax theory, that the optimal income tax rate is increasing in the distributional gain and decreasing in the labor supply elasticity, see for example Atkinson and Stiglitz (1980).

4.3.3 Optimal minimum wage

Again, we first analyze the desirability of a minimum wage, in combination with optimal taxation, by substituting for $u = 0$ in first-order condition (31), and substituting out the partial derivatives and the optimal income tax. A minimum wage is desirable if the following condition holds:

$$\frac{\overline{\Psi'(V^L)} - \overline{\Psi'(V^H)}}{\lambda} - \xi^L \frac{\overline{\Psi'(V^L)}}{\lambda} + \xi^H \frac{\overline{\Psi'(V^H)}}{\lambda} > \left(\frac{\overline{\Psi(V^L)} - \Psi(V^U)}{(1-t)w^L L^L \lambda} N^L + \frac{\xi}{\varepsilon} \right) \frac{\tilde{u}}{\tilde{w}^L}. \quad (43)$$

As before, the first line gives the marginal redistributive gains of a minimum wage, the second line gives the welfare loss associated with higher unemployment, multiplied by the semi-elasticity of unemployment with respect to a higher minimum wage. The first term in the second line again gives the welfare loss associated with the direct utility drop of the workers who lose their job, this time normalized by low-skilled income net of taxes. With positive taxes, low-skilled workers do not reap the full benefits of a higher minimum wage as part of it is taxed away by the government. Hence, a given distributional gain requires a larger increase in the minimum wage and therefore increased unemployment is associated with a higher welfare loss.

The second term in the second line, ξ/ε, is new and captures the welfare costs of a higher minimum wage associated with an erosion of the labor tax base in the tax optimum. Unemployed workers do not pay income taxes, whereas employed workers do. $(\xi/\varepsilon)(\tilde{u}/\tilde{w}_L)$ represents these losses in tax revenue from low-skilled workers as a result of a higher minimum wage. Tax revenue declines more if the increase in unemployment due to the minimum wage is larger. This is captured by the term $\tilde{u}/\tilde{w}_L = (\sigma + \varepsilon - (\sigma - 1)\kappa)/\alpha$, the semi-elasticity of unemployment with respect to the minimum wage. This term has been extensively discussed above. $\xi/\varepsilon = t/(1-t)$ is the tax wedge on low-skilled labor supply. A minimum wage is more distortionary if government sets the labor tax rate at higher levels. The larger are tax distortions on labor supply – i.e., the larger is t – the costlier it is to raise the minimum wage. Thus the minimum wage exacerbates the distortions of the labor tax on low-skilled labor supply by further eroding the tax base.

4.3.4 Minimum wage versus income taxation

As equation (43) shows, the main benefit of a minimum wage is its capacity to redistribute income from high-skilled workers to low-skilled workers. Furthermore, it decreases inequality among the high-skilled but increases inequality among the low-skilled and creates unemployment, causing a drop of utility and tax revenues from those who lose their jobs. A minimum wage is not the only means of redistributing income from high- to low-skilled workers. The same redistribution can be achieved through a higher income-tax rate, while rebating revenue in the form of higher transfers. A minimum wage will be optimal if, and only if, the marginal costs of redistribution through a minimum wage are smaller than the marginal costs of the same redistribution through higher income taxes, evaluated at the tax optimum.

To see whether this is indeed the case, we rewrite equation (42) by substituting for ξ using equation (40):

$$\text{(44)} \quad \frac{t}{1-t}\varepsilon = (\alpha - N^H)\frac{\overline{\Psi'(V^L)} - \overline{\Psi'(V^H)}}{\lambda} + (1-\alpha)\frac{\overline{\Psi'(V^L)}}{\lambda}\xi^L + \alpha\frac{\overline{\Psi'(V^H)}}{\lambda}\xi^H.$$

The left-hand side gives the marginal dead-weight loss of taxation. The right-hand side gives the distributional benefits of taxation by reducing inequality between skill groups (first-term), reducing inequality within the low-skilled group (second term), and reducing inequality within the high-skilled group (third term). Substituting $\left(\overline{\Psi'(V^L)} - \overline{\Psi'(V^H)}\right)/\lambda$ from equation (44) into the desirability condition of a minimum wage, equation (43), illustrates the welfare effects of a binding minimum wage, relative to income taxes that yield the same degree of between-group redistribution. The desirability of the minimum wage can then be written as follows:

$$\text{(45)} \quad \frac{t}{1-t}\left(\frac{\varepsilon}{\alpha - N^H} - \frac{\tilde{u}}{\tilde{w}^L}\right) > \frac{\overline{\Psi(V^L)} - \Psi\left(V^U\right)}{(1-t)L^L w^L \lambda}N^L\frac{\tilde{u}}{\tilde{w}^L} + \frac{1-N^H}{\alpha - N^H}\frac{\overline{\Psi'(V^L)}}{\lambda}\xi^L + \frac{N^H}{\alpha - N^H}\frac{\overline{\Psi'(V^H)}}{\lambda}\xi^H.$$

This expression gives an alternative interpretation regarding the desirability of a binding minimum wage. The left-hand side gives the efficiency costs of attaining a given between-group redistribution through an increase in the tax rate, over and above the efficiency costs of attaining the same between-group redistribution through higher minimum wages. The efficiency costs of a higher tax rate are given by the tax base erosion that takes place due to the distortion of intensive labor supply, $\frac{t}{1-t}\frac{\varepsilon}{\alpha-N^H}$. The intensive labor supply elasticity is multiplied by a term $1/(\alpha - N^H)$. The smaller the high-skilled share of total income, relative to its population share, the more difficult it is to redistribute from high-skilled workers to low-skilled workers by using the tax rate, t, which applies, after all, to both high-skilled and low-skilled workers.

The efficiency costs of a higher minimum wage are given by the tax base erosion that takes place due to the distortion on the extensive employment margin, $\frac{t}{1-t}\frac{\tilde{u}}{\tilde{w}^L}$. Overall, a minimum wage might be more efficient in redistributing income from high-skilled workers to low-skilled workers because it directly raises low-skilled income and decreases high-skilled income. Uniform income taxes, on the contrary, cause net wages of both low-skilled and high-skilled workers to decline in order to redistribute the revenue back in the form of lump-sum transfers to both low-skilled and high-skilled workers.

However, for minimum wages to be optimal, this efficiency gain must outweigh its distributional losses relative to a tax increase. These losses are given by the right-hand side and consist of the direct welfare losses of laid-off individuals (first term) and the within-group distributional gains of a tax increase (last two terms). The first term we encountered and discussed before. The last two terms appear because the uniform income tax rate needs to be raised strongly, relative to the minimum wage, in order to achieve a given between-group redistribution. While this leads to higher efficiency losses, as illustrated by the left-hand side, this also leads to more within-group income equality, a distributional gain.

In short, the minimum wage complements the income tax to reduce between-group income inequality. A minimum wage helps to directly redistribute income from high-skilled workers towards low-skilled workers without the tax-base erosion on the intensive margin associated with taxation, and thereby alleviates the distributional imperfection associated with the uniformity of the income tax.

Proposition 3 *Optimal labor-income taxes increase with the level of earnings inequality and decrease with the elasticity of intensive labor supply. Minimum wages are more distortive if the government sets high taxes on labor earnings, since minimum wages erode the tax base by causing unemployment. Hence, the presence of taxes makes a minimum wage less desirable. The role of minimum wages in an optimal skill-independent tax-benefit system is to complement the tax-benefit system by reducing the distributional imperfections of the income tax. Minimum wages help to redistribute income between skill groups, so that income taxes can be better targeted at reducing inequality within skill groups.*

5 Optimal skill-dependent policies

So far, we assumed that the government cannot differentiate tax instruments according to skill type, whereas it did employ a minimum-wage policy, the enforcement of which requires knowledge on individuals' skill type. One may recognize this as an informational inconsistency. To implement and enforce a minimum wage, government necessarily has information on the individuals' wage rates, but it does not use this information in determining optimal tax policy. This section, therefore, explores the implications of allowing government to optimize a skill-dependent optimal tax and minimum-wage policy.[13] Before deriving expressions for optimal policy, we first repeat the comparative statics analysis as the key elasticities of the model change due to the introduction of skill-specific instruments.

5.1 Comparative statics again

We introduce income taxes, t^L and t^H, that depend on skill type. Moreover, a separate transfer, S, is given to high-skilled workers. Indirect utility is thus given

[13]We do not study participation-dependent policies. That is, the government is still assumed to be unable to condition taxes and transfers based on employment status. Allowing for separate unemployment benefits would necessitate the introduction of a participation margin and hence a second cut-off level for θ. To keep the model tractable we decided not to do so. In Gerritsen and Jacobs (2013), we do model the participation margin along with the skill decision.

by:

$$V_U = T, \tag{46}$$

$$V_\theta^L = T + \frac{\theta^{\beta\varepsilon}((1-t^L)w^L)^{1+\varepsilon}}{1+\varepsilon}, \tag{47}$$

$$V_\theta^H = T + \frac{\theta^{\varepsilon}((1-t^H)w^H)^{1+\varepsilon}}{1+\varepsilon} + S. \tag{48}$$

Note that the only changes as compared to equations (6)-(8) are the substitutions of t^L and t^H for t, and the inclusion of an extra transfer, S, to the high-skilled. As before, the critical value, Θ is determined by $V_\Theta^H = uV^U + (1-u)V_\Theta^L$, and hence:

$$((1-t^H)w^H)^{1+\varepsilon}\Theta^{\varepsilon} - (1-u)((1-t^L)w^L)^{1+\varepsilon}\Theta^{\beta\varepsilon} = -(1+\varepsilon)S. \tag{49}$$

The other equilibrium conditions of the model remain unaltered.

Thus, the comparative statics for Θ are different in the presence of skill-dependent tax instruments. We define $\rho \equiv \frac{1+\varepsilon}{(1-t^H)w^H l_\Theta^H + (1+\varepsilon)S} > 0$, such that ρS gives the transfer to high-skilled workers as a share of the total returns on skill formation. The loglinearized equation for Θ is now given by:

$$(1-\beta-\rho S)\varepsilon\tilde{\Theta} = (1+\varepsilon)(\tilde{w}^L - \tilde{t}^L) - (1-\rho S)(1+\varepsilon)(\tilde{w}^H - \tilde{t}^H) - \rho\tilde{S} - \tilde{u}, \tag{50}$$

where $\tilde{S} \equiv \mathrm{d}S$. Thus, Θ increases in earnings for low-skilled workers and decreases in earnings for high-skilled workers and the unemployment rate.[14]

By redefining $\eta \equiv \frac{(1-\beta-\rho S)\varepsilon}{\Theta g(\Theta)} + \frac{(1-u)l_\Theta^L}{L^L} + \frac{l_\Theta^H}{L^H}$, we can solve the loglinearized model to find the elasticities of the number of skilled workers with respect to the policy variables:

$$\eta\tilde{N}^H = \left(\frac{\sigma - 1 + (1+\varepsilon)(1-\alpha)\rho S}{\alpha}\right)\tilde{w}^L + \tilde{t}^L - (1-(1+\varepsilon)\rho S)\tilde{t}^H + \rho\tilde{S}. \tag{51}$$

Observe that, for $t^L = t^H$ and $S = 0$, the equation collapses to equation (27).

[14]For high-skilled workers to be located at $\theta > \Theta$ and low-skilled workers at $\theta < \Theta$ we require that the difference between high-skilled utility and expected low-skilled utility is increasing in θ. Taking the derivative of the equilibrium condition with respect to θ, we thus obtain the second-order necessary condition: $(1-\beta-\rho S) > 0$.

A transfer to the high skilled alters the result that skill formation increases (decreases) in response to a higher minimum wage if $\sigma > 1$ ($\sigma < 1$). In particular, high-skilled workers now have additional non-wage earnings, which are unaffected by a higher minimum wage. Thus, while the minimum wage depresses high-skilled labor earnings, the effect of minimum wages on skill formation is cushioned due to the presence of non-wage income if $S > 0$. The exact opposite holds for $S < 0$, in which case wage earnings make up for a larger share of income for high-skilled workers.[15] Now that taxes are conditioned on skill type, they do affect skill formation. Quite naturally, skill formation increases with low-skilled taxes t^L and high-skilled transfers S, and decreases with high-skilled taxes t^H, provided that $(1+\varepsilon)\rho S < 1$, which is what we assume.[16] Note that, if high-skilled workers receive higher transfers ($\rho S > 0$ larger), the impact of the high-skilled tax rate on skill formation is lowered, as the transfers remain untaxed.

As before, we can solve the linearized model to find the elasticities of unemployment with respect to the policy variables:

$$(52) \quad \tilde{u} = \left(\frac{\sigma + \varepsilon - \kappa(\sigma - 1) - \kappa(1+\varepsilon)(1-\alpha)\rho S}{\alpha} \right) \tilde{w}^L - (\kappa + \varepsilon)\tilde{t}^L + (\varepsilon + \kappa - \kappa(1+\varepsilon)\rho S)\tilde{t}^H - \kappa\rho\tilde{S}.$$

Again, for $t^L = t^H$ and $S = 0$, this equation collapses to equation (28). The minimum wage has a smaller effect on unemployment if $S > 0$. The reason is that the impact of the minimum wage on skill formation is larger if $S > 0$. Intuitively, compared to the case in which $S = 0$, a minimum wage results in lower low-skilled labor supply, and (through input complementarity) higher low-skilled labor demand. Of course, the opposite holds if $S < 0$. Furthermore, low-skilled taxes decrease unemployment as it discourages low-skilled labor supply; high-skilled taxes

[15]Naturally, to the extent that the subsidy S itself is negatively affected by the minimum wage – for example, through the high-skilled wages of teachers – this effect is smaller and might even disappear.

[16]If $(1+\varepsilon)\rho S > 1$ a higher tax on high-skilled labor earnings leads to lower intensive high-skilled labor supply, lower low-skilled productivity, and higher unemployment, and thereby to more skill formation. This effect would then outweigh the direct negative effect on skill formation. However, notice that this presupposes a share of education subsidies in total high-skilled earnings, ρS, exceeding $1/(1+\varepsilon)$, which for plausible levels of ε would by highly unrealistic. Moreover, we later derive that at the tax optimum $S < 0$.

increase unemployment as it discourages high-skilled labor supply (and the less so if S is larger); and transfers to high-skilled workers decreases unemployment as it encourages high-skilled labor supply.

5.2 Optimal minimum wages, taxes and transfers

5.2.1 First-order conditions of optimal policy

This subsection presents the first-order conditions for the optimal minimum wage, low- and high-skilled income taxes, and lump-sum transfers for low-skilled and high-skilled workers. To interpret the optimal tax expressions, we introduce some simplifying notation. As usual, first-order conditions equate marginal distributional gains with marginal distortionary costs. Distortionary costs are represented by wedges multiplied by elasticities. Wedges in our model are defined as follows:

$$\text{(53)}\quad \Delta^H \equiv \frac{t^H}{1-t^H},$$

$$\text{(54)}\quad \Delta^L \equiv \frac{t^L}{1-t^L},$$

$$\text{(55)}\quad \Delta^U \equiv \frac{\overline{\Psi(V^L)} - \Psi(V^U)}{\lambda(1-t^L)L^L w^L} N^L,$$

$$\text{(56)}\quad \Delta^S \equiv \frac{\Psi(V^H_\Theta) - (1-u)\Psi(V^L_\Theta) - u\Psi(V^U)}{\lambda(1-t^L)L^L w^L} + \frac{t^H w^H l^H_\Theta - S - (1-u)t^L w^L l^L_\Theta}{(1-t^L)L^L w^L}.$$

These wedges measure the welfare gains of marginally higher intensive high-skilled labor supply, higher intensive low-skilled labor supply, lower unemployment, and higher skill formation. The interpretation is straightforward. The first two wedges measure revenue gains from higher labor supply as labor supply is distorted by income taxation. They are both expressed in terms of after-tax income. The third wedge gives the social welfare loss of higher unemployment due to a drop of utility – expressed in monetary terms as a fraction of net low-skilled earnings. The fourth wedge measures the welfare gains of higher skill formation. The first term gives the welfare gain of skill formation associated with the fact that government is more averse to unemployment risk than individuals. The second term gives the revenue gains (or losses) associated with larger skill formation. Whether there are revenue gains or losses depends on whether human capital formation is taxed or subsidized on a net basis, i.e., whether $t^H w^H l^H_\Theta - S - (1-u)t^L w^L l^L_\Theta \gtrless 0$. Both are expressed in terms of net low-skilled labor earnings.

Armed with the additional notation, we can express the first order-conditions as:

$$(57) \quad w^L : (1-\xi^L)\frac{\overline{\Psi'(V^L)}}{\lambda} - \left(\frac{1-t^H}{1-t^L}\right)(1-\xi^H)\frac{\overline{\Psi'(V^H)}}{\lambda} = \left(\frac{t^H-t^L}{1-t^L}\right)(1+\varepsilon) + (\Delta^U+\Delta^L)\frac{\tilde{u}}{\tilde{w}^L} - \Delta^S\frac{\tilde{N}^H}{\tilde{w}^L},$$

$$(58) \quad t^H : 1-(1-\xi^H)\frac{\overline{\Psi'(V^H)}}{\lambda} = \Delta^H\varepsilon + (\varphi\Delta^U+\varphi\Delta^L)\frac{\tilde{u}}{\tilde{t}^H} - \varphi\Delta^S\frac{\tilde{N}^H}{\tilde{t}^H},$$

$$(59) \quad t^L : 1-(1-\xi^L)\frac{\overline{\Psi'(V^L)}}{\lambda} = \Delta^L\varepsilon + (\Delta^U+\Delta^L)\frac{\tilde{u}}{\tilde{t}^L} - \Delta^S\frac{\tilde{N}^H}{\tilde{t}^L},$$

$$(60) \quad S : \frac{\overline{\Psi'(V^H)}}{\lambda} - 1 = (\gamma\Delta^U+\gamma\Delta^L)\frac{\tilde{u}}{\mathrm{d}S/S} - \gamma\Delta^S\frac{\tilde{N}^H}{\mathrm{d}S/S},$$

$$(61) \quad T : 1 - \frac{N^U\Psi'(V^U) + N^L\overline{\Psi'(V^L)} + N^H\overline{\Psi'(V^H)}}{\lambda} = 0,$$

where we denoted $\varphi \equiv \frac{(1-t^L)L^L w^L}{(1-t^H)L^H w^H}$ as total low-skilled labor income relative to total high-skilled labor income, and $\gamma \equiv \frac{(1-t^L)w^L L^L}{N^H S}$ as total low-skilled labor income relative to high-skilled transfers.

Each expression implies that the net redistributive gains (left-hand side) optimally equal distortionary costs (right-hand side). The net redistributive gains always consist of the direct distributional impact (measured in monetary equivalents) of increasing the particular instrument under consideration, plus the impact of redistributing in lump-sum fashion any additional revenue. The efficiency costs are always determined by the behavioral responses on the intensive labor supply margins, the unemployment margin and the skill-formation margin. The first-order condition for T in equation (61) remains unaltered, and will not be discussed any further.

5.2.2 Optimal minimum wage

In equation (57), the first term on the right-hand side is new and the last term is modified. The first term on the right-hand side captures the marginal revenue gains (or losses) of the minimum wage, due to its effects on gross wage rates. It affects tax revenue from low-skilled and high-skilled workers differently if they face different tax rates. We say there is 'tax rate progression' if taxes on skilled labor are higher than on unskilled labor, i.e., if $t^H > t^L$. The minimum wage increases low-skilled wages and lowers high-skilled wages, hence low-skilled labor supply increases and high-skilled labor supply falls. If there is tax rate progression, the minimum wage therefore causes a revenue loss, given that both low-skilled and high-skilled workers have the same labor-supply elasticity. These two effects exactly cancel out in the case of a flat tax rate, i.e., if $t^L = t^H$.

A second difference might originate from the last term, $-\Delta^S \tilde{N}^H / \tilde{w}^L$. This term could now turn negative if skill formation is so highly subsidized that revenue losses outweigh the social insurance gains of larger skill formation. Generally, however, the government would want to redistribute from high-skilled to low-skilled workers, implying a positive net tax on skill formation, such that $\Delta^S > 0$. This also seems to be the empirically relevant case as most industrial countries tax skill formation on a net basis (OECD, 2011). In that case, and assuming that a minimum wage boosts skill formation, ($\tilde{N}^H / \tilde{w}^L > 0$), a higher minimum wage yields higher tax revenues as it causes more skill formation. If $\Delta^S < 0$, skill formation is subsidized on a net basis and, provided that minimum wages boost skill formation, higher minimum wages result in additional revenue losses.

The remainder of the optimal minimum-wage expression is unaffected, compared to the case with skill-independent taxes and transfers.

Proposition 4 *Minimum wages are less likely to be socially desirable under skill-specific instruments if there is tax rate progression ($t^H > t^L$) or, provided that a minimum wage increases skill formation, if skill formation is subsidized on a net basis ($\Delta^S < 0$).*

5.2.3 Optimal tax rate progression

The first-order conditions for t^H and t^L in equations (58) and (59) are similar: the left-hand side gives the net social welfare gains of redistributing a unit of resources by raising the income tax rate. The marginal redistributive gains of high-skilled taxes are larger than the redistributive gains of low-skilled taxes if (i) the average marginal social value of income of high-skilled workers is lower than that of low-skilled workers – i.e., if $\overline{\Psi'(V^H)} < \overline{\Psi'(V^L)}$) – and ($ii$) the government is more concerned about within-group income-inequality in the group of high-skilled workers than in the group of low-skilled workers – i.e., if $\xi^H > \xi^L$. While these conditions depend on the specific social welfare function, they seem intuitively plausible, so that taxes on high-skilled workers should be set higher on the basis of redistributive reasons – not considering the efficiency costs.

The right-hand sides in equations (58) and (59) give the efficiency costs of using either tax instrument in terms of lower intensive labor supply (first term), higher unemployment (second term), and higher skill formation (third term). The formal structure of the first-order conditions is very similar, the implications for the optimal values of the tax rates are not. Although either tax instrument reduces intensive labor supply, as indicated by the first terms $\Delta^H\varepsilon$ and $\Delta^L\varepsilon$, the other elasticities have opposite signs in both equations. In particular, the number of high-skilled workers increases with a higher low-skilled tax rate, but it decreases with a higher high-skilled tax rate. Similarly, the unemployment rate decreases with higher low-skilled taxation, whereas it increases with higher high-skilled taxation. This means that low-skilled taxes alleviate the distortions associated with the minimum wage by reducing unemployment as they stimulate high-skilled labor supply and destimulates low-skilled labor supply. High-skilled taxes, on the other hand, exacerbate distortions of the minimum wage by raising unemployment. Whether there should be tax rate progression is therefore theoretically ambiguous.

Proposition 5 *Distributional concerns tend to call for tax rate progression. Tax rate progression is less desirable if minimum wages are set higher, as tax rate progression exacerbates the labor-market distortions of the minimum wage by increasing low-skilled labor supply. The case for tax rate progression is further weakened (strengthened) if skill formation is taxed (subsidized) on a net basis, i.e., if* $\Delta^S > 0$

($\Delta^S < 0$).

5.2.4 Optimal subsidy on skill formation

The first-order condition for S in equation (60) equates the marginal redistributive costs of directly distributing resources towards high-skilled workers, $\overline{\Psi'(V^H)}/\lambda - 1$, with the marginal welfare gains of lower unemployment and larger skill formation. The distributional gains of providing higher transfers to the high-skilled is negative as it redistributes resources in the wrong direction. Indeed, using the first-order condition for T we can derive that $\overline{\Psi'(V^H)}/\lambda < 1$. Thus, for redistributive reasons, the government would like to tax the high-skilled. However, subsidies on skill formation reduce unemployment, since $\tilde{u}/\tilde{S} = -\kappa\rho < 0$. Hence, subsidies on skill formation alleviate the distortions associated with the minimum wage. Moreover, subsidies on skill formation naturally boost skill formation as $\tilde{N}^H/\tilde{S} = \rho/\eta > 0$. If skill formation is distorted downwards (upwards), such that $\Delta^S > 0$ ($\Delta^S < 0$), subsidizing the high-skilled reduces (exacerbates) the distortion on skill formation. Consequently, it remains unclear whether skill formation should be subsidized on a net basis.

Proposition 6 *Subsidies (taxes) on skill formation result in distributional losses (gains), alleviate the distortions created by the minimum wage, and alleviates (exacerbates) distortions of skill formation if skill formation is taxed (subsidized) on a net basis.*

5.2.5 Minimum wage versus income taxation: a reinterpretation

As a final exercise, we ask the question which instruments are more desirable for income redistribution: minimum wages or income taxes? As can be inferred from the first-order conditions (58), (59), and (57), a properly designed combination of high-skilled and low-skilled income taxes can exactly replicate the distributional effects of a minimum wage. Hence, the question whether minimum wages are desirable in addition to optimal income taxes boils down to the question: do minimum wages entail larger or smaller distortions than income taxes to achieve the same marginal distributional benefits?

To answer this question, we can combine the optimal-tax expressions for the tax rates and the minimum wage to derive a new desirability condition for the minimum wage. This is essentially equivalent to determining the welfare effects of an increase in the minimum wage while offsetting its distributional effects by an appropriate adjustment of the tax rates (i.e., a low-skilled tax increase, combined with a high-skilled tax decrease). This yields the following desirability condition for the minimum wage:

$$\Delta^S \left(\frac{\tilde{N}^H}{\tilde{u}} \right) > \Delta^U + \Delta^L, \tag{62}$$

where $\tilde{N}^H/\tilde{u}$ denotes the partial effect of unemployment on high-skill labor supply. By substituting for the wedges and the elasticity, we find:

$$\left(\frac{t^H z^H_\Theta - S - (1-u)t^L z^L_\Theta}{(1-t^L)L^L w^L} \right) \frac{\Theta g(\Theta)}{(1-\beta-\rho S)\varepsilon} > \left(\frac{\overline{\Psi(V^L)} - \Psi(V^U)}{\lambda(1-t^L)L^L w^L} N^L + \frac{t^L}{1-t^L} \right). \tag{63}$$

The left-hand side gives the welfare gain from higher skill formation. The right-hand side gives the welfare losses due to higher unemployment. These constitute the net welfare effects of a higher minimum wage when the distributional effects are offset by an appropriate adjustment in income taxes. The ratio $\frac{\Theta g(\Theta)}{(1-\beta-\rho S)\varepsilon}$ represents the elasticity of the number of high-skilled workers with respect to a change in the unemployment rate.

The right-hand side gives the dead weight loss of a minimum-wage increase, over and above the costs of a tax change that features the same distributional benefits. The first term represents the utility loss of those low-skilled individuals that lose their jobs because of a higher minimum wage. The second term expresses the marginal welfare loss associated with lower tax revenue, caused by higher unemployment. These social welfare losses can be avoided by using the income tax rather than the minimum wage to redistribute income. The left-hand side gives the marginal welfare gains from the increase in skill formation caused by the higher unemployment rate. [17] There can only be welfare gains of a minimum

[17]Note that, relative to a distributionally equivalent tax reform, a minimum wage increase has no direct effect on the incentives to invest in human capital, i.e., the minimum wage's effect on

wages if skill formation is taxed on a net basis, such that $\Delta^S > 0$. Indeed, if $\Delta^S > 0$, the minimum wage alleviates the net distortion on skill formation, caused by redistributive taxation, by raising human capital investment through higher unemployment. However, if skill formation is subsidized on a net basis, such that $\Delta^S < 0$, a minimum wage can *never* be socially desirable. Indeed, relative to a distributionally equivalent reform of the income tax rates, a minimum wage increase then entails a higher dead-weight loss by causing both higher unemployment and higher skill formation.

The expression for the optimal minimum wage is obtained simply by substituting an equality sign for the inequality sign:

$$\Delta^S \left(\frac{\tilde{N}^H}{\tilde{u}} \right) = \Delta^U + \Delta^L. \tag{64}$$

The only difference with respect to the desirability condition is that the wedge on skill formation Δ^S now also contains the insurance benefit associated with skill formation, $\frac{\Psi(V_\Theta^H)-(1-u)\Psi(V_\Theta^L)-u\Psi(V^U)}{\lambda(1-t^L)w^L L^L}$, which was nil for $u = 0$. This term has been extensively discussed before. Notice that this optimality condition only holds if a binding minimum wage is desirable to start with such that inequality (63) holds for $u = 0$.

Summing up, allowing for skill-specific taxes has some important ramifications regarding the desirability of a minimum wage. To see this, we compare equation (45) with equation (63). By allowing for skill-dependent tax rates and transfers, government can directly redistribute income both within and between high-skilled and low-skilled workers. Unlike in the case of skill-independent tax rates, there is no benefit of having minimum wages to correct for a distributional imperfection of the income-tax system in reducing between-group inequality. The tax-benefit system can achieve exactly the same redistributive impact of a minimum wage, but without the subsequent increase in unemployment. This explains why the redistributive terms, that are still present in equation (45), are absent from equation (63).

However, there is now a new term in equation (63), which is associated with

the skill premium is equivalent to that of the tax reform.

the distortion on skill formation. With skill-independent tax instruments, human capital formation is not distorted by uniform taxes or transfers. When taxes and transfers are skill-dependent, the tax-benefit system is no longer neutral with respect to skill formation. Hence, while tax instruments can be more accurately targeted to distribute from high-skilled workers to low-skilled workers, doing so generates a net tax on skill formation. This distortion was absent in equation (45), but shows up as Δ^S in equation (63). Increasing the minimum wage, and simultaneously offsetting the distributional impact through the income tax system, boosts skill formation by raising unemployment. Thus, minimum wages help to alleviate the distortions of the tax-benefit system on skill formation. Recall that this only holds if investment in human capital is indeed taxed on a net basis.

Proposition 7 *If skill formation is taxed on a net basis, a marginal increase in the minimum wage, compared to a distributionally equivalent tax reform, entails a social welfare loss from higher unemployment and a social welfare gain from higher skill formation. The minimum wage increase is desirable if the social welfare gain outweighs the loss. Minimum wages are more desirable if unemployment has a larger effect on skill formation, if skill formation is more heavily taxed, and if the utility and tax revenue losses associated with higher unemployment are lower. Minimum wages are not desirable if human capital formation is subsidized on a net basis.*

6 Efficient rationing

6.1 Efficient versus inefficient rationing

A binding minimum wage leads to an oversupply of labor and hence to rationing on the labor market. Up to now we only discussed uniform rationing on the extensive margin, according to which every low-skilled individual has the same chance of getting fired. We now also discuss efficient rationing. If individuals are heterogeneous with respect to the disutility of extensive labor supply, rationing is efficient if persons with the highest disutility of work lose their job first. If workers are heterogeneous with respect to disutility of intensive labor supply, as is the case in our model, rationing is efficient if it occurs on the intensive margin, i.e. by restricting the number of hours people work. Every worker equalizes the marginal disutility of work with the marginal utility of higher income. Therefore, a marginal decrease in working hours, forced upon workers by a binding minimum wage, only has second-order effects on individuals' utilities.

Theory provides little guidance when it comes to the efficiency of rationing. In absence of a secondary or "black" market, in which the rationed good is traded, there is little reason to assume the rationed goods are acquired by the individuals who desire them most (Tobin, 1952). Empirically, as noted by Luttmer (2007), this has been confirmed by studies of the U.S. residential market for gas (Davis and Kilian, 2011), the gasoline market (Deacon and Sonstelie, 1989; Frech and Lee, 1987) and on the housing rental market (Glaeser and Luttmer, 2003). As there is no secondary market for jobs or hours of work, it is unlikely that labor rationing due to a minimum wage is efficient. The only more or less direct evidence for the efficiency of lay-offs due to a minimum wage is due to 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 have been interpreted as evidence that workers with the lowest utility surplus of work are rationed first, were it not that in the sensitivity analysis, he finds significant increases in two proxies. Hence, he does not find convincing evidence that the efficiency of job allocation changed due to a change in the minimum wage. He does, however, find some suggestive evidence of a drop in employment due to a higher minimum wage. This supports our assumption of uniform rationing on the extensive margin as low-skilled workers are laid off while the composition of workers, in terms of reservation wages, does not change. The assumption of rationing through lay-offs is further supported by a large body of evidence (Neumark and Wascher, 2006).

There is much less evidence on whether there is rationing on the intensive margin, let alone on its efficiency, and the evidence that exists seems to be conflicting. For example, Zavodny (2000) finds that a minimum wage reduces employment, but *increases* average hours worked, while Couch and Wittenburg (2001) find that a minimum wage reduces both employment and hours worked.

6.2 Model and comparative statics

If rationing occurs exclusively on the intensive margin, there is no unemployment. The intensive labor supply decision of high-skilled workers remains unaltered. However, in the case of a binding minimum wage, low-skilled workers face a restriction on the number of hours they are allowed to work. We denote the effective labor supply – the actual number of hours worked – as l^e, and the maximum number of hours an individual with ability θ is allowed to work as $\bar{l}_\theta$. The size of the minimum wage determines the aggregate number of hours that firms can feasibly employ. How this aggregate hour restriction translates into individual hour rations, and how these rations depend on θ is *a priori* unclear. We assume that the rations are efficient and determine the implications of this for the specific functional form of $\bar{l}_\theta$ below.

Low-skilled workers maximize utility, $V_\theta^L = (1-t)w^L l^e - \frac{1}{\theta^\beta}\frac{(l^e)^{1+1/\varepsilon}}{1+1/\varepsilon}$, with respect to effective labor supply, l^e, subject to the rationing constraint, $\bar{l}_\theta \geq l^e$.[18] We denote the Kuhn-Tucker multiplier for this constraint as $(1-t)w^L\mu$. In equilibrium, μ is the normalized Kuhn-Tucker multiplier which gives the shadow price of relaxing the rationing constraint in terms of the net wage, $(1-t)w^L$. The Lagrangian for the maximization problem of the individual can thus be written as:

$$\mathcal{L} = (1-t)w^L l^e - \frac{1}{\theta^\beta}\frac{(l^e)^{1+1/\varepsilon}}{1+1/\varepsilon} + (1-t)w^L\mu(\bar{l}_\theta - l^e). \tag{65}$$

We denote the optimal effective labor supply for an individual with ability θ as l_θ^e. It is determined by the first-order condition of the Lagrangian with respect to l^e and by the rationing constraint:

$$l_\theta^e = (\theta^\beta(1-\lambda)(1-t)w^L)^\varepsilon \quad if\ \mu = 0, \tag{66}$$

$$l_\theta^e = \bar{l}_\theta \quad if\ \mu > 0. \tag{67}$$

In the absence of rationing the constraint is slack, such that $\mu = 0$, and the

[18] We restrict attention to skill-independent tax instruments. We do not formally analyze the case of skill-dependent tax instruments as the results of it are trivial. The effects of a higher minimum wage can be shown to be exactly mimicked by an increase in the low-skilled tax rate and a decrease in the high-skilled tax rate, leaving the minimum wage redundant (on this, also see Lee and Saez, 2012).

solution for effective labor supply reduces to the one obtained in previous sections: $l^e_\theta = l^L_\theta = (\theta^\beta(1-t)w^L)^\varepsilon$. We call l^L_θ notional labor supply, i.e., the number of hours the worker would optimally like to supply. If the constraint is binding, such that $\mu > 0$, effective labor supply is fully determined by the rationing constraint and $l^e_\theta = \bar{l}_\theta$. Notice from the first-order condition that a minimum wage acts as an implicit tax on labor supply through raising the shadow price of labor supply μ.

Thus, an individual would like to work l^L_θ hours, but if rationed is forced to work $\bar{l}_\theta < l^L_\theta$ instead. Without loss of generality we denote the hours restriction as a proportion of notional labor supply such that $\bar{l}_\theta \equiv (1-u_\theta)l^L_\theta$. We call u_θ the rationing schedule which may or may not depend on θ. It is important to distinguish u_θ from the unemployment rate as we have defined it. While in previous sections u stands for the proportion of low-skilled individuals that are unemployed, in this section u_θ stands for the proportion of hours that are underemployed. To determine individual labor supply we need to know the specific functional form of the rationing schedule. As discussed above, it is empirically unclear how rationing should depend on θ, but we assume in this section that the rationing schedule is efficient. This implies that the functional form of u_θ is such that the marginal utility of an extra hour of work is equal for every unskilled worker. Had this not been the case it would be efficiency improving to marginally decrease rationing of the high marginal utility worker and increase rationing of the low marginal utility worker. The marginal utility of being allowed to work an extra hour of work, in terms of the net wage, is given by the shadow price of labor supply μ. Substituting for $l^e_\theta = (1-u_\theta)(\theta^\beta(1-t)w^L)^\varepsilon$ in the first-order condition, we can write the shadow price as $\mu = 1-(1-u_\theta)^{\frac{1}{\varepsilon}}$. For rationing to be efficient, the shadow price should be independent of θ, and thus we require that:

$$\frac{\mathrm{d}\mu}{\mathrm{d}\theta} = \frac{1}{\varepsilon}(1-u_\theta)^{\frac{1}{\varepsilon}}\frac{\mathrm{d}u_\theta/\mathrm{d}\theta}{1-u_\theta} = 0 \tag{68}$$

This equation tells us that for rationing to be efficient, we necessarily have that $\mathrm{d}u_\theta/\mathrm{d}\theta = 0$. Hence, efficient rationing requires that the ration, as proportion of the notionally supplied number of hours, is equal for every low-skilled worker. The crucial assumption underlying this result is that the compensated elasticity of labor supply, ε, is identical for every low-skilled worker. This assumption implies that substitution effects and thus dead-weight losses are identical for every worker that faces the same ration u_θ. Throughout the remainder of this section we are exclusively interested in efficient rationing schedules and thus write $u_\theta = u$.

Substituting for effective labor supply, l^e_θ, we can write the indirect utility function for low-skilled workers as:

$$V^L_\theta = T + \left(1 - \frac{\varepsilon}{1+\varepsilon}(1-u)^{\frac{1}{\varepsilon}}\right)(1-u)\theta^{\beta\varepsilon}((1-t)w^L)^{1+\varepsilon}. \tag{69}$$

Notice that, for $u = 0$, this collapses to the low-skilled utility in the case of extensive rationing. Furthermore, it can easily be shown that, in the absence of rationing, $\partial V^L_\theta/\partial u = 0$, such that a marginal increase in rationing does not affect low-skilled utility. However, for positive and increasing values of rationing, $\partial V^L_\theta/\partial u$ is negative and decreasing.[19] As a direct consequence, in the absence of rationing, an increase in rationing only has a second-order effect on the cut-off ability level, Θ. Hence, a marginal increase in the minimum wage above the market-clearing wage only affects the human capital decision through a decrease in the skill premium, w^H/w^L. Only for higher levels of the minimum wage, further rationing causes an offsetting response in human capital. See the appendix for a full derivation of the comparative statics.

The comparative statics equation for skill formation is now given by:

$$\tilde{N}^H = \left(\frac{\sigma-1}{\sigma} - (1-u)^{\frac{1}{\varepsilon}}\right)\frac{\sigma}{\eta'}\frac{\tilde{w}^L}{\alpha}, \tag{70}$$

where $\eta' > 0$ is a composite term describing the shape of the income distribution around ability Θ (see appendix). A higher minimum wage will lead to more skill formation if and only if $\frac{\sigma-1}{\sigma} > (1-u)^{\frac{1}{\varepsilon}}$. This is more likely to hold if σ and u are large and if ε is small. Table 1 shows the critical levels of u, above which a higher minimum wage leads to higher skill formation and below which it leads to less skill formation. We show this for values of σ between 1.5 and 2.5. A value of 1.5 seems to be reasonable, although both lower and higher values are found in the literature (Katz and David, 1999). Notice that for $\sigma \leq 1$, a minimum wage always leads

[19] First derivative is given by $\partial V^L_\theta/\partial u = \left((1-u)^{\frac{1}{\varepsilon}} - 1\right)\theta^{\beta\varepsilon}((1-t)w^L)^{1+\varepsilon} < 0$, second derivative by $\partial^2 V^L_\theta/\partial u^2 = -\frac{1}{\varepsilon}(1-u)^{\frac{1-\varepsilon}{\varepsilon}}\theta^{\beta\varepsilon}((1-t)w^L)^{1+\varepsilon} < 0$.

to lower skill formation. Furthermore, we choose values of ε between 0.2 and 0.4, which seems to be a reasonable range (see, e.g., Blundell and MaCurdy, 1999).

The critical values of the intensive unemployment rate lie between 0.10 (for $\varepsilon = 0.2$ and $\sigma = 2.5$) and 0.36 (for $\varepsilon = 0.4$ and $\sigma = 1.5$). Empirical evidence on the degree of working hour restrictions varies widely. There are studies observing employees working less hours than desired (e.g., Kahn and Lang, 1991; Dickens and Lundberg, 1993; Bloemen, 2008) and studies observing employees actually working more hours than they desire (e.g., Stewart and Swaffield, 1997; Böheim and Taylor, 2004). The largest rationing proportion, which is based on a sample of low-income workers, is found by Dickens and Lundberg (1993) and is with 20 percent well within the range of above table. However, under more conservative findings for u of around 10 percent (as in Kahn and Lang, 1991) it is very unlikely that, with efficient rationing, a minimum wage leads to more skill formation.

Table 1: Critical values for u

$u^* = 1 - \left(\frac{\sigma-1}{\sigma}\right)^{\varepsilon}$			
	$\sigma = 1.5$	$\sigma = 2.0$	$\sigma = 2.5$
$\varepsilon = 0.2$	0.20	0.13	0.10
$\varepsilon = 0.3$	0.28	0.19	0.14
$\varepsilon = 0.4$	0.36	0.24	0.18

6.3 Optimal policy

As there are no unemployed when rationing occurs on the intensive margin, the social welfare function simplifies to:

$$(71) \qquad \mathcal{W} \equiv \int_{\underline{\theta}}^{\Theta} \Psi(V_\theta^L)\mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} \Psi(V_\theta^H)\mathrm{d}G(\theta),$$

whereas the government budget is still given by equation (30). Forming the Lagrangian and taking derivatives we find that the first-order conditions for the tax rate and transfer do not change. We find the following first-order condition for the optimal minimum wage:

$$(72) \qquad (1-\xi^L)\overline{\Psi'(V^L)} - (1-\xi^H)\overline{\Psi'(V^H)}$$
$$- \left(1-(1-u)^{\frac{1}{\varepsilon}}\right)(1-\xi^L)\overline{\Psi'(V^L)}\left(\frac{w^L}{1-u}\frac{\mathrm{d}u}{\mathrm{d}w^L} - \varepsilon\right)$$
$$- \lambda\frac{t}{1-t}\left(\frac{w^L}{1-u}\frac{\mathrm{d}u}{\mathrm{d}w^L} - \frac{w^H l_\Theta^H - w^L(1-u)l_\Theta^L}{L^L}\frac{\mathrm{d}N^H}{\mathrm{d}w^L}\right) = 0.$$

The first line gives the redistributional gain of an increase in the minimum wage, which is the same as before. The second line gives the utility loss associated with more rationing due to a higher minimum wage. The third line gives the social welfare loss of an eroding tax base.

To focus on the desirability of a minimum wage, we analyze the first-order condition for $u = 0$. Note that in that case $\left(1-(1-u)^{\frac{1}{\varepsilon}}\right) = 0$ and $w^H l_\Theta^H - w^L(1-u)l_\Theta^L = 0$. Substituting this into the first-order condition, we obtain the following condition for a minimum wage to be desirable:

$$(73) \qquad \frac{\overline{\Psi'(V^L)} - \overline{\Psi'(V^H)}}{\lambda} - \frac{\overline{\Psi'(V^L)}}{\lambda}\xi^L + \frac{\overline{\Psi'(V^H)}}{\lambda}\xi^H > \frac{t}{1-t}\frac{\tilde{u}}{\tilde{w}^L}.$$

This expression is almost identical to the analogue expressions in previous sections. The only term that is missing is the marginal utility loss from rationing which, as we discussed above, is only second-order under efficient rationing. Hence, in the absence of taxation, a marginal increase in the minimum wage above the market clearing wage only has distributional gains, equal to the left-hand side of equation (73). With efficient rationing, and in the absence of taxation, a minimum wage is therefore unambiguously desirable.

If there is a positive tax rate, a minimum wage erodes the tax base as more rationing leads to less workers paying taxes. This welfare loss is represented by the right-hand side of equation (73) and is increasing in the tax rate. In order to determine the desirability of a minimum wage at the optimal tax system we substitute for $\left(\overline{\Psi'(V^L)} - \overline{\Psi'(V^H)}\right)/\lambda$ from the first-order condition for the tax rate, $t/(1-t) = \xi/\varepsilon$. This yields:

$$(74) \qquad \frac{t}{1-t}\left(\frac{\varepsilon}{\alpha - N^H} - \frac{\tilde{u}}{\tilde{w}^L}\right) > \frac{N^L}{\alpha - N^H}\frac{\overline{\Psi'(V^L)}}{\lambda}\xi^L + \frac{N^H}{\alpha - N^H}\frac{\overline{\Psi'(V^H)}}{\lambda}\xi^H.$$

This condition is almost identical to equation (45), the only differences being the utility loss of unemployment, which drops out, and a slightly altered elasticity of unemployment $\tilde{u}/\tilde{w}_L = (\sigma + \varepsilon + \kappa')/\alpha$ (see appendix). The left-hand side gives the efficiency costs of redistributing between skill-groups by using income taxes instead of a minimum wage. In the case of income taxes, the efficiency costs consist of the tax base erosion caused by downward distorted intensive labor supply of both high-skilled and low-skilled workers. In the case of a higher minimum wage, the efficiency costs consist of the tax base erosion associated with underemployment. The right-hand side gives the within-group distributional advantage income taxes have over the minimum wage.

Proposition 8 *Similarly to the case with uniform unemployment, if rationing is efficient, the role of a minimum wage in an optimal skill-independent tax-benefit system is to complement the tax-benefit system by reducing the distributional imperfections of the income tax. Minimum wages help to redistribute more income between skill groups, so that income taxes can be better targeted at reducing inequality within skill groups. The desirability of a minimum wage depends on whether this benefit outweighs the loss in tax revenue due to higher rationing. Contrary to the case with uniform unemployment, there is no direct utility loss associated with a marginally binding minimum wage.*

7 Conclusion

This study indicates that minimum wages typically reduce income inequality between skilled and unskilled workers. It also diminishes inequality within the group of high-skilled workers, but raises inequality within the group of low-skilled workers. If government can differentiate its income tax rates between high-skilled and low-skilled taxes, such redistribution can also be attained by properly adjusting high- and low-skilled taxes. The one feature that distinguishes a minimum wage reform from a distributionally equivalent tax reform is its capacity to create involuntary unemployment among low-skilled workers. Besides the obvious costs of this unemployment, in terms of both private utility and public revenue losses, it generates a potential welfare gain by stimulating skill formation and, thereby, raising high-skilled tax revenue. To put it differently, a minimum wage exacerbates a distortion of redistributive taxes by raising unemployment, resulting in a smaller tax base. Simultaneously, it alleviates a distortion of redistributive taxation by raising skill formation, leading to higher tax revenue. The net effect determines the desirability of a minimum wage.

An important factor determining the optimality of a minimum wage is the degree to which education increases as a result of higher unemployment among the low-skilled. Theoretically this is driven for an important part by the assumption of how strongly the job chances of low-skilled workers on the skill margin are affected by unemployment. We assumed that every low-skilled worker's job chances are affected equally. Assuming, as Lee and Saez (2012) mostly do, that workers on the skill margin are hit first would drastically improve the case for a minimum wage as it would lead to a larger increase in skill formation. The welfare consequences of different rationing schedules are discussed in more detail in Gerritsen and Jacobs (2013) and Gerritsen (2013). In Gerritsen and Jacobs (2013), we study the desirability of a minimum wage in the presence of a more general rationing schedule and empirically calibrate the resulting desirability condition. Gerritsen (2013) treats the wage floor as given, and derives the implications for optimal tax policy.

Appendix 1: Efficient rationing

In the case of efficient rationing, indirect utility is represented by:

$$V_\theta^L = T + \left(1 - \frac{\varepsilon}{1+\varepsilon}(1-u)^{\frac{1}{\varepsilon}}\right)(1-u)\theta^{\beta\varepsilon}((1-t)w^L)^{1+\varepsilon}, \tag{75}$$

$$V_\theta^H = T + \frac{\theta^\varepsilon((1-t)w^H)^{1+\varepsilon}}{1+\varepsilon}. \tag{76}$$

For the cut-off level of ability, Θ, we need $V_\Theta^L = V_\Theta^H$, which implies:

$$\Theta^{(1-\beta)\varepsilon} = \left(1 - \frac{\varepsilon}{1+\varepsilon}(1-u)^{\frac{1}{\varepsilon}}\right)(1-u)(1+\varepsilon)\left(\frac{w^H}{w^L}\right)^{-(1+\varepsilon)}. \tag{77}$$

The rest of the equilibrium conditions consist of the firms' first-order conditions and the market clearing conditions.

$$F_H(L^H, L^L) = w^H, \tag{78}$$

$$F_L(L^H, L^L) = w^L, \tag{79}$$

$$L^H = \int_\Theta^{\overline{\theta}} l_\theta^H \mathrm{d}G(\theta), \tag{80}$$

$$L^L = (1-u)\int_{\underline{\theta}}^{\Theta} l_\theta^L \mathrm{d}G(\theta). \tag{81}$$

Log-linearizing the equilibrium conditions around an initial equilibrium yields the following equations:

$$(1-\beta)\varepsilon\tilde{\Theta} = (1+\varepsilon)(\tilde{w}^L - \tilde{w}^H) - \frac{1-(1-u)^{\frac{1}{\varepsilon}}}{1-\frac{\varepsilon}{1+\varepsilon}(1-u)^{\frac{1}{\varepsilon}}}\tilde{u}, \tag{82}$$

$$\tilde{w}^H = \frac{1-\alpha}{\sigma}(\tilde{L}^L - \tilde{L}^H), \tag{83}$$

$$\tilde{w}^L = \frac{\alpha}{\sigma}(\tilde{L}^H - \tilde{L}^L), \tag{84}$$

$$\tilde{L}^H = -\frac{l_\Theta^H \Theta g(\Theta)}{L^H}\tilde{\Theta} + \varepsilon(\tilde{w}^H - \tilde{t}), \tag{85}$$

$$\tilde{L}^L = \frac{(1-u) l_\Theta^L \Theta g(\Theta)}{L^L}\tilde{\Theta} - \tilde{u} + \varepsilon(\tilde{w}^L - \tilde{t}). \tag{86}$$

Moreover recall that $N^H = 1 - G(\Theta)$ and hence $\tilde{N}^H = -\Theta g(\Theta)\tilde{\Theta}$. Above equations can now be solved to express $\tilde{N}^H$ and $\tilde{u}$ in terms of the exogenous variables, $\tilde{w}^L$ and $\tilde{t}$:

$$\tilde{N}^H = \left(\frac{\sigma-1}{\sigma} - (1-u)^{\frac{1}{\varepsilon}}\right)\frac{\sigma}{\eta'}\frac{\tilde{w}^L}{\alpha}, \tag{87}$$

$$\tilde{u} = \left(\sigma + \varepsilon - \left(\frac{\sigma-1}{\sigma} - (1-u)^{\frac{1}{\varepsilon}}\right)\sigma\kappa'\right)\frac{\tilde{w}^L}{\alpha}, \tag{88}$$

where $\eta' = \left(\left(1 - \frac{\varepsilon}{1+\varepsilon}(1-u)^{\frac{1}{\varepsilon}}\right)\frac{(1-\beta)\varepsilon}{\Theta g(\Theta)} + \left(1-(1-u)^{\frac{1}{\varepsilon}}\right)\left(\frac{l^H_\Theta}{L^H} + \frac{(1-u)l^L_\Theta}{L^L}\right)\right)$ and $\kappa' = \left(\frac{l^H_\Theta}{L^H} + \frac{(1-u)l^L_\Theta}{L^L}\right)\eta'^{-1}$. The interpretation of these two comparative statics equations is similar to the case with extensive rationing. In particular, it shows that a higher minimum wage leads to more skill formation if and only if $(\sigma-1)/\sigma > (1-u)^{\frac{1}{\varepsilon}}$. Furthermore, it can be shown that rationing always increases due to a higher minimum wage.

The Lagrangian for the government's optimization problem is the following:

$$\mathcal{L}(w^L, t, T) = \int_{\underline{\theta}}^{\Theta} \Psi(V^L_\theta)\mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} \Psi(V^H_\theta)\mathrm{d}G(\theta) + \lambda(tw^L L^L + tw^H L^H - T - E). \tag{89}$$

This leads to the following first-order conditions:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial w^L} = {}& (1-t)\left(\int_{\underline{\theta}}^{\Theta} \Psi'(V^L_\theta)(1-u)l^L_\theta \mathrm{d}G(\theta) - \frac{L^L}{L^H}\int_{\Theta}^{\overline{\theta}} \Psi'(V^H_\theta)l^H_\theta \mathrm{d}G(\theta)\right) \\ & - (1-t)\left(1-(1-u)^{\frac{1}{\varepsilon}}\right)\int_{\underline{\theta}}^{\Theta} \Psi'(V^L_\theta)(1-u)l^L_\theta \mathrm{d}G(\theta)\left(\frac{\tilde{u}}{\tilde{w}^L} - \varepsilon\right) \\ & - \lambda t L^L \frac{\tilde{u}}{\tilde{w}^L} + \lambda(tw^H l^H_\Theta - tw^L(1-u)l^L_\Theta)\frac{\mathrm{d}N^H}{\mathrm{d}w^L} = 0,\end{aligned} \tag{90}$$

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial t} = {}& -\int_{\underline{\theta}}^{\Theta} \Psi'(V^L_\theta)w^L(1-u)l^L_\theta \mathrm{d}G(\theta) - \int_{\Theta}^{\overline{\theta}} \Psi'(V^H_\theta)w^H l^H_\theta \mathrm{d}G(\theta) \\ & - \varepsilon\left(1-(1-u)^{\frac{1}{\varepsilon}}\right)\int_{\underline{\theta}}^{\Theta} \Psi'(V^L_\theta)w^L(1-u)l^L_\theta \mathrm{d}G(\theta) \\ & + \lambda\left(w^L L^L + w^H L^H - \varepsilon\frac{t}{1-t}(w^L L^L + w^H L^H)\right) = 0,\end{aligned} \tag{91}$$

$$\frac{\partial \mathcal{L}}{\partial T} = \int_{\underline{\theta}}^{\Theta} \Psi'(V^L_\theta)\mathrm{d}G(\theta) + \int_{\Theta}^{\overline{\theta}} \Psi'(V^H_\theta)\mathrm{d}G(\theta) - \lambda = 0. \tag{92}$$

The first-order condition for the minimum wage can be simplified as follows:

$$(93)\qquad (1-\xi^L)\frac{\overline{\Psi'(V^L)}}{\lambda} - (1-\xi^H)\frac{\overline{\Psi'(V^H)}}{\lambda} = \left(1-(1-u)^{\frac{1}{\varepsilon}}\right)(1-\xi^L)\frac{\overline{\Psi'(V^L)}}{\lambda}\left(\frac{\tilde{u}}{\tilde{w}^L}-\varepsilon\right) + \frac{t}{1-t}\frac{\tilde{u}}{\tilde{w}^L} - \frac{t}{1-t}\frac{w^H l^H_\Theta - w^L(1-u)l^L_\Theta}{w^L L^L}\frac{\tilde{N}^H}{\tilde{w}^L}.$$

The left-hand side gives the distributional gain of a higher minimum wage. The right-hand side gives the costs associated with higher unemployment and the inability of low-skilled workers to react by altering their hours worked (first term) and the costs associated with an exacerbation of the tax distortion due to higher rationing (second term) and lower high-skilled labor supply (third term). In the case of $u = 0$, the first and third terms only imply second-order welfare effects.

Rearranging and substituting for ξ and ξ^L yields the following expression for the optimal income tax:

$$(94)\qquad \frac{t}{1-t} = \frac{\xi}{\varepsilon} - \left(1-(1-u)^{\frac{1}{\varepsilon}}\right)(1-\alpha)(1-\xi^L)\frac{\overline{\Psi'(V^L)}}{\lambda}.$$

Hence, compared to the case with rationing along the extensive margin, there is an additional cost of taxation which pulls down the optimal tax level. In the case of extensive rationing, low-skilled workers coped with higher taxation by working less hours, thereby absorbing part of the direct utility costs. However, if these workers are intensively rationed, they will not reduce their working hours as they already work less than they would prefer. In the case of $u = 0$ this cost naturally vanishes.

References

Acemoglu, Daron, and Jörn-Steffen Pischke. 1999. "The structure of wages and investment in general training." *Journal of Political Economy*, 107(3): 539–572.

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

Aronsson, Thomas, and Erkki Koskela. 2008. "Optimal redistributive taxation and provision of public input goods in an economy with outsourcing and unemployment." *CESifo Working Paper No. 2481.*

Arulampalam, Wiji, Alison L Booth, and Mark L Bryan. 2004. "Training and the new minimum wage." *The Economic Journal*, 114(494): C87–C94.

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

Bhaskar, Venkataraman, and Ted To. 1999. "Minimum wages for Ronald McDonald monopsonies: A theory of monopsonistic competition." *The Economic Journal*, 109(455): 190–203.

Bloemen, Hans G. 2008. "Job search, hours restrictions, and desired hours of work." *Journal of Labor Economics*, 26(1): 137–179.

Blundell, Richard, and Thomas MaCurdy. 1999. "Labor supply: A review of alternative approaches." In *Handbook of Labor Economics*. Vol. 3, , ed. Orley Ashenfelter and David Card, 1559–1695. Amsterdam:Elsevier.

Boadway, Robin, and Katherine Cuff. 2001. "A minimum wage can be welfare-improving and employment-enhancing." *European Economic Review*, 45(3): 553–576.

Böheim, René, and Mark P Taylor. 2004. "Actual and preferred working hours." *British Journal of Industrial Relations*, 42(1): 149–166.

Cahuc, Pierre, and Guy Laroque. 2013. "Optimal taxation and monopsonistic labor market: does monopsony justify the minimum wage?" *Journal of Public Economic Theory.*

Cahuc, Pierre, and Philippe Michel. 1996. "Minimum wage unemployment and growth." *European Economic Review*, 40(7): 1463–1482.

Cahuc, Pierre, Anne Saint-Martin, and André Zylberberg. 2001. “The consequences of the minimum wage when other wages are bargained over.” *European Economic Review*, 45(2): 337–352.

Card, David. 1992. “Using regional variation in wages to measure the effects of the federal minimum wage.” *Industrial and Labor Relations Review*, 46(1): 22–37.

Card, David, and Alan B Krueger. 1995. *Myth and measurement: The new economics of the minimum wage.* Princeton:Princeton University Press.

Chaplin, Duncan D, Mark D Turner, and Andreas D Pape. 2003. “Minimum wages and school enrollment of teenagers: A look at the 1990s.” *Economics of Education Review*, 22(1): 11–21.

Couch, Kenneth A, and David C Wittenburg. 2001. “The response of hours of work to increases in the minimum wage.” *Southern Economic Journal*, 68(1): 171–177.

Crofton, Stephanie O, William L Anderson, and Emily C Rawe. 2009. “Do higher real minimum wages lead to more high school dropouts? Evidence from Maryland across races, 1993–2004.” *American Journal of Economics and Sociology*, 68(2): 445–464.

Cunningham, James. 1981. “The impact of minimum wages on youth employment, hours of work, and school attendance: Cross-sectional evidence from the 1960 and 1970 censuses.” In *The Economics of Legal Minimum Wages.* , ed. Simon Rottenberg, 88–123. Washington, D.C.:American Enterprise Institute for Public Policy Research.

Danziger, Eliav, and Leif Danziger. 2010. “A graduated minimum wage with optimal taxation.” Mimeo, Princeton University and Ben Gurion University.

Davis, Lucas W, and Lutz Kilian. 2011. “The allocative cost of price ceilings in the U.S. residential market for natural gas.” *Journal of Political Economy*, 119(2): 212–241.

Deacon, Robert T, and Jon Sonstelie. 1989. "The welfare costs of rationing by waiting." *Economic Inquiry*, 27(2): 179–196.

Deltas, George. 2007. "Can a minimum wage increase employment and reduce prices in a neoclassical perfect information economy?" *Journal of Mathematical Economics*, 43(6): 657–674.

Dickens, William T, and Shelly J Lundberg. 1993. "Hours Restrictions and Labor Supply." *International Economic Review*, 34(1): 169–92.

Dolado, Juan J, Florentino Felgueroso, and Juan F Jimeno. 2000. "The role of the minimum wage in the welfare state: An appraisal." *IZA Discussion Paper No. 152*.

Dustmann, Christian, and Uta Schönberg. 2009. "Training and union wages." *The Review of Economics and Statistics*, 91(2): 363–376.

Ehrenberg, Ronald G, and Alan J Marcus. 1982. "Minimum wages and teenagers' enrollment-employment outcomes: a multinomial logit model." *Journal of Human Resources*, 17(1): 39–58.

Fairris, David, and Roberto Pedace. 2004. "The impact of minimum wages on job training: an empirical exploration with establishment data." *Southern Economic Journal*, 70(3): 566–583.

Feldstein, Martin S. 1972. "Distributional equity and the optimal structure of public prices." *The American Economic Review*, 62(1/2): 32–36.

Flinn, Christopher J. 2006. "Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates." *Econometrica*, 74(4): 1013–1062.

Frech, Harry E, and William C Lee. 1987. "The welfare cost of rationing-by-queuing across markets: Theory and estimates from the U.S. gasoline crises." *The Quarterly Journal of Economics*, 102(1): 97–108.

Freeman, Richard B. 1996. "The minimum wage as a redistributive tool." *The Economic Journal*, 106(436): 639–649.

Gerritsen, Aart. 2013. "Equity and efficiency in rationed labor markets." Mimeo, Erasmus University Rotterdam.

Gerritsen, Aart, and Bas Jacobs. 2013. "Is a minimum wage an appropriate instrument for redistribution?" Mimeo, Erasmus University Rotterdam.

Glaeser, Edward L., and Erzo F.P. Luttmer. 2003. "The misallocation of housing under rent control." *The American Economic Review*, 93(4): 1027–1046.

Grossberg, Adam J, and Paul Sicilian. 1999. "Minimum wages, on-the-job training, and wage growth." *Southern Economic Journal*, 65(3): 539–556.

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

Hashimoto, Masanori. 1982. "Minimum wage effects on training on the job." *American Economic Review*, 72(5): 1070–1087.

Hosios, Arthur J. 1990. "On the efficiency of matching and related models of search and unemployment." *The Review of Economic Studies*, 57(2): 279–298.

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.

Immervoll, Herwig. 2013. "Minimum wages, minimum labour costs, and the tax treatment of low-wage employment." *IZA Discussion Paper No. 2555.*

Jacobs, Bas, Dirk Schindler, and Hongyan Yang. 2012. "Optimal Taxation of Risky Human Capital." *The Scandinavian Journal of Economics*, 114(3): 908–931.

Kahn, Shulamit, and Kevin Lang. 1991. "The effect of hours constraints on labor supply estimates." *The Review of Economics and Statistics*, 73(4): 605–611.

Katz, Lawrence F, and H David. 1999. "Changes in the wage structure and earnings inequality." In *Handbook of Labor Economics*. Vol. 3, , ed. Orley Ashenfelter and David Card, 1463–1555. Amsterdam:Elsevier.

Landon, Stuart. 1997. "High school enrollment, minimum wages and education spending." *Canadian Public Policy/Analyse de Politiques*, 23(2): 141–163.

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

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

Manning, Alan. 2003. *Monopsony in motion: Imperfect competition in labor markets.* Princeton:Princeton University Press.

Marceau, Nicolas, and Robin Boadway. 1994. "Minimum wage legislation and unemployment insurance as instruments for redistribution." *The Scandinavian Journal of Economics*, 96(1): 67–81.

Mattila, J Peter. 1981. "The impact of minimum wages on teenage schooling and on the part-time/full-time employment of youths." In *The Economics of Legal Minimum Wages.* , ed. Simon Rottenberg, 61–87. Washington, D.C.:American Enterprise Institute for Public Policy Research.

Mincer, Jacob, and Linda S Leighton. 1981. "Effect of minimum wages on human capital formation." In *The Economics of Legal Minimum Wages.* , ed. Simon Rottenberg, 155–173. Washington, D.C.:American Enterprise Institute for Public Policy Research.

Mirrlees, James A. 1971. "An exploration in the theory of optimum income taxation." *Review of Economic Studies*, 38(2): 175–208.

Montmarquette, Claude, Nathalie Viennot-Briot, and Marcel Dagenais. 2007. "Dropout, school performance, and working while in school." *The Review of Economics and Statistics*, 89(4): 752–760.

Neumark, David, and William Wascher. 1995*a*. "Minimum wage effects on employment and school enrollment." *Journal of Business & Economic Statistics*, 13(2): 199–206.

Neumark, David, and William Wascher. 1995*b*. "Minimum-wage effects on school and work transitions of teenagers." *The American Economic Review*, 85(2): 244–249.

Neumark, David, and William Wascher. 2001. "Minimum wages and training revisited." *Journal of Labor Economics*, 19(3): 563–595.

Neumark, David, and William Wascher. 2003. "Minimum wages and skill acquisition: Another look at schooling effects." *Economics of Education Review*, 22(1): 1–10.

Neumark, David, and William Wascher. 2006. "Minimum wages and employment: A review of evidence from the new minimum wage research." *National Bureau of Economic Research Working Paper No. 12663*.

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

Pacheco, Gail A, and Amy A Cruickshank. 2007. "Minimum wage effects on educational enrollments in New Zealand." *Economics of Education Review*, 26(5): 574–587.

Ragan, James F. 1977. "Minimum wages and the youth labor market." *The Review of Economics and Statistics*, 59(2): 129–136.

Rebitzer, James B, and Lowell J Taylor. 1995. "The consequences of minimum wage laws some new theoretical ideas." *Journal of Public Economics*, 56(2): 245–255.

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

Rosen, Sherwin. 1972. "Learning and experience in the labor market." *Journal of Human Resources*, 7(3): 326–342.

Saint-Paul, Gilles. 1996. "Unemployment and increasing private returns to human capital." *Journal of Public Economics*, 61(1): 1–20.

Schiller, Bradley R. 1994. "Moving up: The training and wage gains of minimum-wage entrants." *Social Science Quarterly*, 75(3): 622–636.

Stewart, Mark B, and Joanna K Swaffield. 1997. "Constraints on the desired hours of work of British men." *The Economic Journal*, 107(441): 520–535.

Stigler, George J. 1946. "The economics of minimum wage legislation." *The American Economic Review*, 36(3): 358–365.

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

Tobin, James. 1952. "A Survey of the Theory of Rationing." *Econometrica*, 20(4): 521–553.

Zavodny, Madeline. 2000. "The effect of the minimum wage on employment and hours." *Labour Economics*, 7(6): 729–750.