TI 2016-100/VI
Tinbergen Institute Discussion Paper

Is a Minimum Wage an Appropriate Instrument for Redistribution?

Aart Gerritsen[1]
Bas Jacobs[2]

[1] *Max Planck Institute for Tax Law and Public Finance, Germany;*
[2] *Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute, The Netherlands.*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Is a Minimum Wage an Appropriate Instrument for Redistribution?

Aart Gerritsen and Bas Jacobs*

November 16, 2016

We analyze the redistributional (dis)advantages of a minimum wage over income taxation in competitive labor markets, without imposing assumptions on the (in)efficiency of labor rationing. Compared to a distributionally equivalent tax change, a minimum-wage increase raises involuntary unemployment, but also raises skill formation as some individuals avoid unemployment. A minimum wage is an appropriate instrument for redistribution if and only if the public revenue gains from additional skill formation outweigh both the public revenue losses from additional unemployment and the utility losses of inefficient labor rationing. We show that this critically depends on how labor rationing is distributed among workers. A necessary condition for the desirability of a minimum-wage increase is that the public revenue gains from higher skill formation outweigh the revenue losses from higher unemployment. We write this condition in terms of measurable sufficient statistics. Our empirical analysis suggests that a minimum-wage increase is undesirable in nearly all OECD countries. A reduction in the minimum wage, along with tax adjustments that keep net incomes constant, would yield a Pareto improvement.

JEL: D61; H21; J21; J24; J38
Keywords: Minimum wage; optimal redistribution; unemployment; skill formation

*Gerritsen: Max Planck Institute for Tax Law and Public Finance, Department of Public Economics, Marstallplatz 1, 80539 Munich, Germany, aart.gerritsen@tax.mpg.de. 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/. This is a drastically revised update of an earlier working paper version that circulated under the same title. We thank Robin Boadway, Johann Brunner, Carlos da Costa, Laurence Jacquet, Etienne Lehmann, Emmanuel Saez, Coen Teulings, Hendrik Vrijburg, Floris Zoutman, and participants of numerous seminars and conferences for helpful comments and discussions. Financial support from The Netherlands Organization for Scientific Research (NWO Vidi Grant No. 452-07-013) is gratefully acknowledged.

1 Introduction

Is a minimum wage an appropriate instrument for redistribution? Proponents of the minimum wage emphasize its distributional benefits as it raises the earnings of low-skilled workers. However, the valuation of these distributional effects requires an intrinsically political judgment on which economists have little to say. Opponents of the minimum wage emphasize that it reduces employment, although there is no empirical consensus on the adverse employment effects of the minimum wage (cf. Card and Krueger, 1995; Neumark and Wascher, 2006; Schmitt, 2013). Therefore, as long as the minimum-wage debate is framed in terms of a trade-off between more redistribution of income and higher unemployment, the answer to this paper's title remains contentious. We avoid these discussions by analyzing the minimum wage not in isolation from, but in comparison to redistribution via the tax and transfer system. This allows us to assess the desirability of a minimum wage without relying on either political judgments about the value of redistribution or controversial estimates of the labor-demand effects of a minimum wage.

We develop a relatively standard model of occupational choice and optimal income redistribution, based on Diamond (1980) and Saez (2002), which we augment with a minimum-wage policy. Firms demand high- and low-skilled labor in perfectly competitive labor markets. A minimum wage might be binding for low-skilled workers, but not for high-skilled workers. High-skilled workers in our model could roughly be thought of as workers that completed upper secondary education. Individuals are assumed to be heterogeneous in their disutility of work in low-skilled and high-skilled occupations. Depending on their disutility, they decide to do high-skilled work or low-skilled work, or to be (voluntarily) unemployed. A binding minimum wage fixes the low-skilled wage, thereby rationing some individuals out of the low-skilled labor market. Rationed individuals are unable to find a low-skilled job, and are, therefore, forced to choose between the high-skilled occupation and (involuntary) unemployment. We remain entirely agnostic about which individuals are, and which individuals are not, able to find a low-skilled job. Hence, we adopt a fully general 'rationing schedule' that determines how rationing is distributed among individuals with different disutilities of work. This contrasts with most of the theoretical literature, which typically assumes that labor rationing is efficient, i.e., that rationing is exclusively concentrated on individuals with the highest disutility of work. The general rationing schedule we adopt is more in line with the fact that we lack good empirical evidence on the distribution of labor rationing.

An important feature of our model is that the distributional consequences of a minimum-wage increase can be perfectly replicated by a change in income taxes. Perfectly competitive, profit-maximizing firms pay for the minimum wage by reducing high-skilled wages. As a result, the minimum wage redistributes income from individuals with a high income to individuals with a low income. The income tax could achieve the same income redistribution by reducing low-income taxes and raising high-income taxes. The relevant question therefore is: can a minimum wage achieve a given amount of income redistribution from high- to low-skilled workers at lower efficiency costs than the income tax? We obtain the answer to this question by analyzing a policy reform that raises the minimum wage and simultaneously adjusts taxes to leave net incomes of both high- and low-skilled workers unaffected. We label this policy reform a *net-income-neutral* (NIN) minimum-wage increase. The effects of a NIN minimum-wage increase indicate how a minimum-wage increase differs from a distributionally equivalent change in taxes.[1] As the policy reform leaves net incomes constant, it allows us to assess the desirability of a minimum-wage increase exclusively in terms of economic efficiency, without resorting to political judgments regarding the desirability of income redistribution. Our paper makes four contributions.

First, we show that a minimum-wage increase differs from a distributionally equivalent change in taxes by creating more unemployment *and* more high-skilled employment. Intuitively, a NIN increase in the minimum wage raises the wage costs of low-skilled workers. This reduces low-skilled labor demand and, therefore, rations some individuals out of the low-skilled labor market. While some of these rationed individuals become unemployed, others might prefer high-skilled employment over unemployment, and, therefore, choose to become high-skilled. The magnitude of the effects on unemployment and high-skilled employment are crucially determined by the rationing schedule. If rationing is mostly concentrated on individuals with a high (low) disutility of high-skilled work, then rationing mostly leads to higher unemployment (high-skilled employment). All other effects of a minimum-wage increase are identical to the effects of a distributionally equivalent tax change.

Second, we derive a simple condition under which the minimum wage is an appropriate instrument for redistribution. This is the case if a minimum-wage increase is more desirable than a distributionally equivalent change in taxes. We derive the desirability condition by showing that a NIN increase in the minimum wage has three welfare-relevant effects. (i) Labor rationing reduces utility as long as the rationed individuals strictly prefer low-skilled work over unemployment or high-skilled work. (ii) Increased unemployment reduces public revenue if the unemployed pay less taxes than low-skilled workers. (iii) Increased high-skilled employment raises public revenue if high-skilled workers pay more taxes than low-skilled workers. A minimum-wage increase is more desirable than a distributionally equivalent change in taxes if and only if the revenue gains from increased high-skilled employment are sufficiently high to compensate for the revenue losses of increased unemployment and the utility losses from inefficient rationing. Moreover, if the desirability condition holds in the tax optimum without a minimum wage, then a

[1] Our conclusions do not in any way depend on this particular reform, but it allows for the most transparent comparison between a minimum wage and the tax and transfer system.

minimum wage is necessarily part of the overall policy optimum. Intuitively, the second-best role of the minimum wage is to alleviate tax distortions on skill formation. Both optimal and observed taxes tend to increase with income, and therefore distort labor participation and skill decisions downwards. By raising high-skilled employment, a NIN minimum-wage increase alleviates the tax distortion on skill formation. However, by raising unemployment, it also exacerbates the tax distortion on labor participation. A minimum-wage increase is more desirable than a distributionally equivalent tax change if and only if the gains from smaller distortions of skill formation outweigh the costs from both larger distortions of participation and the utility losses associated with inefficient rationing.

Third, we derive a *necessary* condition for the desirability of a minimum-wage increase that is solely expressed in terms of empirically recoverable statistics. Because a NIN minimum-wage increase leads to utility losses from inefficient rationing, the revenue gains from increased high-skilled employment must at least outweigh the revenue losses from increased unemployment for the minimum-wage increase to be desirable. This only holds if the increase in high-skilled employment is large enough relative to the increase in unemployment. That is, rationing should be sufficiently concentrated on individuals that prefer high-skilled employment over unemployment. The rationing schedule is therefore a crucial determinant of the desirability of a minimum wage. While we lack any empirical evidence on the rationing schedule, we show that the effect of unemployment on skill formation can function as a *sufficient statistic*, foregoing the need to determine the rationing schedule. Consequently, we can express the necessary condition for the desirability of a minimum-wage increase in terms of three empirically recoverable sufficient statistics: the tax wedge on participation, the tax wedge on skill formation, and the effect of unemployment on skill formation. This condition does not rely on controversial estimates of the labor-demand effects of a minimum wage. The reason is that both the benefits (more high-skilled employment) and the costs (more unemployment) of the NIN minimum-wage increase are proportional to the reduction in labor demand.

Fourth, we bring the necessary condition for the desirability of a NIN minimum-wage increase to the data, and do so for a large number of OECD countries. We review empirical estimates of the semi-elasticity of school enrollment rates with respect to low-skilled unemployment rates, and use these estimates to calibrate the effect of unemployment on high-skilled employment. We use OECD data to calibrate the tax wedges on participation and skill formation. For most countries, though possibly not for the United States, we find that an increase in the minimum wage is strictly less desirable than a distributionally equivalent change in taxes. Thus, an increase in income redistribution could in these countries be most efficiently achieved by adjusting income taxes. Moreover, those countries could obtain a Pareto improvement by reducing the minimum wage while adjusting income taxes to neutralize the effects on the income distribution. Such a combined reform would raise both public revenue and utility by reducing rationing. Our empirical analysis therefore suggests that a minimum wage is *not* an appropriate instrument for redistribution.[2]

The remainder of our paper is structured as follows. Section 2 discusses earlier literature. Section 3 introduces the theoretical model. Section 4 defines the NIN minimum-wage increase and derives its comparative statics. Section 5 derives the welfare effects of a NIN minimum-wage increase, and the conditions under which a minimum wage-increase is more desirable than a distributionally equivalent tax change. It also provides a detailed discussion on how these conditions relate to findings in previous studies, and considers the robustness of our results with respect to relaxing a number of theoretical assumptions. Section 6 determines whether a binding minimum wage could be a desirable supplement to the tax optimum. Section 7 brings the necessary condition for the desirability of a NIN minimum-wage increase to the data. Section 8 concludes with some final thoughts.

2 Earlier literature

This paper contributes to the literature that studies minimum wages in models of optimal income redistribution and competitive labor markets.[3] Most studies consider the two-type optimal-tax framework of Stern (1982) and Stiglitz (1982). Using this framework, Allen (1987) and Guesnerie and Roberts (1987) find that a binding minimum wage is undesirable if the government could set nonlinear income taxes.[4] While they only consider underemployment on the labor-hours margin, Marceau and Boadway (1994) extend these analyses by considering involuntary unemployment on the extensive margin. They find that a minimum wage can only be a desirable policy if the unemployed receive smaller transfers than the low-skilled employed. In that case, labor participation is distorted upwards and a minimum wage alleviates this distortion by pushing some low-skilled individuals out of the labor market. We contribute to these studies by endogenizing the skill decisions of individuals. As a result, a minimum wage not only pushes some individuals

[2] This does not imply that the minimum wage might not be desirable for reasons that are not directly related to income redistribution, such as correcting market imperfections caused by monopsony power or insufficient bargaining power.

[3] A separate literature analyzes the welfare properties of a minimum wage in non-competitive labor markets. Even if our results indicate that a minimum wage might not be an appropriate instrument for redistribution, this literature suggests that minimum wages could still be desirable to reduce labor-market frictions. Notable studies include Hungerbühler and Lehmann (2009) and Cahuc and Laroque (2013), who both consider a minimum wage along with optimal taxes. Cahuc and Laroque (2013) show that a minimum wage is not useful to reduce monopsony problems on the labor market as long as the government has sufficient tax instruments at its disposal. Hungerbühler and Lehmann (2009) do find a role for a minimum wage alongside optimal nonlinear labor income taxes if workers' bargaining power is inefficiently low and the government cannot directly control bargaining power.

[4] They do find a potentially useful role for the minimum wage if income taxation is restricted to a linear tax rate. The reason is that a minimum wage can redistribute income in a way that a linear income tax cannot, see also Gerritsen and Jacobs (2013).

into unemployment, but it also pushes some others into high-skilled employment. This implies that a minimum wage can be desirable even if the unemployed receive higher transfers than the low-skilled employed, as long as the high-skilled pay more taxes than the low-skilled. If skill formation is distorted downwards, the minimum wage helps to alleviate the distortions on skill formation. These distortions are absent in previous studies because they assume that the skill distribution is exogenous.

Lee and Saez (2012) is most closely related to our study. Like us, they introduce a minimum wage to the occupational-choice model of Diamond (1980) and Saez (2002). They assume that labor rationing is efficient, i.e., that rationing is concentrated on individuals with the lowest utility surplus of low-skilled work. A binding minimum wage allows the government to reduce taxes on the low-skilled without creating any distortions in participation or skill formation. Intuitively, efficient rationing ensures that neither the unemployed nor the high-skilled employed would be able to find a low-skilled job. In their Proposition 2, Lee and Saez (2012) find that a minimum wage is desirable if low-skilled workers have a marginal social welfare weight that is larger than the average, which equals 1 in the optimum. Their Proposition 3 replicates Marceau and Boadway (1994). If skills are exogenous, a minimum wage can only be desirable if the unemployed receive smaller transfers than the low-skilled employed. We show that both results can be seen as special cases of our more general desirability condition. We make five contributions to the analysis in Lee and Saez (2012).

First, our analysis demonstrates that a minimum-wage increase might be desirable because it alleviates the tax distortions on skill formation. While Lee and Saez (2012) do allow for endogenous skill formation in their Proposition 2, the critical importance of skill distortions is not made clear because their desirability condition is written in terms of marginal social welfare weights, rather than distortions. Second, we derive a desirability condition for the minimum wage that is valid under any arbitrary rationing schedule, whereas the desirability condition of Lee and Saez (2012) is only valid under the assumption of efficient rationing. This is important because it is theoretically and empirically unclear why a minimum wage would only ration workers with the lowest willingness to work (Luttmer, 2007).[5] We show that the rationing schedule is of critical importance for the desirability of a minimum wage. Third, our desirability condition for a minimum-wage increase is valid for any initial allocation. In contrast, the condition in Lee and Saez (2012) is only informative of the desirability of raising the minimum wage if taxes are optimally set. Fourth, our desirability condition is written in terms of empirically measurable statistics, whereas the desirability condition of Lee and Saez (2012) is written in terms of marginal social welfare weights – which are not objectively

[5] Lott (1990) and Palda (2000) also draw attention to inefficient labor rationing caused by minimum wages. See Gerritsen (2016) for a discussion of the consequences of inefficient rationing for optimal taxes and transfers.

measurable. Fifth, unlike Lee and Saez (2012), we bring our desirability condition to the data.

Finally, a number of further studies show that a minimum wage could be desirable if combined with specific other policies, or if low-skilled workers are heterogeneous in multiple dimensions. Boadway and Cuff (2001) consider the framework of Mirrlees (1971) and find that a minimum wage is desirable if it can be combined with a policy that forces the unemployed to accept any job that they can find. Danziger and Danziger (2015) find a useful role for the minimum wage if it can be combined with a policy that forces firms to hire a certain number of low-skilled workers, even if their marginal productivity is below the minimum wage. Blumkin and Danziger (2014) consider a case in which the government redistributes from 'lazy' to 'hard-working' low-skilled workers that earn the same wage rate, but vary in the number of hours they work. They find that a minimum wage is desirable if it reduces labor hours of the lazy, as this makes it harder for lazy individuals to mimic hard-working individuals.

3 Model

This section describes labor-supply decisions of individuals, labor-demand decisions of firms, as well as the objective of the government. We develop a variation of the occupational-choice models of Diamond (1980) and Saez (2002), extended with a binding minimum wage.

3.1 Individuals

We consider a continuum of individuals of mass one. The baseline model assumes that individuals differ in their ability θ and their occupation $i \in \{H, L, U\}$, which denotes whether individuals are high-skilled (H), low-skilled (L), or unemployed (U). Ability is continuously distributed on support $[0, \overline{\theta}]$ according to a cumulative distribution function $G(\theta)$ with a corresponding density function $g(\theta)$. Based on their ability θ, individuals decide to participate as a high-skilled worker and earn wage income w^H, to participate as a low-skilled worker and earn wage income w^L, where $w^H > w^L$, or not to participate at all and earn no wage income: $w^U \equiv 0$. The government can impose a minimum wage by fixing the low-skilled wage w^L. Moreover, it levies differentiated income taxes τ^i for all observed levels of wage income w^i. If taxes for the unemployed are negative, they receive an unemployment benefit $-\tau^U$. We assume that the government cannot distinguish between the voluntary unemployed and the involuntary unemployed so they both receive the same unemployment benefits. Individuals spend all their net income on consumption: $c^i = w^i - \tau^i$.

Utility from consumption is given by an increasing and strictly concave function of consumption c^i, which is identical for all individuals: $v(c^i)$, $v'(\cdot) > 0$, $v''(\cdot) < 0$. When a worker of ability θ becomes low-skilled, she suffers disutility of work $1/\theta$. When she becomes high-skilled, her disutility equals $(1 + \beta)/\theta$. The parameter β is a constant disutility markup of being high-skilled, which represents the effort costs of becoming high-skilled. Disutility of work is decreasing in ability θ, and more so for high-skilled work than for low-skilled work. As a result, individuals with higher ability have a comparative advantage in high-skilled work. Non-participants do not incur any disutility of work. In Appendix D, we generalize our model to allow for two-dimensional heterogeneity, where individuals incur disutility of low-skilled work $1/\theta^L$ and disutility of high-skilled work $1/\theta^H$, with θ^L and θ^H drawn from a joint distribution $G(\theta^L, \theta^H)$.[6] All our main results carry over to the more general setting. We come back to this when discussing the robustness of our results.

Utility of the high skilled, the low skilled, and the unemployed are thus given by:

$$V_\theta^H \equiv v(w^H - \tau^H) - \frac{1+\beta}{\theta}, \tag{1}$$

$$V_\theta^L \equiv v(w^L - \tau^L) - \frac{1}{\theta}, \tag{2}$$

$$V^U \equiv v(-\tau^U), \tag{3}$$

where subscripts indicate that both high- and low-skilled utility depend on individual ability θ. Each individual optimally decides whether to participate in the labor market and whether to work as a low-skilled or a high-skilled worker. Figure 1 provides a stylized graph of utility as a function of ability. $\Theta_1 \equiv \{\theta : V^U = V_\theta^L\}$ is the ability level at which an individual is indifferent between unemployment and low-skilled employment. $\Theta_2 \equiv \{\theta : V^U = V_\theta^H\}$ is the ability level at which an individual is indifferent between unemployment and high-skilled employment. $\Theta_3 \equiv \{\theta : V_\theta^L = V_\theta^H\}$ is the ability level at which an individual is indifferent between low-skilled and high-skilled employment. Eqs. (1)–(3) imply that Θ_1, Θ_2 and Θ_3 are uniquely determined and equal to:

$$\Theta_1 = \frac{1}{v(w^L - \tau^L) - v(-\tau^U)}, \tag{4}$$

$$\Theta_2 = \frac{1+\beta}{v(w^H - \tau^H) - v(-\tau^U)}, \tag{5}$$

$$\Theta_3 = \frac{\beta}{v(w^H - \tau^H) - v(w^L - \tau^L)}. \tag{6}$$

Higher unemployment benefits $(-\tau^U)$ make non-participation more attractive and there-

[6]Our baseline model corresponds to the occupational-choice model of Saez (2002) with a low-skilled participation margin and a skill margin, but without a high-skilled participation margin. The model in Appendix D corresponds to the general case of Saez (2002) with both high- and low-skilled participation margins and a skill margin.

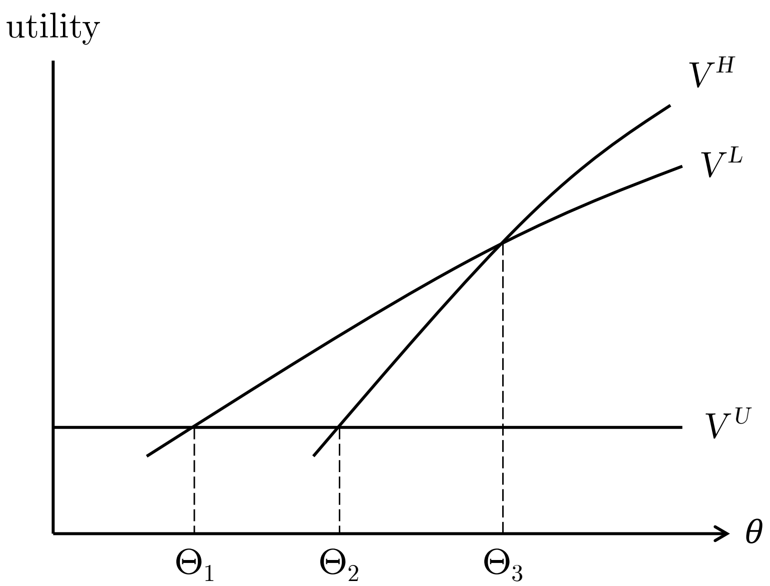

Figure 1: Utility as a function of ability

fore raise Θ_1 and Θ_2. Similarly, an increase in low-skilled net income $(w^L - \tau^L)$ makes low-skilled employment more attractive and therefore lowers Θ_1 and raises Θ_3. Finally, an increase in high-skilled net income $(w^H - \tau^H)$ makes high-skilled employment more attractive, and therefore lowers both Θ_2 and Θ_3. We focus on nontrivial equilibria that contain at least some high- and low-skilled workers, which requires that $\bar{\theta} > \Theta_3 > \Theta_2 > \Theta_1$.[7] In terms of Figure 1, this implies that high- and low-skilled utility curves cross somewhere above the line for unemployed utility.

Labor-supply decisions of individuals are determined by their preference orderings over the different occupations. Figure 1 clearly illustrates how these preferences depend on ability θ. Individuals with low ability $\theta \in [0, \Theta_1)$ prefer non-participation over anything else. Individuals with high ability $(\Theta_3, \overline{\theta}]$ prefer high-skilled employment over anything else. And individuals with intermediate ability $\theta \in [\Theta_1, \Theta_3]$ prefer low-skilled employment over anything else. However, with a binding minimum wage, not every individual with ability $\theta \in [\Theta_1, \Theta_3]$ is able to find a low-skilled job. This is because a minimum wage w^L makes low-skilled employment more attractive, while – as we show below – reducing low-skilled labor demand. This results in low-skilled labor rationing.

In our baseline model, we assume that individuals know whether they are able to find a low-skilled job before making their labor-supply decisions. When individuals figure out that they cannot find a low-skilled job, they must decide between (involuntary) unemployment and high-skilled employment. As illustrated in Figure 1, rationed individuals

[7] If this condition is violated, either no individual wants to be high skilled or no individual wants to be low-skilled. In particular, if $\Theta_3 > \bar{\theta}$, every individual prefers to be either low skilled or unemployed. Moreover, if $\Theta_3 < \Theta_1$, every individual prefers to be either high skilled or unemployed. Finally, if $\bar{\theta} > \Theta_3 > \Theta_1$, then eqs. (4)–(6) imply that $\Theta_3 > \Theta_2 > \Theta_1$.

with ability $\theta \in [\Theta_1, \Theta_2)$ prefer unemployment over high-skilled employment and thus become unemployed. Rationed individuals with ability $\theta \in [\Theta_2, \Theta_3]$ prefer high-skilled employment over unemployment and become high-skilled employed.

Alternatively, we could assume that individuals make their labor-supply decisions before knowing whether they are able to find a low-skilled job. Individual labor supply then depends on the *expected* utility of low-skilled work, which itself depends on her probability of finding a job. We show in Appendix C that our main results are not affected by this alternative sequencing of rationing and labor supply decisions.

3.2 Rationing and aggregate labor supply

Aggregate supply of high- and low-skilled labor depends on which workers are not able to find a low-skilled job. The proportion of individuals with ability θ that is not able to find a job is denoted by the rationing rate u_θ.[8] The rationing schedule is the set of all rationing rates.

Definition 1 *The rationing schedule $\{u_\theta\}$ assigns a rationing rate u_θ to every ability level $\theta \in [\Theta_1, \Theta_3]$, which specifies the proportion of individuals with ability θ that is not able to find a low-skilled job.*

By imposing no structure on the rationing schedule, we remain agnostic about which workers are unable to find a low-skilled job due to a binding minimum wage. We later demonstrate that the rationing schedule critically affects the desirability of a minimum wage. The literature on minimum wages generally makes specific assumptions regarding labor rationing. Studies most often assume efficient rationing, which implies that labor rationing only affects individuals that have the lowest utility surplus of work (e.g., Marceau and Boadway, 1994; Lee and Saez, 2012; Blumkin and Danziger, 2014). In our model, this would imply that only individuals with ability close to either Θ_1 or Θ_3 are unable to find a low-skilled job, since they are indifferent between low-skilled employment on the one hand, and involuntary unemployment or high-skilled employment on the other. However, it is *a priori* unclear whether labor rationing is efficient, since there is generally no secondary market for jobs that could (re)allocate jobs to individuals with the highest utility surplus of work. Uniform or random rationing is often analyzed as an alternative to efficient rationing (e.g., Lee and Saez, 2008; Gerritsen and Jacobs, 2013). Uniform rationing implies that the rationing rate is independent from ability θ, hence $u_\theta = u$ for all individuals with $\theta \in [\Theta_1, \Theta_3]$. However, there is also little reason to expect that rationing is uniformly distributed among individuals with different ability.

[8]The rationing rate does not necessarily correspond to the standard definition of the unemployment rate, because rationed individuals with ability $\theta \in [\Theta_2, \Theta_3]$ decide to become high-skilled rather than unemployed.

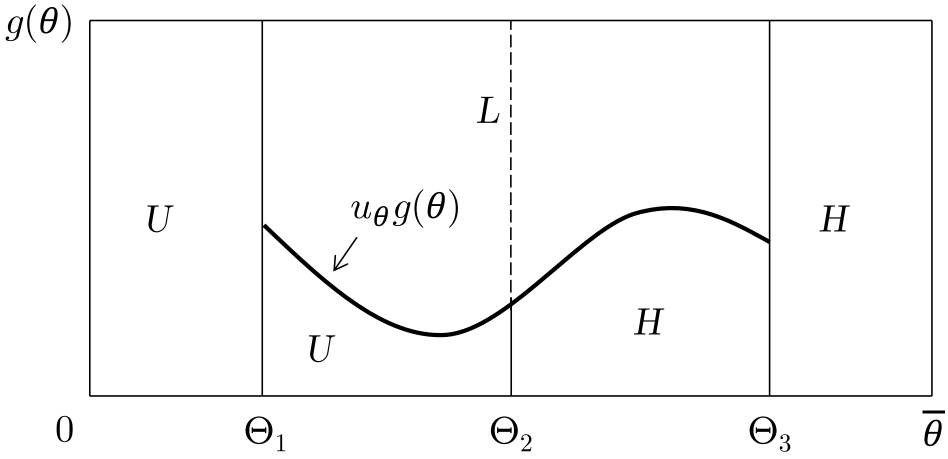

Figure 2: A stylized representation of equilibrium

Aggregating the labor supply of all individuals, while using the definition of the rationing schedule, yields the following expressions for aggregate high- and low-skilled labor supply, and unemployment:

$$H = 1 - G(\Theta_3) + \int_{\Theta_2}^{\Theta_3} u_\theta \mathrm{d}G(\theta), \tag{7}$$

$$L = \int_{\Theta_1}^{\Theta_3} (1 - u_\theta) \mathrm{d}G(\theta), \tag{8}$$

$$U = G(\Theta_1) + \int_{\Theta_1}^{\Theta_2} u_\theta \mathrm{d}G(\theta). \tag{9}$$

Figure 2 provides a graphical illustration of aggregate labor supply. For illustrative purposes, we show the case of a uniform ability distribution, so that $g(\theta)$ is constant. The areas denoted by "U" represent voluntary and involuntary unemployment, the area denoted by "L" represents low-skilled employment, and the areas denoted by "H" represent high-skilled employment. Notice how the rationing schedule determines which individuals become unemployed or high skilled due to rationing.

3.3 Firms

Aggregate demand for high-skilled workers is denoted by H^d and aggregate demand for low-skilled workers by L^d. A competitive, representative firm takes wages as given and demands high- and low-skilled labor to maximize profits. Workers of the same skill type – but different ability θ – are perfect substitutes in production, whereas high-skilled labor and low-skilled labor are imperfect substitutes in production. The production technology $F(\cdot)$ is homogeneous of degree one, and given by:

$$F(H^d, L^d), \qquad F_H, F_L > 0, \qquad F_{HH}, F_{LL} < 0, \qquad F_{HL} > 0, \tag{10}$$

where subscripts denote partial derivatives. Production features positive but diminishing marginal products of both factors. High- and low-skilled labor are cooperant factors of production. Necessary (and sufficient) conditions for profit maximization imply that marginal labor products equal wages:

$$F_H(H^d, L^d) = w^H, \tag{11}$$

$$F_L(H^d, L^d) = w^L. \tag{12}$$

Due to constant returns to scale in production, there are no pure profits in equilibrium. Eqs. (11)–(12) imply downward sloping labor demand curves. As a result, a higher minimum wage is associated with reduced low-skilled labor demand.

3.4 Equilibrium

For a given government policy $\{w^L, \tau^U, \tau^L, \tau^H\}$, eqs. (7)–(9) describe aggregate high- and low-skilled labor supply as functions of the high-skilled wage and the rationing schedule. Eqs. (11)–(12) describe aggregate labor demand as a function of the high-skilled wage. The economy is in general equilibrium when the high-skilled wage and the rationing schedule are such that aggregate labor supply equals labor demand: $H = H^d$ and $L = L^d$. Because the minimum wage is only binding for low-skilled labor, the high-skilled wage freely adjusts to ensure that the high-skilled labor market clears in equilibrium. However, a binding minimum wage makes low-skilled wage adjustments impossible. Instead, labor rationing adjusts to ensure equilibrium on the low-skilled labor market. As can be seen from Figure 2, there is an infinite number of possible rationing schedules $\{u_\theta\}$ that yield the same equilibrium levels of aggregate employment. This indeterminacy has important implications for the comparative statics and welfare effects of the minimum wage, as we show in the following sections.

3.5 Government

The government sets a minimum wage w^L and income taxes $\{\tau^H, \tau^L, \tau^U\}$. The critical informational assumption of our analysis is that individual earnings are verifiable. The government can thus simultaneously implement an income tax and enforce a binding minimum wage for the low-skilled. Our approach is informationally consistent because the implementation of both the income tax and the minimum wage require the same information on individual earnings.[9]

[9]This contrasts with a number of previous studies that are 'informationally inconsistent' (e.g., Guesnerie and Roberts, 1987; Allen, 1987; Marceau and Boadway, 1994; Boadway and Cuff, 2001; Blumkin and Danziger, 2014; Danziger and Danziger, 2015). These studies assume that information on individual wages can be used to enforce a minimum wage, but not to condition taxes and transfers on wages, since this would allow the government to reach first best. In our case, as in Lee and Saez (2012), first best

We assume that the social welfare function $\mathcal{W}$ is utilitarian:[10]

$$\mathcal{W} \equiv UV^U + \int_{\Theta_1}^{\Theta_2} (1-u_\theta) V_\theta^L \mathrm{d}G(\theta) + \int_{\Theta_2}^{\Theta_3} u_\theta V_\theta^H \mathrm{d}G(\theta) + \int_{\Theta_3}^{\overline{\theta}} V_\theta^H \mathrm{d}G(\theta). \tag{13}$$

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

$$\mathcal{B} \equiv U\tau^U + L\tau^L + H\tau^H - R = 0, \tag{14}$$

where R is an exogenous revenue requirement.

4 A net-income-neutral minimum-wage increase

4.1 Defining the net-income-neutral minimum-wage increase

A minimum wage can raise the income of low-skilled workers, but so can the tax system.[11] The relevant question therefore is: how does a change in the minimum wage differ from a distributionally equivalent change in the income tax system? This question can be approached in two different ways. One could derive the effects of a change in the minimum wage in isolation, as well as the effects of a distributionally equivalent tax change, and then compare the two. Or one could derive the effects of a combined reform that simultaneously raises the minimum wage and adjusts taxes to leave net wages unaffected. Both approaches are logically equivalent and therefore yield identical results. We take the latter approach and label the combined reform a *net-income-neutral* (NIN) increase in the minimum wage.

Analyzing a NIN increase in the minimum wage has several important advantages. First, the NIN minimum-wage increase allows us to ignore many behavioral effects that a minimum-wage increase has in common with a distributionally equivalent change in taxes. This greatly reduces the analytical complexity associated with deriving the comparative statics.[12] Second, because the reform leaves net incomes unaffected, we can focus the welfare analysis exclusively on the efficiency gains and losses of a minimum-wage increase relative to a distributionally equivalent change in taxes. This allows us to analyze the

cannot be reached even with wage-specific taxes because individuals are heterogeneous with respect to disutility of work instead of wages.

[10]None of our findings depend on the assumption of a utilitarian social welfare function. Stronger redistributional concerns can be introduced by, for example, weighing individual utilities with Pareto weights or by summing over a concave transformation of individual utilities.

[11]For recent empirical evidence on the distributional effects of the U.S. minimum wage, see Autor, Manning, and Smith (2016). For the distributional effects of the U.S. tax system, see Piketty and Saez (2007).

[12]As we demonstrate in Gerritsen and Jacobs (2013), the comparative statics and welfare effects of changes in the minimum wage in isolation from taxes, or taxes in isolation from the minimum wage, are straightforward but mathematically tedious and notationally elaborate.

desirability of a minimum wage without taking a stance on inherently subjective political preferences for income redistribution.[13] Third, the welfare analysis of the reform gives a necessary condition for the relative desirability of a minimum wage that is expressed solely in terms of *sufficient statistics* that could be measured empirically. Later we bring this condition to the data.

The following Lemma formally defines the NIN minimum-wage increase and derives the changes in taxes that are necessary to maintain net-income neutrality.

Lemma 1 *A net-income-neutral increase in the minimum wage raises the minimum wage by* $\mathrm{d}w^L > 0$*, keeps the unemployment benefit constant* $\mathrm{d}\tau^U = 0$*, raises the low-skilled tax by* $\mathrm{d}\tau^L = \mathrm{d}w^L$*, and lowers the high-skilled tax such that* $\mathrm{d}\tau^H = \mathrm{d}w^H = -(L/H)\mathrm{d}w^L < 0$.

Proof. Equate the total derivative of net income $w^i - \tau^i$ to zero to find $\mathrm{d}\tau^i = \mathrm{d}w^i$ for $i \in \{H, L, U\}$. Linear homogeneity of the production function together with eqs. (11)–(12) implies zero equilibrium profits: $F(H, L) - w^H H - w^L L = 0$. Take the total derivative and rearrange to find $H\mathrm{d}w^H = -L\mathrm{d}w^L$. Rewrite to obtain the Lemma. ■

A minimum wage compresses the wage differential between high- and low-skilled workers due to complementarity of labor types in production (i.e., $F_{HL} > 0$). Intuitively, an increase in the low-skilled wage drives down low-skilled labor demand, which, in turn, lowers the productivity and wages of high-skilled workers. The increase in the low-skilled wage is fully paid for by a decrease in high-skilled wages ($H\mathrm{d}w^H = -L\mathrm{d}w^L$). This logically follows from the absence of profits due to constant returns to scale in production. To fully neutralize the changes in gross wages, the NIN minimum-wage reform therefore raises low-skilled taxes and lowers high-skilled taxes, while keeping unemployment benefits constant.

4.2 Comparative statics

The behavioral effects of the NIN minimum-wage increase are critical for the welfare analysis we conduct below. We graphically illustrate the effects of the NIN minimum-wage increase in Figure 3. The policy reform has no effect on individual preferences for different occupations. Since net wages do not change, eqs. (4)–(6) imply that the cut-offs Θ_1, Θ_2 and Θ_3 remain unaffected. The only effect of the policy reform is that individuals may change their occupation due to a change in low-skilled labor rationing. The NIN minimum-wage increase raises low-skilled labor costs, and, therefore, results in a reduction of low-skilled employment ($\mathrm{d}L < 0$). This is indicated by the upward

[13] Our approach is comparable to Christiansen (1981, 1984) and Kaplow (2008), among others. They study combined reforms that raise a consumption tax (or public good provision, the tax on capital income, etc.) while offsetting all distributional implications by appropriate changes in the non-linear income tax. Like us, they obtain simple desirability conditions that do not depend on social preferences for income redistribution.

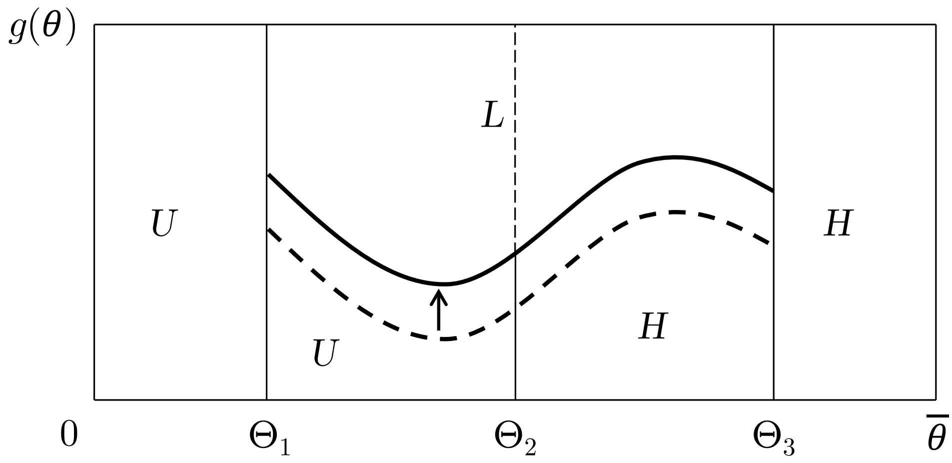

Figure 3: Comparative statics of a net-income-neutral minimum wage increase

shift of the rationing schedule in Figure 3. As a result, some individuals with ability $\theta \in [\Theta_1, \Theta_2)$ might lose their low-skilled jobs and move into unemployment ($\mathrm{d}U \geq 0$). These rationed workers prefer unemployment over high-skilled employment. Similarly, some individuals with ability $\theta \in [\Theta_2, \Theta_3]$ might also lose their low-skilled job, and move into high-skilled employment ($\mathrm{d}H \geq 0$). These workers prefer high-skilled employment over unemployment.

What happens to unemployment and high-skilled employment crucially depends on how the increase in rationing is distributed among low-skilled workers. That is, it depends on the change in the rationing schedule. In what follows, we use ρ as a measure of the proportion of additional rationing that is concentrated on individuals that prefer high-skilled employment over unemployment:

$$\rho \equiv \frac{\int_{\Theta_2}^{\Theta_3} \mathrm{d}u_\theta \mathrm{d}G(\theta)}{\int_{\Theta_1}^{\Theta_3} \mathrm{d}u_\theta \mathrm{d}G(\theta)} = \frac{\mathrm{d}H}{\mathrm{d}L} \in [0, 1]. \tag{15}$$

In terms of Figure 3, ρ measures the increase in the area denoted by "H" relative to the decrease in the area denoted by "L." The larger is ρ, the more a given increase in labor rationing translates into an increase in high-skilled employment. We assume that $\mathrm{d}u_\theta \geq 0$ for all θ, which ensures that $\rho \in [0, 1]$. At one extreme, if $\rho = 1$, all additional rationing is concentrated on low-skilled workers with relatively high ability. In that case, rationing leads to more high-skilled employment without causing any increase in unemployment. At the other extreme, if $\rho = 0$, all additional rationing is concentrated on low-skilled workers with relatively low ability. In that case, rationing does not affect high-skilled employment but only raises unemployment. Armed with the definition of ρ, we can formally state the comparative statics of a NIN minimum-wage increase in the following Lemma.

Lemma 2 *The general-equilibrium comparative statics of the NIN minimum-wage increase, as described by Lemma 1, are:*

$$\mathrm{d}V^U = \mathrm{d}V_\theta^L = \mathrm{d}V_\theta^H = \mathrm{d}\Theta_1 = \mathrm{d}\Theta_2 = \mathrm{d}\Theta_3 = 0. \tag{16}$$

$$\mathrm{d}H = \int_{\Theta_2}^{\Theta_3} \mathrm{d}u_\theta \mathrm{d}G(\theta) = \rho\alpha\varepsilon\frac{\mathrm{d}w^L}{w^L} \geq 0, \tag{17}$$

$$\mathrm{d}L = -\int_{\Theta_1}^{\Theta_3} \mathrm{d}u_\theta \mathrm{d}G(\theta) = -\alpha\varepsilon\frac{\mathrm{d}w^L}{w^L} < 0, \tag{18}$$

$$\mathrm{d}U = \int_{\Theta_1}^{\Theta_2} \mathrm{d}u_\theta \mathrm{d}G(\theta) = (1-\rho)\alpha\varepsilon\frac{\mathrm{d}w^L}{w^L} \geq 0, \tag{19}$$

where $\varepsilon \equiv -F_L/(LF_{LL}) > 0$ *is the labor demand elasticity and* $\alpha \equiv (1/L + \rho/H)^{-1} > 0$ *is a share parameter.*

Proof. See Appendix A. ■

Relative to a distributionally equivalent change in taxes, a minimum-wage increase leads to a reduction in low-skilled employment ($\mathrm{d}L < 0$). This (weakly) increases both unemployment ($\mathrm{d}U \geq 0$) and high-skilled employment ($\mathrm{d}H \geq 0$). The reduction in low-skilled employment is equal to the total increase in rationing due to a NIN minimum-wage increase, as shown by eq. (18). Unsurprisingly, the higher the labor-demand elasticity ε, the larger the increase in low-skilled labor rationing. A fraction ρ of the additional labor rationing reflects an increase in high-skilled employment, i.e., $\mathrm{d}H = -\rho\mathrm{d}L \geq 0$ as shown in eq. (17). As long as $\rho > 0$, the NIN reform leads to more high-skilled employment. Similarly, a fraction $1-\rho$ of the additional rationing reflects higher unemployment, i.e., $\mathrm{d}U = -(1-\rho)\mathrm{d}L \geq 0$ as shown in eq. (19). As long as $\rho < 1$, the NIN reform leads to more unemployment.

5 Welfare analysis

5.1 A desirability condition for the minimum wage

The following Proposition is the main result of the paper. It provides the condition under which a minimum-wage increase is more desirable than a distributionally equivalent change in the tax system.

Proposition 1 *A minimum-wage increase is more desirable than a distributionally equivalent change in the tax system if and only if a NIN minimum-wage increase raises social* *welfare, so that the following condition is satisfied::*

$$\rho(\tau^H - \tau^L) - (1-\rho)(\tau^L - \tau^U) > (1-\rho)\left(\frac{\bar{V}_{12}^L - V^U}{\lambda}\right) + \rho\left(\frac{\bar{V}_{23}^L - \bar{V}_{23}^H}{\lambda}\right), \tag{20}$$

where λ is the shadow value of public resources, $\bar{V}_{12}^L \equiv \int_{\Theta_1}^{\Theta_2} V_\theta^L \mathrm{d}u_\theta \mathrm{d}G(\theta) / \int_{\Theta_1}^{\Theta_2} \mathrm{d}u_\theta \mathrm{d}G(\theta)$ and $\bar{V}_{23}^L \equiv \int_{\Theta_2}^{\Theta_3} V_\theta^L \mathrm{d}u_\theta \mathrm{d}G(\theta) / \int_{\Theta_2}^{\Theta_3} \mathrm{d}u_\theta \mathrm{d}G(\theta)$ are the average low-skilled utility of the individuals that are rationed by the reform and have ability $\theta \in [\Theta_1, \Theta_2)$ and $\theta \in [\Theta_2, \Theta_3]$ respectively, and $\bar{V}_{23}^H \equiv \int_{\Theta_2}^{\Theta_3} V_\theta^H \mathrm{d}u_\theta \mathrm{d}G(\theta) / \int_{\Theta_2}^{\Theta_3} \mathrm{d}u_\theta \mathrm{d}G(\theta)$ is the average high-skilled utility of the individuals that are rationed by the reform and have ability $\theta \in [\Theta_2, \Theta_3]$.

Proof. See Appendix A. ■

In the previous section, Lemma 2 established that a NIN increase in the minimum wage raises unemployment and high-skilled employment at the cost of reduced low-skilled employment, while leaving utility within any occupation unaffected. In line with this, Proposition 1 establishes that the NIN minimum-wage reform only affects social welfare through changes in utility and public revenue that are caused by the increases in high-skilled employment and unemployment. The left-hand side of eq. (20) captures the welfare effects of the potential public revenue gains ($\mathrm{d}\mathcal{B}$), whereas the right-hand side represents the welfare effects of the potential utility losses ($-\mathrm{d}\mathcal{W}/\lambda$). We can distinguish four welfare-relevant effects of the NIN increase in the minimum wage: (i) it raises public revenue from individuals that become high-skilled if $\tau^H > \tau^L$ (first left-hand-side term), (ii) it reduces public revenue from individuals that become unemployed if $\tau^L > \tau^U$ (second left-hand-side term), (iii) it lowers the utility of individuals that become involuntarily unemployed (first right-hand-side term), and (iv) it lowers the utility of individuals that become involuntarily high-skilled employed (second right-hand-side term).

The left-hand side of eq. (20) indicates that the NIN minimum-wage increase has an ambiguous effect on public revenue ($\mathrm{d}\mathcal{B} \gtrless 0$). On the one hand, the increase in high-skilled employment leads to higher revenue, provided that the high-skilled pay more taxes than the low-skilled (i.e., if $\tau^H > \tau^L$). This increase in revenue is larger if rationing induces more individuals to become high-skilled (i.e., if ρ is larger). On the other hand, the increase in unemployment leads to a reduction in public revenue if the low-skilled pay more taxes than the unemployed (i.e., if $\tau^L > \tau^U$). The reduction in revenue is larger if rationing leads to more unemployment (i.e., if ρ is smaller). Thus, the net effect on revenue crucially depends on the tax wedges ($\tau^H - \tau^L$) and ($\tau^L - \tau^U$), *and* on the fraction of rationing ρ that is concentrated on individuals that prefer high-skilled employment over unemployment.

The right-hand side of eq. (20) demonstrates that the NIN minimum-wage increase potentially generates utility losses ($\mathrm{d}\mathcal{W}/\lambda \leq 0$). Some individuals are rationed out of the low-skilled labor market and decide to become unemployed. Since these individuals prefer low-skilled employment over unemployment, they suffer utility losses that are – expressed in monetary units – on average equal to $(\bar{V}_{12}^L - V^U)/\lambda$. To obtain the total welfare effect, this term is multiplied by the proportion of rationed individuals that become unemployed $(1-\rho)$. Similarly, some of the rationed individuals switch to high-skilled employment. Since these individuals prefer low-skilled employment over high-skilled employment, they also suffer utility losses that are – expressed in monetary units – on average equal to $(\bar{V}_{23}^L - \bar{V}_{23}^H)/\lambda$. This term is multiplied by the proportion of rationed individuals that become high-skilled (ρ). The total utility loss of the NIN minimum-wage increase crucially depends on the efficiency of the rationing schedule. In the case of efficient rationing, all rationing is concentrated on individuals that are indifferent between low-skilled employment and high-skilled employment or unemployment (i.e., on those that have ability Θ_1 or Θ_3). In that case $\bar{V}_{12}^L = V^U$ and $\bar{V}_{23}^L = \bar{V}_{23}^H$ so that $\mathrm{d}\mathcal{W} = 0$ and the right-hand side of eq. (20) vanishes. In any other case, the more inefficient rationing – the more rationing is concentrated around ability Θ_2 – the higher the utility losses of the NIN increase in the minimum wage, and therefore the larger the right-hand side of eq. (20).

The utility losses that occur because of inefficient labor rationing received ample attention in the literature (e.g. Lott, 1990; Marceau and Boadway, 1994; Palda, 2000; Luttmer, 2007; Lee and Saez, 2012; Gerritsen, 2016).[14] But the rationing schedule – as represented by ρ – also crucially determines labor-supply responses, and thereby the magnitude of the revenue effects associated with the minimum wage. That is, the rationing schedule determines whether rationing mostly leads to higher unemployment and revenue losses (if ρ is small) or to increases in high-skilled employment and revenue gains (if ρ is large). To the best of our knowledge, this role of labor rationing has gone largely unnoticed.

In our model, the minimum wage is a second-best instrument to alleviate tax distortions on the skill margin. If the government redistributes from the high-skilled to the low-skilled and from the low-skilled to the unemployed, it does so by setting distortionary taxes on skill formation $(\tau^H > \tau^L)$ and low-skilled labor participation $(\tau^L > \tau^U)$. By raising high-skilled employment and unemployment, the NIN minimum-wage increase alleviates the tax-induced distortion on skill formation and exacerbates the tax-induced distortion on low-skilled participation. Proposition 1 establishes that there is a second-best role for the minimum wage if the revenue gains from a reduction in education distortions are large enough to compensate for the revenue losses of increased participation distortions and the utility losses of inefficient rationing. In a first-best world, redistribution would take place only through individualized lump-sum taxes and transfers, and there would be no distortionary taxation $(\tau^H - \tau^L = \tau^L - \tau^U = 0)$. In that case, Proposition 1 shows that a minimum wage is not desirable as it would merely lead to utility losses

[14] This is in line with the literature that emphasizes the utility losses associated with inefficient rationing of specific commodities, such as rental houses (Glaeser and Luttmer, 2003), gasoline (Frech and Lee, 1987), or residential gas (Davis and Kilian, 2011).

from inefficient rationing.

5.2 A necessary condition

As discussed, the utility losses of a NIN minimum-wage increase – given by the right-hand side of eq. (20) – are weakly positive. For a minimum-wage increase to be more desirable than a distributionally equivalent change in taxes, the left-hand side of eq. (20) must therefore be positive as well. This yields the following *necessary* condition for the desirability of a minimum-wage increase.

Corollary 1 *A necessary condition for a minimum-wage increase to be more desirable than a distributionally equivalent change in the tax system is that a NIN minimum-wage increase raises public revenue (*$\mathrm{d}\mathcal{B} > 0$*):*

$$(\tau^H - \tau^L)\frac{\rho}{1-\rho} > (\tau^L - \tau^U). \tag{21}$$

Or, equivalently:

$$(\tau^H - \tau^L)\frac{\mathrm{d}H}{\mathrm{d}U} > (\tau^L - \tau^U). \tag{22}$$

Proof. Equate the right-hand side of eq. (20) to zero and rewrite the expression to obtain eq. (21). Substitute for eqs. (17) and (19) to obtain eq. (22). ■

Intuitively, a NIN minimum-wage increase can only be desirable if it leads to an increase in public revenue in order to compensate for any reductions in individual utility due to inefficient rationing. The reform is more likely to raise revenue when the tax wedge on skill formation $(\tau^H - \tau^L)$ is higher, the tax wedge on low-skilled participation $(\tau^L - \tau^U)$ is lower, and rationing leads to more high-skilled employment (i.e., when ρ or equivalently $\mathrm{d}H/\mathrm{d}U$ is higher). If the conditions in Corollary 1 do *not* hold, a minimum wage increase is less desirable than an equally redistributive change in taxes. Furthermore, a NIN *decrease* of the minimum wage would in that case constitute a Pareto improvement as it would raise revenue as well as (weakly) increase the utility of every individual.

There are two special cases in which $\mathrm{d}\mathcal{W}/\lambda = 0$, so that eqs. (21)–(22) are both necessary *and* sufficient conditions for a minimum-wage increase to be more desirable than a distributionally equivalent change in taxes. The first special case is when rationing is efficient and there is no pre-existing rationing. In that case, rationed individuals are indifferent between low-skilled employment and unemployment or high-skilled employment, and therefore do not suffer any utility losses. The second special case is when social preferences are Rawlsian rather than utilitarian, so that we can write $\mathcal{W} = V^U$ instead of eq. (13). With Rawlsian preferences, the government simply does not care about the utility losses of rationed low-skilled workers, so that $\mathrm{d}\mathcal{W}/\lambda = 0$.

The condition in eq. (21) illustrates once more the critical importance of the rationing schedule as captured by ρ. The more rationing is concentrated on individuals that prefer high-skilled employment over unemployment, the more likely it is that a minimum-wage increase leads to more public revenue than an equally redistributive tax change. Given that we lack information on the rationing schedule, $\rho/(1-\rho)$ could be anywhere between zero and infinity. Hence, it is *a priori* unclear whether the condition in eq. (21) is satisfied or not. However, as indicated by eq. (22), we do know that $\rho/(1-\rho)$ is equal to the increase in high-skilled employment relative to the increase in unemployment, $\mathrm{d}H/\mathrm{d}U$. This means that, if we have an empirical estimate of $\mathrm{d}H/\mathrm{d}U$, we do not need specific knowledge on the rationing schedule to determine whether the necessary conditions in Corollary 1 are satisfied. In other words, the effect of unemployment on high-skilled employment – for given net wages – can function as a *sufficient statistic* for the rationing schedule (cf. Chetty, 2009). We return to this in Section 7 when we bring eq. (22) to the data.

Corollary 1 allows us to judge the desirability of a minimum-wage increase without taking a stance on either intrinsically subjective political preferences for redistribution, or the empirical magnitude of the aggregate low-skilled employment reduction associated with a minimum-wage increase. To see this, notice that neither social preferences nor the aggregate low-skilled employment effects enter the necessary conditions in eqs. (21)–(22). Intuitively, we can refrain from making political judgements regarding income redistribution because we compare a minimum-wage increase with a distributionally equivalent tax reform. The only relevant question is whether the minimum wage redistributes income more efficiently than income taxes. The answer to this question does not depend on controversial estimates of the aggregate low-skilled employment effects of a minimum wage because the relative costs of a minimum-wage increase (revenue losses of more unemployment and utility losses of rationing) and the relative benefits of a minimum-wage increase (revenue gains of more high-skilled employment) are both proportional to the total reduction in low-skilled employment. It is the factor of proportionality ρ, which determines how much of the low-skilled employment loss reflects an increase in high-skilled employment, that is crucial for the desirability of a minimum-wage increase.

5.3 Relation to the literature

Earlier literature on the desirability of a minimum wage typically considers whether a binding minimum wage could improve upon the optimal tax system. While we consider optimal taxes in the next section, it is worth emphasizing that our results in Proposition 1 and Corollary 1 are valid irrespective of whether the tax system is optimized or not, generalizing earlier literature.

Moreover, we contribute to the literature in a number of other ways. Most analyses consider individual skills as exogenously given, and therefore do not find the same second-best role for a minimum wage as we do. Indeed, the classical studies by Allen (1987) and Guesnerie and Roberts (1987) find that a minimum wage is not a useful instrument for redistribution. In their framework, labor hours of low-skilled workers are optimally distorted downwards by positive marginal taxes. A minimum wage exacerbates this distortion by efficiently rationing the number of low-skilled labor hours. Thus, compared to a distributionally equivalent change in taxes, a minimum-wage increase leads to a reduction in public revenue. Because skills are assumed to be exogenous, the minimum-wage increase does not generate any offsetting revenue gains and is therefore undesirable.

Marceau and Boadway (1994) also consider exogenous skills. However, rationing takes place on the extensive margin, as in our paper, and thus causes involuntary unemployment. They find that a minimum wage can only be desirable if the low-skilled employed pay less taxes than the unemployed, so that an increase in unemployment leads to revenue gains. Lee and Saez (2012) draw the same conclusion in their Proposition 3, which also considers exogenously given skills. We capture these results as special cases. To see this, consider $\rho = 0$, which would imply that the skill distribution is unaffected by rationing. In that case, Proposition 1 shows that a minimum wage can only be desirable if the low-skilled employed pay less taxes than the unemployed ($\tau^L < \tau^U$) – as in Marceau and Boadway (1994) and Proposition 3 in Lee and Saez (2012). This could be the case with sufficiently generous in-work benefits. Intuitively, if tax policy yields an upward distortion of low-skilled participation, a minimum wage alleviates this distortion by rationing individuals out of the low-skilled labor market.[15]

We generalize these findings by allowing for endogenous skill formation. Contrary to earlier studies, we show in Proposition 1 that a minimum wage can be desirable even if participation is taxed. The reason is that the rationing associated with a minimum-wage increase not only raises unemployment, but also encourages skill formation. The latter effect increases revenue as long as high-skilled workers pay more taxes than low-skilled workers. While the effect of rationing on high-skilled employment is relatively understudied within the minimum-wage literature, it is a potentially important channel through which a minimum wage could positively affect social welfare. Indeed, a sizable empirical literature demonstrates the importance of involuntary unemployment for individual decisions to invest in education. We discuss the evidence from this literature in Section 7.

Lee and Saez (2012) do allow for endogenous skill formation in their Proposition 2, while assuming that rationing is efficient and that taxes are set optimally. Moreover, they consider a different policy reform. Instead of a minimum-wage increase, they study an increase in the low-skilled transfer, while keeping low-skilled wages constant with a

[15]See also Hummel and Jacobs (2016), who demonstrate that labor unions can be desirable for income redistribution to alleviate the distortions of excessive labor participation caused by participation subsidies.

minimum wage. A binding minimum wage must be part of the policy optimum if this reform is desirable in the tax optimum without a minimum wage. They find that this is the case if the social welfare weight of the low-skilled exceeds one, or in our notation if $b^L \equiv v'(c^L)/\lambda > 1$. In the next section, we show that the desirability condition in Proposition 1 can only be reduced to $b^L > 1$ if rationing is indeed efficient and taxes are set optimally. Thus, we generalize Proposition 2 of Lee and Saez (2012) by deriving a desirability condition for a minimum-wage increase that is valid irrespective of whether rationing is efficient or not, and irrespective of whether taxes are optimized or not. Furthermore, our necessary condition in Corollary 1 does not depend on social welfare weights, but only consists of variables with empirical counterparts. This allows us to bring this condition to the data in Section 7.

5.4 Robustness

In deriving our results, we made a number of assumptions that warrant further discussion.

Sequencing of rationing and participation decisions We assumed that individuals make their labor-supply decisions after they find out whether they are rationed out of the low-skilled labor market. Consequently, they can still move into high-skilled employment if they are unable to find a low-skilled job. Alternatively, we could assume that individuals make their participation decisions before knowing whether they will be rationed out of the low-skilled labor market. In that case, individuals take into account that there is a positive probability that they lose their low-skilled job. In Appendix C, we show that Proposition 1 and Corollary 1 carry over in modified form to this different sequencing of labor-supply decisions and rationing realizations. Our results remain qualitatively unaffected when we adopt this alternative sequencing of rationing realizations and labor-supply decisions. In particular, compared to a change in income taxes, a minimum-wage increase causes additional rationing. This raises the probability of low-skilled unemployment and thereby induces some individuals to switch to high-skilled employment. Welfare effects consist of utility losses from rationing and public revenue gains and losses from increased high-skilled employment and unemployment. As a result, a minimum-wage increase is desirable if and only if the revenue gains from additional high-skilled employment are large enough to offset any revenue and utility losses from higher unemployment.

Multidimensional ability distribution We restricted our model by assuming that all differences between individuals can be captured by ability θ. A more general setup allows individuals to differ in low-skilled ability θ^L and high-skilled ability θ^H, both drawn from some joint distribution $G(\theta^L, \theta^H)$. Disutility of low-skilled work would equal $1/\theta^L$ and disutility of high-skilled work would equal $1/\theta^H$.[16] Labor-supply decisions then depend on both θ^L and θ^H. This more general setup allows for a high-skilled participation margin, i.e., it allows for individuals that are indifferent between unemployment and high-skilled employment. This is ruled out in the case of one-dimensional heterogeneity. In Appendix D, we extend our model to allow for two-dimensional heterogeneity. We demonstrate that Proposition 1 and Corollary 1 perfectly carry over to this more general setting.

Number of skill types We could extend the model by allowing for more than two skill types in production as well as for multiple other factors of production. In that case, only the lowest skill type would be subject to a binding minimum wage. Our results would remain unaffected, provided that the government is able to tax every factor of production separately, so that the redistributive effects of a minimum wage can be perfectly mimicked by a distributionally equivalent tax reform. The same argument is made by Lee and Saez (2012).

Intensive margin We could allow individuals to decide on both their skill type and the number of hours they work. This would generate income inequality among individuals with the same skill type if their ability also affects disutility of working hours. Given that the government could enforce a binding minimum wage, informational consistency requires that the government sets skill-specific income taxes.[17] We show in Gerritsen and Jacobs (2014) that this would leave our results on the desirability of a NIN minimum-wage increase unaffected.

6 Minimum-wage desirability in the tax optimum

There are two intuitively appealing allocations at which to evaluate the desirability of a NIN minimum-wage increase. One is the allocation as observed in real-world economies. This would inform us about the desirability of raising the minimum wage in these economies. A second possible allocation is the tax optimum in the absence of a binding minimum wage. From this we might learn whether a binding minimum wage is part of the overall policy optimum. This would also tightly link our results to the earlier literature, which typically restricts attention to the optimal tax system. We focus on the optimal tax system in this section, and consider real-world economies in the next.

[16] Notice that the model in the main text is a special case of this more general setup, in which $\theta^H = \theta^L/(1+\beta)$.

[17] As is well known from Diamond (1980), skill-specific income taxes are no longer identical to skill-independent income taxes when labor is also supplied on the intensive margin. The informationally consistent tax system might therefore be at odds with the tax system we typically observe in reality.

6.1 General rationing

Deriving the optimal tax system in the absence of a binding minimum wage, and evaluating the desirability condition of the minimum wage at this allocation, yields the following Proposition.

Proposition 2 *In the absence of a binding minimum wage, the optimal tax system features taxes that are increasing with income: $\tau^H > \tau^L > \tau^U$. As a result, a minimum-wage increase is more desirable than a distributionally equivalent tax change if and only if rationing is sufficiently efficient and ρ sufficiently high.*

Proof. We prove the first part of the proposition in Appendix A. The second part immediately follows from imposing $\tau^H > \tau^L > \tau^U$ on the desirability condition in Proposition 1. ■

The first part of Proposition 2 is intuitively straightforward. The government optimally redistributes from the employed to the unemployed because the marginal utility of consumption of the employed is smaller than that of the unemployed. It therefore sets $\tau^L > \tau^U$. Similarly, the government optimally redistributes from the high-skilled to the low-skilled because the marginal utility of consumption of the high-skilled is smaller than that of the low-skilled. It therefore sets $\tau^H > \tau^L$. Given that optimal taxes are increasing with income, increases in high-skilled employment lead to revenue gains and increases in unemployment to revenue losses. From Proposition 1, it then follows that a binding minimum wage is potentially part of the overall optimum. Starting from the overall tax optimum, it is optimal to introduce a minimum wage if ρ is sufficiently high – so that a NIN minimum-wage increase leads to revenue gains – and rationing is sufficiently efficient – so that the utility losses are limited.

Recall from the previous section that both Proposition 1 and Corollary 1 remain valid in a more general framework with multidimensional heterogeneity and a high-skilled participation margin, as shown in Appendix D. However, the first part of Proposition 2 need not hold in such a more general model. As we know from Diamond (1980), Saez (2002), and Christiansen (2015), the optimal low-skilled tax might in that case be either greater or smaller than the unemployment tax. Intuitively, the government might want to redistribute from the unemployed to the low-skilled employed in order to alleviate distortions on the high-skilled participation margin. If $\tau^H > \tau^U > \tau^L$ in the tax optimum, a minimum wage would always generate more revenue than a distributionally equivalent tax change. Intuitively, as low-skilled participation is distorted *upwards* and skill formation is distorted downwards, rationing alleviates both distortions by raising unemployment and high-skilled employment. As a result, the necessary condition of Corollary 1 would in that case be fulfilled. However, Proposition 1 still requires rationing to be sufficiently efficient for a minimum-wage increase to be more desirable than a distributionally equivalent change in taxation.

6.2 Efficient rationing

Can we say more about the desirability of a minimum-wage increase if we assume that taxes are set optimally *and* that rationing is efficient? As it turns out, only under these two assumptions does our desirability condition collapse to the one formulated by Lee and Saez (2012). This is established by the following Proposition.

Proposition 3 *In the absence of a binding minimum wage, if taxes are set optimally and if rationing is efficient, a minimum-wage increase is more desirable than a distributionally equivalent tax change if and only if the following condition holds:*

$$(\tau^H - \tau^L)\frac{\rho}{1-\rho} > (\tau^L - \tau^U) \qquad \Leftrightarrow \qquad b^L \equiv \frac{v'(c^L)}{\lambda} > 1 \tag{23}$$

Proof. See Appendix A. ■

Two things are worth mentioning to obtain an intuitive understanding of the results in Proposition 3. First, notice that $b^L > 1$ implies that the marginal distributional benefits of the low-skilled tax are negative. In the optimum, this must mean that the marginal distortion of the low-skilled tax is negative as well. In other words, $b^L > 1$ implies that the distortion on skill formation is larger than the distortion on participation. Consequently, the behavioral responses associated with an increase in the low-skilled tax must generate a net gain in public revenue. Second, notice that, under efficient rationing, only those individuals with the lowest utility surplus of low-skilled work are rationed out of the labor market. These are the same individuals that would alter their labor supply decisions in response to an increase in the low-skilled tax. Efficient rationing therefore causes the same behavioral responses as an increase in the low-skilled tax. Taken together, these two things imply that, if taxes are optimal, $b^L > 1$, and rationing is efficient, a NIN minimum-wage increase raises public revenue and leaves utility unaffected. A minimum-wage increase must therefore be more desirable than a distributionally equivalent change in taxes.

There are three reasons to focus on our desirability conditions in eqs. (20)–(22) rather than the one of Lee and Saez (2012) in eq. (23). The first is that eqs. (20)–(22) are valid under any specification of the rationing schedule, whereas eq. (23) is only valid under efficient rationing. The second is that eqs. (20)–(22) are valid for any tax-transfer system, including non-optimal ones, whereas eq. (23) is only valid under the optimal tax system. Third, most statistics in eqs. (20)–(21) and all statistics in eq. (22) have empirical counterparts, whereas the statistic in eq. (23) depends on one's political judgments regarding the social marginal value of low-skilled income.

7 Bringing the desirability condition to the data

7.1 A sufficient-statistics approach

Corollary 1 provides necessary conditions for a minimum-wage increase to be more desirable than a distributionally equivalent tax change. Eq. (21) indicates that this condition crucially depends on the rationing schedule (ρ), about which we know little. Fortunately, eq. (22) indicates that we could use an empirical measure of $\mathrm{d}H/\mathrm{d}U$ as a *sufficient statistic* for the rationing schedule. In this section, we use estimates from the empirical literature on education and unemployment to measure this statistic. Along with tax data from a wide range of different OECD countries, this allows us to calibrate the necessary condition for the desirability of a minimum-wage increase.

Notice that $\mathrm{d}H/\mathrm{d}U$ measures the relationship between high-skilled employment and unemployment for given net wages. We assume that the low-skilled in our model correspond to individuals that did not complete upper-secondary education. High-skilled individuals then correspond to individuals that did complete upper-secondary education.[18] There is a sizable empirical literature that regresses the upper-secondary enrollment rate (e) on the low-skilled unemployment rate ($\bar{u}$) for a panel of regions. It does so while controlling for other factors that influence enrollment rates – most notably the monetary returns to different levels of education. These studies thus produce the following estimate:

$$\eta \equiv \frac{\mathrm{d}e}{\mathrm{d}\bar{u}}. \tag{24}$$

η measures the effect of a percentage-point increase in the low-skilled unemployment rate on the high-school enrollment rate for given net wage incomes. To go from changes in enrollment rates to changes in the high-skilled population, one needs to multiply by the cohort size of people that are eligible for graduation. To go from changes in low-skilled unemployment rates to changes in unemployment, one needs to multiply by the low-skilled labor force. The following Lemma uses this to rewrite the necessary condition of eq. (22) in terms of empirically measurable statistics.

Proposition 4 *A necessary condition for a minimum-wage increase to be more desirable than a distributionally equivalent tax change is:*

$$\eta > \eta^* \equiv \frac{P}{S}\frac{\Delta_L}{\Delta_H}, \tag{25}$$

[18] Admittedly, this is a strong assumption, since dropping out of secondary school does not condemn one to working for a minimum wage, and upper-secondary education is hardly a guarantee for a job at a higher wage rate. Nevertheless, it has long been recognized that schooling is an important factor driving both labor earnings and employment opportunities (e.g. Nickell, 1979; Card, 1999).

where $P \equiv L + \int_{\Theta_1}^{\Theta_2} u_\theta \mathrm{d}\theta$ *gives the low-skilled labor force,* S *is the cohort size of people that are eligible for graduation,* $\Delta_L \equiv \tau^L - \tau^U$ *is the marginal public revenue loss of an increase in unemployment for given high-skilled employment,* $\Delta_H \equiv \tau^H - (1-\bar{u})\tau^L - \bar{u}\tau^U$ *is the marginal public revenue gain of an increase in the number of high-school graduates for a given low-skilled unemployment rate, and* $\bar{u} \equiv \int_{\Theta_1}^{\Theta_2} u_\theta \mathrm{d}G(\theta)/P$ *is the low-skilled unemployment rate.*

Proof. See Appendix A. ■

Thus, the effect of low-skilled unemployment rates on high-school enrollment rates (η) must exceed some critical treshold (η^*) for a minimum-wage increase to be more desirable than a distributionally equivalent tax change. The treshold η^* is increasing in the low-skilled labor force P and the marginal revenue loss from unemployment Δ_L. Intuitively, both factors raise the public revenue losses associated with an increase in the low-skilled unemployment rate and therefore raise the bar for the minimum wage. The treshold η^* is decreasing in the student cohort size S and the marginal revenue gains from high-school graduates Δ_H. Intuitively, both factors raise the public revenue gains associated with an increase in the enrollment rate and therefore lowers the bar for the minimum wage. Below we bring all the parts of eq. (25) to the data.

7.2 Calibrating the necessary condition

7.2.1 Measuring the labor force and cohort size

Since upper-secondary education is typically completed around the age of 18, S is taken to be the size of the 18-year-old population cohort. Values for S are obtained from Eurostat and national statistics offices and reported for a number of OECD countries in the first column of Table 1. The second column gives P, the size of the labor force that completed at most primary education, values of which are obtained from Eurostat, the ILO, and national statistics offices. Both variables are divided by the total labor force (denoted by LF) to enhance comparability across countries. The sample of countries is restricted only by the availability of data on Δ_H and Δ_L.

7.2.2 Measuring the revenue gains from schooling

Δ_H measures the gain in public revenue if one additional person obtains an upper-secondary educational degree. These revenue gains are provided by OECD (2011a, pp. 172-73) for a number of countries and reported in column 3 of Table 1.[19] The OECD considers revenue gains from higher income taxes and employees' social-security contributions, lower transfers, and higher labor utilization (i.e., lower unemployment rates),

[19] For the Netherlands we use data from Ter Rele (2007) because data for this country are not provided by OECD (2011a).

and the revenue losses from direct costs of financing education and the foregone taxes on earnings associated with education. Gains and losses are calculated over the entire life cycle and discounted at a three percent annual real interest rate to obtain the public net present value of an additional high-skilled worker.

7.2.3 Measuring the revenue losses from unemployment

Δ_L measures the loss in public revenue if one additional person becomes unemployed. Statistics on the revenue losses from low-skilled unemployment are extracted from OECD (2011b, p. 56). The OECD reports the participation tax rate of an individual moving from full-time work at 50 percent of the average wage to short-term unemployment. These values take into account the losses from lower income taxes and social-security contributions and higher social, housing, family, and unemployment benefits, together with the gains from lower in-work tax benefits, if applicable. Multiplying these participation tax rates with the average minimum wage income, also from the OECD, we obtain values for Δ_L.[20] These are reported in the fourth column of Table 1.

7.2.4 The effect of unemployment on schooling

There is a sizable empirical literature on the relationship between school enrollment rates and unemployment rates. Clark (2011) surveys this literature for the United Kingdom and presents own estimates. In Appendix C, we survey several other empirical studies. Overall, we find that an increase in the unemployment rate of one percentage point tends to lead to an increase in the enrollment rate of 0.1 to 0.6 percentage points. All except one of the estimated semi-elasticities are well below the highest estimate of 0.6. It is important to keep in mind that proper identification remains a potential concern, since it may be hard to find a proper instrument for the unemployment rate. Instead of using instrumental variables, the literature tends to rely on adding control variables and region- and time-fixed effects to avoid endogeneity bias. Importantly, these control variables include the monetary returns to various levels of education – which our model identifies as an important confounding explanatory variable for skill formation.

7.3 The desirability of a minimum-wage increase

It must be stressed that any implications that are drawn from our calibration exercise are only as accurate as the measurement of the key statistics. While we have made conservative assumptions wherever we could – biasing results in favor of a minimum wage – there are significant uncertainties involved in measuring the costs and benefits of unemployment and education. Moreover, our analysis relies on extrapolating the available

[20] For countries without a minimum wage, we multiply the participation tax rate by 25 percent of average wage income. This represents our sample's lower bound for the minimum wage.

estimates of η towards countries for which no similar estimates have been produced. Keeping these caveats in mind, column 5 in Table 1 provides values of η^*, the right-hand side of the condition in eq. (25).

Our analysis suggests that a NIN minimum-wage increase is desirable only if the semi-elasticity of low-skilled unemployment on enrollment η exceeds this critical value, such that $\eta > \eta^*$. The critical values of η^* range from 0.4 for the United States to 10.1 for Spain.[21] It is useful to consider the two extreme cases in some more detail. In the United States, a minimum-wage increase can only be more desirable than a distributionally equivalent change in taxation if a one percentage-point increase in the unemployment rate leads to at least 0.4 percentage points higher enrollment in education. At the other extreme, a minimum-wage increase in Spain only enhances welfare if a one percentage-point increase in the unemployment rate leads to at least 10.1 percentage points higher enrollment. The reasons for these differences between the United States and Spain are readily observable from Table 1. For the United States, we see that the public benefits of more workers with secondary education (Δ_H) are relatively large. On top of that, the size of the labor force with only primary education (P) is relatively small, such that a one percentage-point increase in the low-skilled unemployment rate is less costly. Spain, on the other hand, shows a relatively small public return to secondary education and a relatively large unskilled population, raising the costs of an increase in the unemployment rate.

To determine whether a minimum-wage increase might be beneficial for the countries under consideration, we compare the calibrated values of η^* with the actual estimates for η. For the United States, our empirical calibration indicates that η should exceed 0.4 for a higher minimum wage to be desirable. As this is within the range of the empirical estimates we found, we cannot reject that a minimum-wage increase might be more beneficial for the United States than a distributionally equivalent change in taxes. However, given that the necessary condition does not take into account direct utility losses of rationing, we do not consider the case for a higher U.S. minimum wage very strong. For all other countries, a one percentage-point increase in the unemployment rate should lead to at least a 0.5 percentage-point increase in enrollment rates. As this approaches the upper bound of the empirical estimates, our results suggest that in these countries a minimum-wage increase is inferior to a distributionally equivalent change in the tax rate. Indeed, they suggest that a *decrease* of the minimum wage, along with compensating tax changes to keep net wages constant, would lead to a Pareto improvement. Such a net-income-neutral decrease of the minimum wage is expected to increase both individuals'

[21] In France, as can be seen from the bottom row of Table 1, net public revenues from a person completing upper-secondary education are in fact negative. Hence, regardless of the value of η, a net-income-neutral *decrease* in the minimum wage leads to a Pareto improvement in France as lower unemployment and lower education both lead to higher public revenue.

Table 1: Calibrating the desirability condition

Country	S/LF (1)	P/LF (2)	Δ_H (3)	Δ_L (4)	η^* (5)	Minimum wage (6)
United States	0.03	0.09	60.5	8.0	0.4	Y
Czech Republic	0.03	0.05	20.0	5.1	0.5	Y
Hungary	0.03	0.12	33.1	6.3	0.7	Y
Germany	0.03	0.14	65.1	8.1	0.7	N
United Kingdom	0.03	0.19	95.0	11.4	0.8	Y
Austria	0.03	0.16	64.7	8.4	0.8	N
Netherlands	0.02	0.25	139.0	16.1	1.2	Y
Poland	0.04	0.08	9.5	5.7	1.3	Y
Sweden	0.03	0.16	29.8	8.3	1.5	N
Norway	0.03	0.20	33.6	10.2	2.3	N
Italy	0.03	0.36	36.6	6.5	2.4	N
Canada	0.03	0.13	25.0	12.9	2.6	Y
Denmark	0.02	0.26	44.5	11.2	2.6	N
Finland	0.03	0.15	15.8	8.0	2.8	N
Slovenia	0.02	0.13	22.7	11.9	2.8	Y
Ireland	0.03	0.20	33.3	14.9	3.1	Y
Australia	0.03	0.27	29.8	9.8	3.2	Y
Portugal	0.03	0.66	43.0	7.8	4.7	Y
Spain	0.02	0.40	15.3	9.4	10.1	Y
France	0.03	0.23	-5.6	13.9	$\Delta_\Theta < 0$	Y

All values 2009 or latest. Δ_H and Δ_L are measured at 2009 prices, in thousands of PPP equivalent USD. Δ_H is an average of male and female values using shares in age-18 cohorts as weights. Δ_L is the unweighted average of revenue losses from an additional unemployed minimum- wage earning single, single parent with 2 children, one-earner married with 2 children, and two-earner married with 2 children with a spouse earning 67 percent of the average wage. All data are available in a separate spreadsheet, available upon request from the authors.

Sources: OECD (2011a, pp. 172-73), OECD (2011b, p. 56), Ter Rele (2007, p. 353), International Labour Organization (2014), Eurostat (2009b), Eurostat (2009a), U.S. Bureau of Labor Statistics (2009), U.S. Census Bureau (2009), Statistics Canada (2009), Australian Bureau of Statistics (2009).

utility and public revenue.[22]

8 Conclusions

Is a minimum wage an appropriate instrument for redistribution? To answer this question, this paper compares the minimum wage with income taxation in an occupational-choice model with competitive labor markets. Compared to a distributionally equivalent change in taxes, a higher minimum wage raises low-skilled wage costs, which leads to low-skilled labor rationing. The distribution of labor rationing critically determines the desirability of a minimum wage. Depending on which individuals are rationed out of the labor market, some of them become unemployed while others decide to upgrade their skills to avoid unemployment and find high-skilled employment instead. If taxes are increasing with income, rationed individuals that become unemployed pay less taxes while those that become high-skilled pay more taxes. Moreover, if rationing is inefficient, it causes utility losses because rationed individuals become involuntarily unemployed or involuntarily high-skilled employed. A minimum-wage increase is more desirable than a distributionally equivalent tax change if and only if the revenue gains from increased high-skilled employment outweigh both the revenue losses from increased unemployment and the utility losses from inefficient rationing.

A necessary condition for the desirability of a minimum wage is that the revenue gains from increased high-skilled employment outweigh the revenue losses from increased unemployment. We express this condition in terms of three sufficient statistics: the tax wedge on participation, the tax wedge on skill formation, and the increase in high-skilled employment for a given increase in unemployment. Calibration of the desirability condition suggests that an increase in income redistribution could be achieved more efficiently by an income-tax reform than by a minimum-wage increase for almost all countries we consider. These countries could achieve a Pareto improvement by reducing their minimum wage while adjusting taxes to keep net incomes unaffected. We therefore conclude that a minimum wage is *not* an appropriate instrument for redistribution.[23]

We suggest a number of promising avenues for further research. First, obtaining more and better-identified estimates of the effect of unemployment on skill formation could greatly enhance our understanding of the welfare effects of a minimum wage relative to taxation. Second, future research may fruitfully apply our general rationing schedule to analyze the desirability of minimum wages in settings with non-competitive labor markets. Third, to rationalize the use of a minimum wage for income redistribution,

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

[23] This does not imply that minimum wages might not be desirable for purposes that are unrelated to redistribution, for example, for correcting market imperfections such as monopsony power or inefficient bargaining power.

further research may need to resort to non-welfarist notions of justice or non-standard individual preferences. That is, a non-welfarist may argue that all working individuals should be able to earn a sufficiently high 'living wage' in the private sector without relying on government transfers. Alternatively, individuals might care about the source of their income, and derive more satisfaction from wage payments than from government transfers (for evidence, see Akay et al., 2012). Fourth, further research might find a useful role for a minimum wage that applies to only a subset of the working population, such as industry- or age-specific minimum wages (e.g., see Kabátek, 2015). Even if the rationing caused by a general minimum wage does not generate sufficient skill formation to render it desirable, the same might not be true for, say, a minimum wage that only applies to working youths.

References

Akay, Alpaslan, Olivier Bargain, Mathias Dolls, Dirk Neumann, Andreas Peichl, and Sebastian Siegloch. 2012. "Happy taxpayers? Income taxation and well-being." IZA Discussion Paper No. 6999, Bonn: IZA.

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

Australian Bureau of Statistics. 2009. "3101.0 Australian Demographic Statistics, Table 59." Canberra. `http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3101.0Jun%202015`.

Autor, David H., Alan Manning, and Christopher L. Smith. 2016. "The contribution of the minimum wage to US wage inequality over three decades: a reassessment." *American Economic Journal: Applied Economics* 8 (1):58–99.

Blumkin, Tomer and Leif Danziger. 2014. "Deserving poor and the desirability of a minimum wage." IZA Discussion Paper No. 8418, Bonn: IZA.

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

Cahuc, Pierre and Guy Laroque. 2013. "Optimal taxation and monopsonistic labor market: Does monopsony justify the minimum wage?" *Journal of Public Economic Theory* 16 (2):259–273.

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

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

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

Chetty, Raj. 2009. "Sufficient statistics for welfare analysis: a bridge between structural and reduced-form methods." *Annual Review of Economics* 1 (2):31–52.

Christiansen, Vidar. 1981. "Evaluation of public projects under optimal taxation." *Review of Economic Studies* 48 (3):447–457.

———. 1984. "Which commodity taxes should supplement the income tax?" *Journal of Public Economics* 24 (2):195–220.

———. 2015. "Optimal Participation Taxes." *Economica* 82 (328):595–612.

Clark, Damon. 2011. "Do recessions keep students in school? The impact of youth unemployment on enrolment in post-compulsory education in England." *Economica* 78 (311):523–545.

Danziger, Eliav and Leif Danziger. 2015. "A Pareto-improving minimum wage." *Economica* 82 (326):236–252.

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.

Diamond, Peter. 1980. "Income taxation with fixed hours of work." *Journal of Public Economics* 13 (1):101–110.

Eurostat. 2009a. "Employment by sex, age and educational attainment level." Brussels: European Commission. `http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfsa_egaed`.

———. 2009b. "Population on 1 January by age and sex." Brussels: European Commission. `http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_pjan`.

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." *Quarterly Journal of Economics* 102 (1):97–108.

Gerritsen, Aart. 2016. "Equity and efficiency in rationed labor markets." Working Paper No. 2016–4, Munich: Max Planck Institute for Tax Law and Public Finance.

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

———. 2014. "Is a minimum wage an appropriate instrument for redistribution?" CESifo Working Paper No. 4588, Munich: CESifo.

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

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

Hummel, Albert Jan and Bas Jacobs. 2016. "Optimal income taxation in unionized labor markets." Mimeo, Erasmus University Rotterdam.

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.

International Labour Organization. 2014. "Key Indicators of the Labour Market." `http://ilo.org/legacy/english/global-reports/kilm2015/kilm14.xlsx`.

Kabátek, Jan. 2015. "Happy birthday, you're fired! The effects of age-dependent minimum wage on youth employment flows in the Netherlands." IZA Discussion Paper No. 9528, Bonn: IZA.

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

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

Lee, David and Emmanuel Saez. 2008. "Optimal minimum wage policy in competitive labor markets." NBER Working Paper No. 14320, Cambridge-MA: NBER.

———. 2012. "Optimal minimum wage policy in competitive labor markets." *Journal of Public Economics* 96 (9–10):739–749.

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

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

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

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

Neumark, David and William Wascher. 2006. "Minimum wages and employment: A review of evidence from the new minimum wage research." NBER Working Paper No. 12663, Cambridge-MA: NBER.

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

OECD. 2011a. "Education at a glance 2011: OECD indicators." Paris: OECD Publishing.

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

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

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

Piketty, Thomas and Emmanuel Saez. 2007. "How progressive is the U.S. federal tax system? A historical and international perspective." *Journal of Economic Perspectives* 21 (1):3–24.

Saez, Emmanuel. 2002. "Optimal income transfer programs: Intensive versus extensive labor supply responses." *Quarterly Journal of Economics* 117 (3):1039–1073.

Schmitt, John. 2013. "Why does the minimum wage have no discernible effect on employment?" Center for Economic and Policy Research, Washington-DC.

Statistics Canada. 2009. "Table 051-0001 - Estimates of population, by age group and sex for July 1." Ontario. `http://www5.statcan.gc.ca/cansim/a26?id=510001`.

Stern, Nicholas. 1982. "Optimum taxation with errors in administration." *Journal of Public Economics* 17 (2):181–211.

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

Ter Rele, Harry. 2007. "Measuring the lifetime redistribution achieved by Dutch taxation, cash transfer and non-cash benefits programs." *Review of Income and Wealth* 53 (2):335–362.

Tumino, Alberto and Mark Taylor. 2015. "The impact of local labour market conditions on school leaving decisions." ISER Working Paper No. 2015-14, Essex: ISER.

U.S. Bureau of Labor Statistics. 2009. "Labor Force Statistics from the Current Population Survey." Washington-DC. `http://www.bls.gov/cps/cps_aa2009.htm`.

U.S. Census Bureau. 2009. "Current Population Survey, Annual Social and Economic Supplement." Washington-DC. `https://www.census.gov/population/age/data/2009comp.html`.

A Proofs

A.1 Proof of Lemma 2

By Lemma 1, the NIN minimum-wage increase features $\mathrm{d}(w^i - \tau^i) = 0$ for $i \in \{H, L, U\}$. Substitute this into the derivatives of eqs. (1)–(6) to obtain eq. (16). Take derivatives of eqs. (7)–(9) and substitute for eq. (16) to find the first equalities in eqs. (17)–(19). Take the total derivative of eq. (12) and substitute for $HF_{LH} = -LF_{LL}$, which follows from homogeneity of degree zero of the marginal products, to find $\mathrm{d}H/H - \mathrm{d}L/L = \varepsilon \mathrm{d}w^L/w^L$, where $\varepsilon \equiv -F_L/(LF_{LL}) > 0$. Substitute for $\mathrm{d}H$ and $\mathrm{d}L$ from the first equalities of eqs. (17)–(18) and for the definition of ρ from eq. (15) to obtain $\int_{\Theta_1}^{\Theta_3} \mathrm{d}u_\theta \mathrm{d}G(\theta) = \alpha\varepsilon \mathrm{d}w^L/w^L$, with $\alpha \equiv (1/L + \rho/H)^{-1} > 0$. Use the last result, along with the definition of ρ, to find the second equalities in eqs. (17)–(19). The final inequalities follow from $\alpha, \varepsilon > 0$ and $\rho \in [0, 1]$.

A.2 Proof of Proposition 1

The total effect of the NIN minimum-wage increase on social welfare equals $\mathrm{d}\mathcal{W}/\lambda + \mathrm{d}\mathcal{B}$. $\mathrm{d}\mathcal{W}/\lambda$ is obtained by taking the total derivative of eq. (13):

$$\text{(A.1)} \qquad \frac{\mathrm{d}\mathcal{W}}{\lambda} = -\int_{\Theta_1}^{\Theta_2} \left(\frac{V_\theta^L - V^U}{\lambda}\right) \mathrm{d}u_\theta \mathrm{d}G(\theta) - \int_{\Theta_2}^{\Theta_3} \left(\frac{V_\theta^L - V_\theta^H}{\lambda}\right) \mathrm{d}u_\theta \mathrm{d}G(\theta).$$

We used eqs. (4)–(6), which imply that marginal changes in the critical levels of ability do not affect social welfare. Multiply and divide the first term by $\int_{\Theta_1}^{\Theta_2} \mathrm{d}u_\theta \mathrm{d}G(\theta)$, and the second term by $\int_{\Theta_2}^{\Theta_3} \mathrm{d}u_\theta \mathrm{d}G(\theta)$, substitute for eqs. (17) and (19), and use the definitions of $\bar{V}_{12}^L$, $\bar{V}_{23}^L$, and $\bar{V}_{23}^H$ to obtain:

$$\text{(A.2)} \qquad \frac{\mathrm{d}\mathcal{W}}{\lambda} = -\left[(1-\rho)\left(\frac{\bar{V}_{12}^L - V^U}{\lambda}\right) + \rho\left(\frac{\bar{V}_{23}^L - \bar{V}_{23}^H}{\lambda}\right)\right] \alpha\varepsilon \frac{\mathrm{d}w^L}{w^L}.$$

The effect of the NIN minimum-wage increase on $\mathcal{B}$ is obtained by taking the total derivative of eq. (14):

$$\text{(A.3)} \qquad \mathrm{d}\mathcal{B} = \tau^U \mathrm{d}U + \tau^L \mathrm{d}L + \tau^H \mathrm{d}H + L\mathrm{d}\tau^L + H\mathrm{d}\tau^H.$$

Substitute for $\mathrm{d}H$, $\mathrm{d}L$ and $\mathrm{d}U$ from eqs. (17)–(19), and for $\mathrm{d}\tau^H = -(L/H)\mathrm{d}\tau^L$ from Lemma 1, to obtain:

$$\text{(A.4)} \qquad \mathrm{d}\mathcal{B} = \left[\rho(\tau^H - \tau^L) - (1-\rho)(\tau^L - \tau^U)\right] \alpha\varepsilon \frac{\mathrm{d}w^L}{w^L}.$$

Finally, the NIN increase in the minimum wage raises social welfare if and only if $\mathrm{d}\mathcal{W}/\lambda + \mathrm{d}\mathcal{B} > 0$. Substitute for eqs. (A.2) and (A.4) to establish the proposition.

A.3 Proof of Proposition 2

In the absence of a binding minimum wage, there is no rationing (i.e., $u_\theta = 0$ for all θ) and the low-skilled wage is endogenously determined by eq. (12). Without rationing, eqs. (7) and (9) indicate that unemployment is determined by the critical ability level Θ_1, and high-skilled employment by the critical level Θ_3. Moreover, from eqs. (4) and (6) we know that Θ_1 only depends on the net incomes of the unemployed and the low-skilled employed, and Θ_3 on the net incomes of the low-skilled employed and the high-skilled employed. Accordingly, we can write unemployment as a function $U = \tilde{U}(c^U, c^L)$, which is increasing in $c^U \equiv -\tau^U$ and decreasing in $c^L \equiv w^L - \tau^L$. Similarly, we can write high-skilled employment as a function $H = \tilde{H}(c^L, c^H)$, which is decreasing in c^L and increasing in $c^H \equiv w^H - \tau^H$. Low-skilled employment follows residually from $L = 1 - U - H$.

It is easiest to solve for the tax optimum by using the primal approach. That is, we maximize social welfare subject to the economy's resource constraint with respect to the allocation (i.e., net wages c^i) rather than taxes τ^i. Substituting for eqs. (1)–(3), $\tau^i = w^i - c^i$, and $u_\theta = 0$ for all θ into the social welfare function of eq. (13) yields:

$$\text{(A.5)}\quad \mathcal{W} = \int_0^{\Theta_1} v(c^U)\mathrm{d}G(\theta) + \int_{\Theta_1}^{\Theta_3} \left(v(c^L) - \frac{1}{\theta} \right) \mathrm{d}G(\theta) + \int_{\Theta_3}^{\overline{\theta}} \left(v(c^H) - \frac{1+\beta}{\theta} \right) \mathrm{d}G(\theta).$$

Notice that eqs. (4) and (6) imply that marginal changes in Θ_1 or Θ_3 do not affect social welfare. Substituting for $\tau^i = w^i - c^i$ and $U = \tilde{U}(c^U, c^L)$, $H = \tilde{H}(c^L, c^H)$ and $L = 1 - U - H$ into the budget constraint of eq. (14) yields the economy's resoure constraint:

$$\text{(A.6)}\quad \mathcal{B} = w^L - c^L - (w^L - c^L + c^U)\tilde{U}(c^U, c^L) + (w^H - c^H - w^L + c^L)\tilde{H}(c^L, c^H) - R.$$

Notice that marginal changes in gross wages do not directly affect the budget constraint. To see this, recall from Lemma 1 that $\mathrm{d}w^H = -(L/H)\mathrm{d}w^L$. Taking the derivative of eq. (A.6) with respect to gross wages, while leaving net wages constant, thus yields $\mathrm{d}\mathcal{B} = (1 - U - H)\mathrm{d}w^L + H\mathrm{d}w^H = L\mathrm{d}w^L + H\mathrm{d}w^H = 0$.

As usual, the Lagrangian for the government optimization problem can be written as $\mathcal{L} \equiv \mathcal{W} + \lambda\mathcal{B}$. Substituting for eqs. (A.5) and (A.6), and taking derivatives with respect to c^H, c^L, and c^U yields the following first-order conditions:

$$\text{(A.7)}\quad \frac{1}{\lambda}\frac{\partial \mathcal{L}}{\partial c^H} = \left(\frac{v'(c^H)}{\lambda} - 1\right) H + (w^H - c^H - w^L + c^L)\frac{\partial \tilde{H}}{\partial c^H} = 0,$$

$$\text{(A.8)}\quad \frac{1}{\lambda}\frac{\partial \mathcal{L}}{\partial c^L} = \left(\frac{v'(c^L)}{\lambda} - 1\right) L - (w^L - c^L + c^U)\frac{\partial \tilde{U}}{\partial c^L} + (w^H - c^H - w^L + c^L)\frac{\partial \tilde{H}}{\partial c^L} = 0,$$

$$\text{(A.9)}\quad \frac{1}{\lambda}\frac{\partial \mathcal{L}}{\partial c^U} = \left(\frac{v'(c^U)}{\lambda} - 1\right) U - (w^L - c^L + c^U)\frac{\partial \tilde{U}}{\partial c^U} = 0.$$

To obtain these first-order conditions, we used the fact that marginal changes in critical ability levels and gross wages do not affect the Lagrangian. These first-order conditions can be rewritten in terms of welfare weights, tax wedges, and elasticities. Substituting for social welfare weights $b^i \equiv \frac{v'(c^i)}{\lambda}$, net income $c^i \equiv w^i - \tau^i$, and elasticities $\varepsilon^H_{\tau^H} \equiv \frac{\partial \tilde{H}}{\partial c^H}\frac{c^H - c^L}{H} > 0$, $\varepsilon^H_{\tau^L} \equiv -\frac{\partial \tilde{H}}{\partial c^L}\frac{c^H - c^L}{H} > 0$, $\varepsilon^U_{\tau^L} \equiv -\frac{\partial \tilde{U}}{\partial c^L}\frac{c^L - c^U}{U} > 0$, and $\varepsilon^U_{\tau^U} \equiv \frac{\partial \tilde{U}}{\partial c^U}\frac{c^L - c^U}{U} > 0$, yields:

$$\text{(A.10)}\qquad 1 - b^H = \left(\frac{\tau^H - \tau^L}{c^H - c^L}\right)\varepsilon^H_{\tau^H},$$

$$\text{(A.11)}\qquad 1 - b^L = \left(\frac{\tau^L - \tau^U}{c^L - c^U}\right)\frac{U}{L}\varepsilon^U_{\tau^L} - \left(\frac{\tau^H - \tau^L}{c^H - c^L}\right)\frac{H}{L}\varepsilon^H_{\tau^L},$$

$$\text{(A.12)}\qquad 1 - b^U = -\left(\frac{\tau^L - \tau^U}{c^L - c^U}\right)\varepsilon^U_{\tau^U},$$

In the tax optimum, the redistributional gains of each tax instrument – given by the left-hand sides of eqs. (A.10)–(A.12) – must equal their marginal dead-weight losses – given by the right-hand sides of eqs. (A.10)–(A.12).

We can prove by contradiction that eqs. (A.10)–(A.12) imply $\tau^H > \tau^L > \tau^U$. First note that any equilibrium with both high- and low-skilled workers must necessarily have $w^H - \tau^H > w^L - \tau^L > -\tau^U$. Together with strict concavity of $v(\cdot)$, this implies that social welfare weights are strictly decreasing in income: $b^U > b^L > b^H$. Now, consider $\tau^U \geq \tau^L$. Eq. (A.12) then implies that $b^U \leq 1$. Moreover, eqs. (A.10) and (A.11) then imply that $b^L \geq 1$ if $\tau^H \geq \tau^L$ and $b^H \geq 1$ if $\tau^H \leq \tau^L$. This contradicts decreasing welfare weights. Thus, in the optimum we must have that $\tau^L > \tau^U$. Now consider $\tau^L \geq \tau^H$. Eq. (A.10) then implies that $b^H \geq 1$. Moreover, eqs. (A.11) and (A.12) then imply that $b^L \leq 1$ if $\tau^L \geq \tau^U$ and $b^U \leq 1$ if $\tau^L \leq \tau^U$. This again contradicts decreasing welfare weights. Thus, in the optimum we must have that $\tau^H > \tau^L$.

A.4 Proof of Proposition 3

In the proof of Proposition 2, we show that the optimal low-skilled tax equates the distributional benefits of taxing the low-skilled to the marginal distortion of taxing the low-skilled. Indeed, substitute for the social welfare weight $b^L \equiv \frac{v'(c^L)}{\lambda}$ and net income

$c^i \equiv w^i - \tau^i$ into eq. (A.8) to get:

$$(1-b^L)L = (\tau^L - \tau^U)\frac{\partial \tilde{U}}{-\partial c^L} - (\tau^H - \tau^L)\frac{\partial \tilde{H}}{-\partial c^L}, \tag{A.13}$$

where $U = \tilde{U}(c^U, c^L)$ and $H = \tilde{H}(c^L, c^H)$ are unemployment and high-skilled employment, in the absence of a binding minimum wage, as functions of net wages. Moreover, if rationing is efficient, the relative increases in high-skilled employment and unemployment that are caused by rationing are identical to the relative increases caused by the low-skilled tax:

$$\frac{\rho}{1-\rho} = \frac{\mathrm{d}H}{\mathrm{d}L} = \frac{\partial \tilde{H}/\partial c^L}{\partial \tilde{U}/\partial c^L}. \tag{A.14}$$

The first condition in eq. (23) follows directly from Proposition 1. The equivalence follows from substituting for eqs. (A.13) and (A.14).

A.5 Proof of Proposition 4

Enrollment rates are related to changes in high-skilled employment as:[24]

$$\mathrm{d}H = S\mathrm{d}e. \tag{A.15}$$

The low-skilled unemployment rate is given by $\bar{u} \equiv \int_{\Theta_1}^{\Theta_2} u_\theta \mathrm{d}G(\theta)/P$, with $P \equiv L + \int_{\Theta_1}^{\Theta_2} u_\theta \mathrm{d}G(\theta) = 1 - G(\Theta_1) - H$. For given net incomes (and thus for given critical levels of ability), the change in the low-skilled unemployment rate is obtained by taking the derivative of $\bar{u}$:

$$\mathrm{d}\bar{u} = \frac{\mathrm{d}U + \bar{u}\mathrm{d}H}{P} \quad \Leftrightarrow \quad \mathrm{d}U = P\mathrm{d}\bar{u} - \bar{u}\mathrm{d}H, \tag{A.16}$$

where we used $\mathrm{d}U = \int_{\Theta_1}^{\Theta_2} \mathrm{d}u_\theta \mathrm{d}G(\theta)$ from eq. (19). Substitute eqs. (A.15)–(A.16) into eq. (22) and rearrange to get:

$$(\tau^H - (1-\bar{u})\tau^L - \bar{u}\tau^U)\frac{S}{P}\frac{\mathrm{d}e}{\mathrm{d}\bar{u}} > \tau^L - \tau^U \tag{A.17}$$

Substitute $\Delta_H \equiv \tau^H - (1-\bar{u})\tau^L - \bar{u}\tau^U$ and $\Delta_L \equiv \tau^L - \tau^U$, and eq. (24) into (A.17) to establish the Proposition.

[24] We assume that drop-out rates are negligible so that every enrolled student graduates. By doing so, we overstate the effect of unemployment on high-skilled employment, biasing our results in favor of the minimum wage.

B Empirical literature on η

Table 2 gives an overview of empirical studies on the impact of unemployment on school enrollment. Earlier UK evidence is surveyed in a similar overview by Clark (2011), in his Table 1. The first column indicates the study of interest; the second column indicates the country of analysis; the third column indicates the time span of the analysis; the fourth column indicates whether the schooling variable refers to enrollment rates (E), or high-school graduation rates (G), whether it refers to boys (b), girls (g), or both (bg), and the age group under consideration; the fifth column indicates to which age-group the unemployment variable refers; the final column gives the estimate of η. The estimated effects of a percentage-point increase in the unemployment rate on the upper-secondary enrollment rate ranges from 0.1 to 0.6. All estimates except for one are well below 0.6.[25]

Empirical studies typically use youth unemployment rates as a proxy for the low-skilled unemployment rate. For example, Clark (2011) estimates the impact of the youth unemployment rate among workers aged 18 and 19 on the enrollment rate for 16-year-olds for a panel of English regions between 1975 and 2005. He finds that a one percentage-point increase in the unemployment rate leads to a 0.32 (0.45) percentage-point increase in the enrollment rate for boys (girls). Moreover, he concludes that this estimate is at least twice as large as those found in previous studies. In another recent study on the UK, Tumino and Taylor (2015) find an effect similar to that of Clark (2011), namely an increase in enrollment of 0.48 percentage point.

While most studies on the impact of unemployment on school enrollment focus on the UK, a few studies analyze the relationship for the United States and Spain. For the US, Card and Lemieux (2001) use variations over states and years to estimate the effect of unemployment on enrollment rates, and find that a one percentage-point increase in the unemployment rate raises school enrollment rates of 17-year-olds by 0.40 percentage point. They also determine the effect of the unemployment rate in the state of birth at age 17 on educational attainment, and find that a one percentage-point increase in the unemployment rate leads to a 0.17 (0.18) percentage-point increase in the share of high-school graduates among boys (girls).[26] In a study on African-American students, Kane (1994) finds the effect to be as large as 0.6.

The disadvantage of the US studies is that data availability confines them to using the prime-age unemployment rate, which is arguably a worse proxy for low-skilled unemploy-

[25]The estimated enrollment responses might capture a combination of structural and temporary business-cycle effects, but it is not clear in which direction this would bias the estimates. η is underestimated (overestimated) if enrollment responds stronger (weaker) to structural unemployment than to temporary unemployment caused by business cycles. We are not aware of empirical evidence that separates both effects on enrollment decisions.

[26]This estimate is likely to suffer from attenuation bias because of interstate migration. After all, the unemployment rate in the state of birth is not likely to affect the schooling decision of a person that moved to another state. On the basis of interstate migration data, Card and Lemieux suspect this attenuation bias to be in the order of 10-25 percent.

Table 2: Empirical estimates of η

Study	Country	Time	Schooling	Unemployment	η
Clark (2011)	UK	1975-	E, b, 16 y/o	18-19 y/o	0.32
		2005	E, g, 16 y/o	18-19 y/o	0.45
Tumino and Taylor (2015)	UK	1991-2009	E, bg, 16 y/o	16-21 y/o	0.4
Card and Lemieux (2001)	US	1968-	E, bg, 15-16 y/o	25-54 y/o	0.14
		1996	E, bg, 17 y/o	25-54 y/o	0.40
		1954-	G, b	25-54 y/o	0.17
		1964	G, g	25-54 y/o	0.18
Kane (1994)	US	1973-1988	G, bg, 18-19 y/o	Total	0.60
Petrongolo and San Segundo (2002)	ES	1991	E, b, 16-17 y/o	16-24 y/o	0.44

The column on schooling indicates whether the dependent variable was the enrollment rate (E) or the high-school graduation rate (G), whether it concerned boys (b), girls (g), or both (bg), and the age group to which the schooling variable refers. Note that Kane (1994) uses the graduation-rate of African Americans. Estimates of Clark (2011) are found in his Tables 2 and 3 on pages 533-534. Estimates for Tumino and Taylor (2015) are found in their Table 3 on page 29. Estimates for Card and Lemieux (2001) are found in their table 9.4 on page 467 and table 9.6 on page 471. Estimates for Kane (1994) are found in his text, page 890. Estimates for Petrongolo and San Segundo (2002) are found in their text, page 364.

ment than the youth unemployment rate used in UK studies. Similar to the UK studies, Petrongolo and San Segundo (2002) analyze the impact of youth unemployment on school enrollment in Spain and find that a one percentage-point increase in the unemployment rate leads to an increase in the enrollment rate for boys of 0.44.

C Labor supply decisions before rationing is realized [Online Appendix]

In the main text we allow individuals to adjust their labor-supply decisions on the basis of whether or not they are able to find a low-skilled job. As a result, every rationed individual with ability $\theta \in [\Theta_2, \Theta_3]$ works as a high-skilled worker. In this Appendix, we show that Proposition 1 and Corollary 1 carry over in modified form when individuals face uncertainty about whether they can obtain a job, and have to make their final labor-supply decision before this uncertainty materializes.

C.1 Model

Utility in each occupation is still given by eqs. (1)–(3). As before, we postulate a general rationing schedule. An individual with ability θ that decides to supply labor as a low-skilled worker becomes unemployed with probability u_θ and is able to find a low-skilled job with probability $1 - u_\theta$. As a result, the expected utility of someone that supplies low-skilled labor equals:

$$V_\theta^{EL} \equiv (1 - u_\theta)V_\theta^L + u_\theta V^U. \tag{C.1}$$

Individuals maximize expected utility when making their occupational choices. Their expected utility equals V^U as a non-participant, V_θ^{EL} when supplying low-skilled labor, and V_θ^H when supplying high-skilled labor. In equilibrium, there is a critical ability level $\Phi_1 \equiv \{\theta : V_\theta^{EL} = V^U \Leftrightarrow V_\theta^L = V^U\}$ at which individuals are indifferent between non-participation and low-skilled labor, and a critical level $\Phi_2 \equiv \{\theta : V_\theta^H = V_\theta^{EL}\}$, at which individuals are indifferent between low-skilled and high-skilled labor supply. These critical ability levels are given by:

$$\Phi_1 = \frac{1}{v(w^L - \tau^L) - v(-\tau^U)}, \tag{C.2}$$

$$\Phi_2 = \frac{\beta + u_{\Phi_2}}{v(w^H - \tau^H) - (1 - u_{\Phi_2})v(w^L - \tau^L) - u_{\Phi_2}v(-\tau^U)}. \tag{C.3}$$

We again focus on the nontrivial case in which at least some individuals strictly prefer low-skilled employment over their outside options, while at least some others strictly prefer high-skilled employment. This requires that $\overline{\theta} > \Phi_2 > \Phi_1$. Moreover, we assume that Φ_2 is unique. In the absence of rationing of low-skilled workers, this is ensured by the assumptions we made on individual preferences. In the presence of low-skilled rationing, however, this requires the shape of the rationing schedule to be 'regular enough'.[27]

[27] Without additional assumptions on u_θ we cannot rule out multiple levels of θ that satisfy eq. (C.3). Also see Gerritsen and Jacobs (2014) for a more complete discussion of these issues.

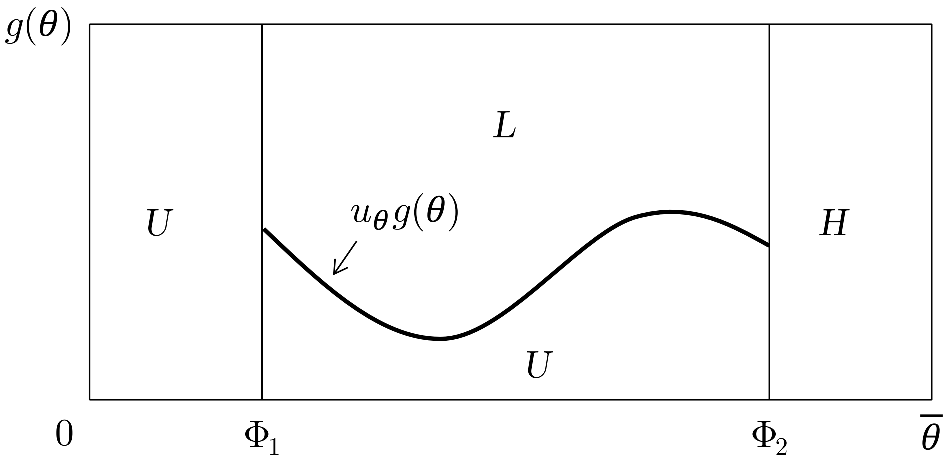

Figure 4: A graphical illustration of an equilibrium

In equilibrium, individuals with ability $\theta \in [0, \Phi_1)$ decide to be voluntarily unemployed, individuals with ability $\theta \in [\Phi_1, \Phi_2]$ decide to supply labor as a low-skilled worker and possibly become involuntarily unemployed, and individuals with ability $\theta \in (\Theta_3, \overline{\theta}]$ decide to supply labor as a high-skilled worker. See Figure 4 for a graphical illustration of an equilibrium. In contrast to the main text, the area under the rationing schedule only represents unemployment. Equilibrium numbers of high-skilled workers, low-skilled workers, and non-participants are given by:

$$H = 1 - G(\Phi_2), \tag{C.4}$$

$$L = \int_{\Phi_1}^{\Phi_2} (1 - u_\theta) \mathrm{d}G(\theta), \tag{C.5}$$

$$U = G(\Phi_1) + \int_{\Phi_1}^{\Phi_2} u_\theta \mathrm{d}G(\theta). \tag{C.6}$$

From eq. (C.4) follows that the number of high-skilled workers only depends on Φ_2. Using eq. (C.3) we can thus write:

$$H = H(u_{\Phi_2}, w^H - \tau^H, w^L - \tau^L, -\tau^U), \tag{C.7}$$

By taking the derivative of eq. (C.3), it can be proven that Φ_2 is decreasing in the rationing rate at ability Φ_2.[28] This implies that high-skilled employment is increasing in the rationing rate at ability level Φ_2, such that $\partial H / \partial u_{\Phi_2} > 0$. Intuitively, an increased probability of low-skilled unemployment for individuals with ability Φ_2 reduces their

[28] In particular, eq. (C.3) allows us to write the critical ability level Φ_2 as a function of net wages and the unemployment rate at Φ_2, such that $\Phi_2 = \Phi_2(u_{\Phi_2}, w^H - \tau^H, w^L - \tau^L, -\tau^U)$. Taking the partial derivative, we obtain:

$$\frac{\partial \Phi_2}{\partial u_{\Phi_2}} = -\frac{1 - \Phi_2/\Phi_1}{v(c^H) - (1 - u_{\Phi_2})v(c^L) - u_{\Phi_2} v(c^U)} < 0.$$

expected low-skilled utility, thereby causing them to supply high-skilled rather than low-skilled labor.

Labor demand is still determined by the firm's first-order conditions in eqs. (11)–(12), and the government's budget constraint by eq. (14). Social welfare now equals:

$$\mathcal{W} \equiv \int_{0}^{\Phi_1} V^U \mathrm{d}G(\theta) + \int_{\Phi_1}^{\Phi_2} V_\theta^{EL} \mathrm{d}G(\theta) + \int_{\Phi_2}^{\overline{\theta}} V_\theta^H \mathrm{d}G(\theta). \tag{C.8}$$

Individual utility maximization implies that social welfare is unaffected by marginal changes in either Φ_1 or Φ_2.

C.2 Comparative statics

We want to know how a minimum-wage increase differs from a distributionally equivalent change in taxes. Again, we determine this by deriving the comparative statics of a net-income-neutral (NIN) minimum-wage increase as defined in Lemma 1. The comparative statics of a NIN minimum-wage increase are graphically illustrated in Figure 5. The minimum wage increase reduces low-skilled employment as indicated by the upward shift of the rationing schedule. Recall that the area under the rationing schedule only captures unemployment. Thus, in contrast to the main text, the upward shift of the rationing schedule purely represents an increase in unemployment.

As in the main text, the reform does not affect the low-skilled participation decision and therefore leaves Φ_1 unaffected. Intuitively, these individuals are indifferent between low-skilled employment and unemployment, so their expected low-skilled utility does not depend on the probability of unemployment. The same does not hold for individuals with ability level Φ_2, whose expected low-skilled utility is strictly decreasing in the probability of unemployment u_{Φ_2}. Thus, the NIN reform leads to an increase in high-skilled employment as long as it raises the unemployment probability at the critical ability level Φ_2. Intuitively, low-skilled individuals with ability Φ_2 escape the higher probability of becoming unemployed by upgrading their skills. This increase in high-skilled employment is illustrated by the leftward shift of Φ_2 in Figure 5.

The equilibrium effects of the NIN minimum-wage increase are thus comparable to the main text. The reform reduces low-skilled employment ($\mathrm{d}L < 0$) and (weakly) raises high-skilled employment ($\mathrm{d}H \geq 0$). While the increase in rationing raises unemployment, the increase in high-skilled employment reduces unemployment as long as some of the individuals with Φ_2 that become high-skilled were previously unemployed. Thus, as long as $u_{\Phi_2} > 0$, the net effect on unemployment is ambiguous ($\mathrm{d}U \lesseqgtr 0$). To formally derive these comparative statics, it is useful to introduce the concept of *rationing incidence*.

Definition 2 *The rationing incidence I_θ at ability level θ gives the increase in the rationing rate at that ability level,* $\mathrm{d}u_\theta$*, as a fraction of the total increase in rationing*

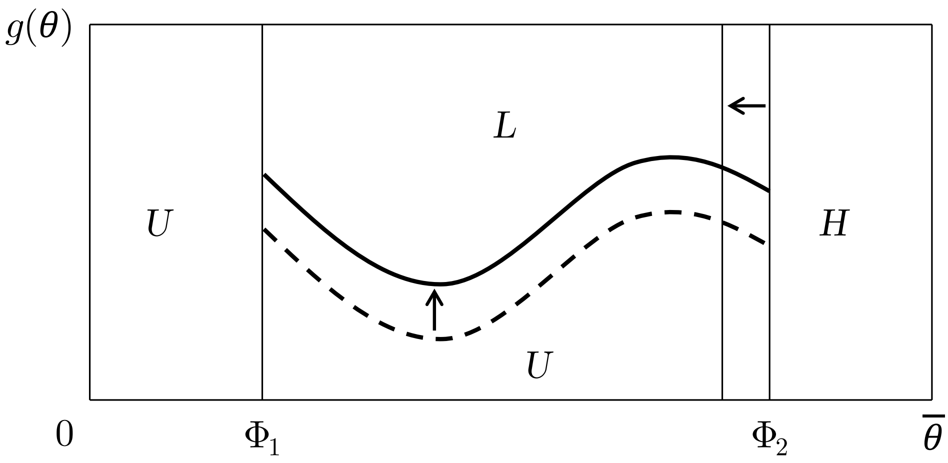

Figure 5: Comparative statics of a net-income-neutral minimum wage increase

$\int_{\Phi_1}^{\Phi_2} \mathrm{d}u_\theta \mathrm{d}G(\theta)$*:*

$$I_\theta \equiv \frac{\mathrm{d}u_\theta}{\int_{\Phi_1}^{\Phi_2} \mathrm{d}u_\vartheta \mathrm{d}G(\vartheta)} \in [0, \infty). \tag{C.9}$$

We assume that $I_\theta \geq 0$ for all θ. By eq. (C.9), this implies that $\mathrm{d}u_\theta \geq 0$ for all θ if $\int_{\Phi_1}^{\Phi_2} \mathrm{d}u_\vartheta \mathrm{d}G(\vartheta) > 0$. As a result, if one unemployment probability increases, all unemployment probabilities weakly increase. The rationing incidence measures the extent to which the NIN minimum-wage increase raises unemployment probabilities at every ability level θ. If $I_\theta = 0$, then individuals with ability θ do not face higher unemployment probabilities, but others generally do. If I_θ approaches infinity, only workers with ability θ see their unemployment probabilities increase, but nobody else. Using this definition of the rationing incidence, we can derive the comparative statics in the following Lemma, which is analogous to Lemma 2 in the main text.

Lemma C.1 *The general-equilibrium comparative statics of the NIN minimum-wage increase, as described by Lemma 1, are:*

$$\mathrm{d}V^U = \mathrm{d}V_\theta^L = \mathrm{d}V_\theta^H = \mathrm{d}\Phi_1 = 0. \tag{C.10}$$

$$\mathrm{d}H = \left(\frac{\partial H}{\partial u_{\Phi_2}} I_{\Phi_2}\right) \alpha\varepsilon \frac{\mathrm{d}w^L}{w^L} \geq 0, \tag{C.11}$$

$$\mathrm{d}L = -\left(1 + (1 - u_{\Phi_2}) \frac{\partial H}{\partial u_{\Phi_2}} I_{\Phi_2}\right) \alpha\varepsilon \frac{\mathrm{d}w^L}{w^L} < 0, \tag{C.12}$$

$$\mathrm{d}U = \left(1 - u_{\Phi_2} \frac{\partial H}{\partial u_{\Phi_2}} I_{\Phi_2}\right) \alpha\varepsilon \frac{\mathrm{d}w^L}{w^L} \lesseqgtr 0, \tag{C.13}$$

where $\varepsilon \equiv -F_L/(LF_{LL}) > 0$ *is the labor demand elasticity and*

$\alpha \equiv \left(\frac{1}{L} + \left(\frac{1-u_{\Phi_2}}{L} + \frac{1}{H}\right)\frac{\partial H}{\partial u_{\Phi_2}} I_{\Phi_2}\right)^{-1} > 0$ *is a share parameter.*

Proof. By Lemma 1, the NIN minimum-wage increase features $\mathrm{d}(w^i - \tau^i) = 0$ for $i \in \{H, L, U\}$. Substitute this into the derivatives of eqs. (1)–(3) and (C.2) to obtain eq. (C.10). Take derivatives of eqs. (C.7) and (C.5)–(C.6) to obtain:

$$\mathrm{d}H = \frac{\partial H}{\partial u_{\Phi_2}} \mathrm{d}u_{\Phi_2}, \tag{C.14}$$

$$\mathrm{d}L = -\int_{\Phi_1}^{\Phi_2} \mathrm{d}u_\theta \mathrm{d}G(\theta) - (1 - u_{\Phi_2})\frac{\partial H}{\partial u_{\Phi_2}} \mathrm{d}u_{\Phi_2}, \tag{C.15}$$

$$\mathrm{d}U = \int_{\Phi_1}^{\Phi_2} \mathrm{d}u_\theta \mathrm{d}G(\theta) - u_{\Phi_2}\frac{\partial H}{\partial u_{\Phi_2}} \mathrm{d}u_{\Phi_2}. \tag{C.16}$$

To obtain eqs. (C.15) and (C.16), we used the derivative of eq. (C.4) which together with eq. (C.14) implies that $\mathrm{d}G(\Phi_2) = -\mathrm{d}H = -\frac{\partial H}{\partial u_{\Phi_2}}\mathrm{d}u_{\Phi_2}$. As in the Proof of Lemma 2, the total derivative of eq. (12) yields $\mathrm{d}H/H - \mathrm{d}L/L = \varepsilon \mathrm{d}w^L/w^L$, with $\varepsilon \equiv -F_L/(LF_{LL}) > 0$. Substitute for $\mathrm{d}H$ and $\mathrm{d}L$ from eqs. (C.14) and (C.15), and the rationing incidence from eq. (C.9), and rearrange to obtain $\int_{\Phi_1}^{\Phi_2} \mathrm{d}u_\theta \mathrm{d}G(\theta) = \alpha\varepsilon \frac{\mathrm{d}w^L}{w^L}$, with $\alpha \equiv \left(\frac{1}{L} + \left(\frac{1-u_{\Phi_2}}{L} + \frac{1}{H}\right)\frac{\partial H}{\partial u_{\Phi_2}} I_{\Phi_2}\right)^{-1}$. The definition of rationing incidence allows us to write $\mathrm{d}u_{\Phi_2} = I_{\Phi_2}\int_{\Phi_1}^{\Phi_2} \mathrm{d}u_\theta \mathrm{d}G(\theta) = I_{\Phi_2}\alpha\varepsilon\frac{\mathrm{d}w^L}{w^L}$. Substitute this into eqs. (C.14)–(C.16) and rearrange to obtain eqs. (C.11)–(C.13). The final inequalities follow from $\alpha, \varepsilon > 0$ and $I_{\Phi_2} \in [0, \infty)$. ■

As net wages remain unaffected, utility remains the same for each occupation, and so does the critical level of ability Φ_1. Moreover, with constant net wages, the change in high-skilled employment is purely determined by the increase in the unemployment probability at ability level Φ_2. As long as there is some rationing at the skill margin, so that I_{Φ_2} is strictly positive, a minimum-wage increase leads to more high-skilled employment than a distributionally equivalent change in taxes.

C.3 Welfare analysis

The next Proposition and Corollary are analogous to Proposition 1 and Corollary 1 in the main text.

Proposition C.1 *A minimum-wage increase is more desirable than a distributionally equivalent change in the tax system if and only if a NIN minimum-wage increase raises social welfare, so that the following condition is satisfied:*

$$(\tau^H - (1 - u_{\Phi_2})\tau^L - u_{\Phi_2}\tau^U)\frac{\partial H}{\partial u_{\Phi_2}} I_{\Phi_2} - (\tau^L - \tau^U) > \frac{\bar{V}^L - V^U}{\lambda}, \tag{C.17}$$

where λ is the shadow value of public resources, and $\bar{V}^L \equiv \int_{\Phi_1}^{\Phi_2} V_\theta^L \mathrm{d}u_\theta \mathrm{d}G(\theta) / \int_{\Phi_1}^{\Phi_2} \mathrm{d}u_\theta \mathrm{d}G(\theta)$ is the average expected low-skilled utility of the individuals that are rationed by the reform.

Proof. The total effect of the NIN minimum-wage increase on social welfare equals $\mathrm{d}\mathcal{W}/\lambda + \mathrm{d}\mathcal{B}$. $\mathrm{d}\mathcal{W}/\lambda$ is obtained by taking the total derivative of eq. (C.8) after substituting for $V_\theta^{EL} \equiv (1-u_\theta)V_\theta^L + u_\theta V^U$:

$$\text{(C.18)} \qquad \frac{\mathrm{d}\mathcal{W}}{\lambda} = -\int_{\Phi_1}^{\Phi_2} \left(\frac{V_\theta^L - V^U}{\lambda} \right) \mathrm{d}u_\theta \mathrm{d}G(\theta).$$

Multiply and divide by $\int_{\Phi_1}^{\Phi_2} \mathrm{d}u_\theta \mathrm{d}G(\theta) = \alpha\varepsilon \frac{\mathrm{d}w^L}{w^L}$ and use the definition of $\bar{V}^L$ to find:

$$\text{(C.19)} \qquad \frac{\mathrm{d}\mathcal{W}}{\lambda} = -\left(\frac{\bar{V}^L - V^U}{\lambda} \right) \alpha\varepsilon \frac{\mathrm{d}w^L}{w^L}.$$

The effect of the NIN minimum-wage increase on $\mathcal{B}$ is obtained by taking the total derivative of eq. (14):

$$\text{(C.20)} \qquad \mathrm{d}\mathcal{B} = \tau^U \mathrm{d}U + \tau^L \mathrm{d}L + \tau^H \mathrm{d}H + L\mathrm{d}\tau^L + H\mathrm{d}\tau^H.$$

Substitute for $\mathrm{d}H$, $\mathrm{d}L$, and $\mathrm{d}U$ from eqs. (C.11)–(C.13) and for $\mathrm{d}\tau^H = -(L/H)\mathrm{d}\tau^L$ from Lemma 1, and rearrange to obtain:

$$\text{(C.21)} \qquad \mathrm{d}\mathcal{B} = \left((\tau^H - (1-u_{\Phi_2})\tau^L - u_{\Phi_2}\tau^U) \frac{\partial H}{\partial u_{\Phi_2}} I_{\Phi_2} - (\tau^L - \tau^U) \right) \varepsilon\alpha \frac{\mathrm{d}w^L}{w^L}$$

Finally, the NIN increase in the minimum wage raises social welfare if and only if $\mathrm{d}\mathcal{W}/\lambda + \mathrm{d}\mathcal{B} > 0$. Substitute for eqs. (C.19) and (C.21) to establish the proposition. ■

Corollary C.1 *A necessary condition for a minimum-wage increase to be more desirable than a distributionally equivalent change in the tax system is that public revenue increases ($\mathrm{d}\mathcal{B} > 0$):*

$$\text{(C.22)} \qquad (\tau^H - (1-u_{\Phi_2})\tau^L - u_{\Phi_2}\tau^U) \frac{\partial H}{\partial u_{\Phi_2}} I_{\Phi_2} > (\tau^L - \tau^U).$$

Or, equivalently:

$$\text{(C.23)} \qquad (\tau^H - \tau^L) \frac{\mathrm{d}H}{\mathrm{d}U} > (\tau^L - \tau^U).$$

Proof. Equate the right-hand side of eq. (C.17) to zero and rearrange to obtain eq. (C.22). Further rearrange to obtain $(\tau^H - \tau^L)\frac{\partial H}{\partial u_{\Phi_2}} I_{\Phi_2} > \left(1 - u_{\Phi_2}\frac{\partial H}{\partial u_{\Phi_2}} I_{\Phi_2}\right)(\tau^L - \tau^U)$. Substitute for eqs. (C.11) and (C.13) to obtain the condition in eq. (C.23). ■

The desirability condition of the minum wage has an identical interpretation as in the main text. The left-hand side of eq. (C.17) represents the net effect of the NIN minimum-wage increase on public revenue, while the right-hand side of eq. (C.17) gives the utility losses due to inefficient rationing ($\mathrm{d}\mathcal{W}/\lambda \leq 0$). Under efficient rationing and absent any pre-existing rationing, we have $\bar{V}^L = V^U$ so that $\mathrm{d}\mathcal{W}/\lambda = 0$. As before, under inefficient rationing we have $\bar{V}^L > V^U$, hence $\mathrm{d}\mathcal{W} < 0$. With increasing taxes ($\tau^H > \tau^L > \tau^U$), a minimum wage is desirable if rationing is concentrated to a large enough extent on individuals with ability Φ_2. In that case, a minimum wage has a strong positive effect on high-skilled employment – yielding revenue gains. A minor difference with the main text is in the first tax wedge on the left-hand-side of eq. (C.17), which now includes the unemployment probability at ability level Φ_2 and the unemployment tax. Intuitively, a marginal increase in high-skilled employment implies that individuals with ability Φ_2 switch towards high-skilled employment. Because a fraction $1 - u_{\Phi_2}$ of these individuals would otherwise be low-skilled employed, and a fraction u_{Φ_2} would otherwise be unemployed, the appropriate tax wedge on skill formation equals $\tau^H - (1 - u_{\Phi_2})\tau^L - u_{\Phi_2}\tau^U$.

Ignoring utility losses, the necessary condition for a minimum wage increase to be more desirable than an equally redistributive change in taxes is once more $\mathrm{d}\mathcal{B} > 0$. Eq. (C.22) is similar in spirit to eq. (21). Whereas the incidence of rationing on ability levels $\theta \in [\Theta_2, \Theta_3]$ is crucial in the main text, it is the incidence on ability level Φ_2 that is crucial in this appendix. In both cases, the effect of unemployment on high-skilled employment can function as a sufficient statistic for the rationing schedule. Indeed, the necessary condition in eq. (C.23) is identical to the one in eq. (22). The empirical analysis in the main text, which is based on the necessary condition in eq. (22), therefore remains entirely unaffected by our assumption on the timing of rationing.

Finally, absent rationing, the model is identical to the one in the main text. As a result, the tax optimum without binding minimum wage is characterized by the same conditions in eqs. (A.10)–(A.12). This implies that Propositions 2 and 3 are still valid. For a given optimal tax schedule ($\tau^H > \tau^L > \tau^U$), a minimum wage is more likely to be desirable if rationing is more efficient and more concentrated on ability level Φ_2.

D Multidimensional heterogeneity [Online Appendix]

In the main text we assume that individuals are heterogeneous with respect to a single parameter θ. Moreover, preferences are such that there is no high-skilled participation margin, i.e., no individuals are indifferent between high-skilled employment and unemployment. In this Appendix, we show that Proposition 1 and Corollary 1 carry over to the case with multidimensional heterogeneity and a high-skilled participation margin.

D.1 Model

We consider a continuum of individuals of mass one that are heterogeneous with respect to low-skilled ability θ^L and high-skilled ability θ^H, drawn from a joint distribution function $G(\theta^H, \theta^L)$ with domain $[0, \infty) \times [0, \infty)$. The marginal distribution function of θ^L is given by $G^L(\theta^L)$, and the conditional distribution function of θ^H by $G^H(\theta^H|\theta^L)$. Utility from high-skilled employment, low-skilled employment, and unemployment are given by:

$$V^H_{\theta^H} \equiv v(w^H - \tau^H) - \frac{1}{\theta^H}, \tag{D.1}$$

$$V^L_{\theta^L} \equiv v(w^L - \tau^L) - \frac{1}{\theta^L}, \tag{D.2}$$

$$V^U \equiv v(-\tau^U). \tag{D.3}$$

We assume that w^L is determined by the minimum wage and that $w^L < w^H$. This assumption is without loss of generality. If high-skilled wages would be larger than low-skilled wages, we could simply relabel the high-skilled as low-skilled and the low-skilled as high-skilled.

We define three critical ability levels. First, we define $\Theta^L \equiv \{\theta^L : V^U = V^L_{\theta^L}\}$ as the unique level of low-skilled ability at which an individual is indifferent between unemployment and low-skilled employment:

$$\Theta^L = \frac{1}{v(w^L - \tau^L) - v(-\tau^U)}. \tag{D.4}$$

This critical ability level corresponds perfectly to Θ_1 in the main text. Second, we define $\Theta^H_1 \equiv \{\theta^H : V^U = V^H_{\theta^H}\}$ as the unique level of high-skilled ability at which an individual is indifferent between unemployment and high-skilled employment:

$$\Theta^H_1 = \frac{1}{v(w^H - \tau^H) - v(-\tau^U)}. \tag{D.5}$$

Finally, we define $\Theta^H_2 \equiv \{\theta^H : V^L_{\theta^L} = V^H_{\theta^H}\}$ as the high-skilled ability level at which an individual is indifferent between high- and low-skilled employment. Since Θ^H_2 also depends on the individual's low-skilled ability, we can write it as a function $\Theta^H_2(\theta^L)$, such that:

$$\Theta_2^H(\theta^L) = \frac{1}{v(w^H - \tau^H) - v(w^L - \tau^L) + \frac{1}{\theta^L}}. \tag{D.6}$$

The critical ability levels separate the population into non-participants, low-skilled workers, and high-skilled workers. An individual prefers to be low-skilled employed if: (i) his low-skilled ability is large enough to prefer low-skilled employment over unemployment, $\theta^L > \Theta^L$, and (ii) his high-skilled ability is low enough to prefer low-skilled over high-skilled employment, $\theta^H < \Theta_2^H(\theta^L)$. We denote the vector of individual abilities as $\boldsymbol{\theta} \equiv (\theta^L, \theta^H)$ and write the rationing rate at ability $\boldsymbol{\theta}$ as $u_{\boldsymbol{\theta}} \equiv u(\theta^L, \theta^H)$. This rationing rate denotes the fraction of all individuals with ability levels $\boldsymbol{\theta}$ that are rationed out of the low-skilled labor market. The rationing schedule assigns a rationing rate to every possible combination of ability levels $\boldsymbol{\theta}$. We can thus write low-skilled employment as:

$$L = \int_{\Theta^L}^{\infty} \left(\int_0^{\Theta_2^H(\theta^L)} (1 - u_{\boldsymbol{\theta}}) \mathrm{d}G^H(\theta^H|\theta^L) \right) \mathrm{d}G^L(\theta^L). \tag{D.7}$$

An individual becomes voluntarily unemployed if: (i) his low-skilled ability is sufficienty low, $\theta^L < \Theta^L$, and (ii) his high-skilled ability is sufficiently low, $\theta^H < \Theta_1^H$. Furthermore, a rationed individual prefers involuntary unemployment over high-skilled employment if his high-skilled ability is sufficiently low, $\theta^H < \Theta_1^H$. Hence, we can write unemployment as:

$$U = \int_0^{\Theta^L} \left(\int_0^{\Theta_1^H} \mathrm{d}G^H(\theta^H|\theta^L) \right) \mathrm{d}G^L(\theta^L) + \int_{\Theta^L}^{\infty} \left(\int_0^{\Theta_1^H} u_{\boldsymbol{\theta}} \mathrm{d}G^H(\theta^H|\theta^L) \right) \mathrm{d}G^L(\theta^L). \tag{D.8}$$

Finally, an individual voluntarily decides to become high-skilled in one of the following two cases. First, if (i) his low-skilled ability is low enough to prefer unemployment over low-skilled employment, $\theta^L < \Theta^L$, and (ii) his high-skilled ability is high enough to prefer high-skilled employment over unemployment, $\theta^H > \Theta_1^H$. Second, if (i) his low-skilled ability is high enough to prefer low-skilled employment over unemployment, $\theta^L > \Theta^L$, and (ii) his high-skilled ability is high enough to prefer high-skilled employment over low-skilled employment, $\theta^H > \Theta_2^H(\theta^L)$. Moreover, any rationed individual prefers high-skilled employment over unemployment if his high-skilled ability is sufficiently large, $\theta^H > \Theta_1^H$.

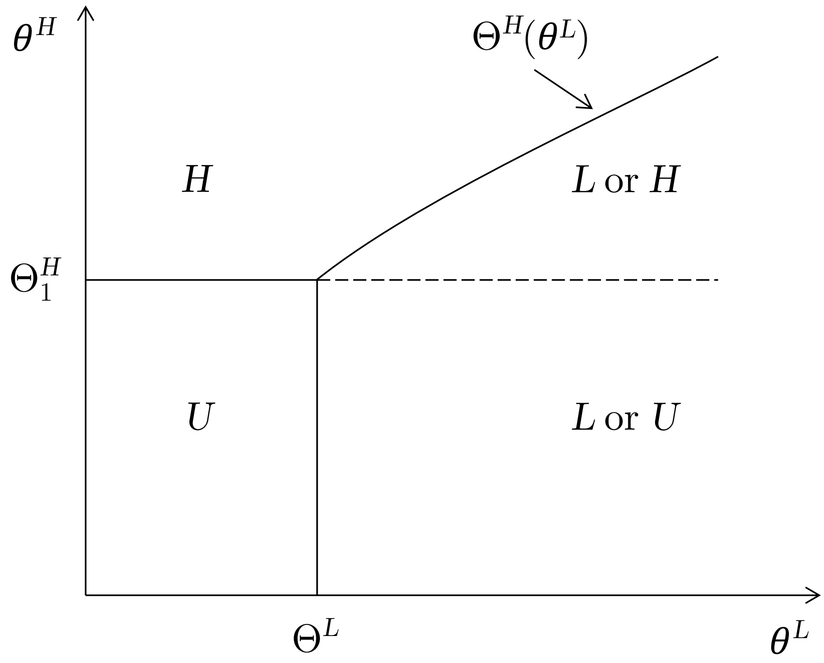

Figure 6: A graphical illustration of an equilibrium

Hence, high-skilled employment is given by:[29]

$$\text{(D.9)}\quad H = \int_0^{\Theta^L} \left(\int_{\Theta_1^H}^{\infty} \mathrm{d}G^H(\theta^H|\theta^L) \right) \mathrm{d}G^L(\theta^L) + \int_{\Theta^L}^{\infty} \left(\int_{\Theta_2^H(\theta^L)}^{\infty} \mathrm{d}G^H(\theta^H|\theta^L) \right) \mathrm{d}G^L(\theta^L) + \int_{\Theta^L}^{\infty} \left(\int_{\Theta_1^H}^{\Theta_2^H(\theta^L)} u_{\boldsymbol{\theta}} \mathrm{d}G^H(\theta^H|\theta^L) \right) \mathrm{d}G^L(\theta^L).$$

Figure 6 gives a graphical illustration of the equilibrium in (θ^L, θ^H)-space. The area denoted by "U" represents the voluntarily unemployed individuals that have relatively low ability for both low- and high-skilled work. The area denoted by "H"represents individuals that prefer high-skilled employment over anything else. Their high-skilled ability is relatively large and their low-skilled ability sufficiently small. The area denoted by "L or U" are individuals that prefer being low-skilled employed, but choose unemployment over high-skilled employment when rationed. Finally, the area denoted by "L or H" are individuals that prefer being low-skilled employed, but choose high-skilled employment over unemployment when rationed.

As in the main text, labor demand is determined by the first-order conditions in eqs. (11)–(12), and the government budget constraint is given by eq. (14). Social welfare now

[29]With respect to the third term in eq. (D.9), notice that $\Theta_2^H(\theta^L) \geq \Theta_1^H$ for all $\theta^L \geq \Theta^L$. Thus, the upper bound of the inner integral always exceeds the lower bound.

equals:

(D.10)

$$\mathcal{W} \equiv UV^U + \int_{\Theta^L}^{\infty} \left(\int_0^{\Theta_2^H(\theta^L)} (1-u_{\boldsymbol{\theta}}) V_{\theta^L}^L \mathrm{d}G^H(\theta^H|\theta^L) \right) \mathrm{d}G^L(\theta^L)$$
$$+ \int_0^{\Theta^L} \left(\int_{\Theta_1^H}^{\infty} V_{\theta^H}^H \mathrm{d}G^H(\theta^H|\theta^L) \right) \mathrm{d}G^L(\theta^L) + \int_{\Theta^L}^{\infty} \left(\int_{\Theta_2^H(\theta^L)}^{\infty} V_{\theta^H}^H \mathrm{d}G^H(\theta^H|\theta^L) \right) \mathrm{d}G^L(\theta^L)$$
$$+ \int_{\Theta^L}^{\infty} \left(\int_{\Theta_1^H}^{\Theta_2^H(\theta^L)} u_{\boldsymbol{\theta}} V_{\theta^H}^H \mathrm{d}G^H(\theta^H|\theta^L) \right) \mathrm{d}G^L(\theta^L).$$

The first term aggregates utility of the unemployed, the second term aggregates utility of low-skilled workers, and the final three terms aggregate utility of high-skilled workers. Individual utility maximization ensures that social welfare is unaffected by marginal changes in any of the critical ability levels.

D.2 Comparative statics

We derive the comparative statics of a net-income-neutral (NIN) increase in the minimum wage to determine how a minimum wage increase differs from a distributionally equivalent change in taxes. As in the main text, the NIN minimum-wage increase reduces low-skilled employment by rationing low-skilled workers out of the labor market. Recall that we defined ρ as the proportion of additional rationing that is concentrated on individuals that prefer high-skilled employment over unemployment. In the multidimensional case, we can thus write ρ as:

(D.11)
$$\rho \equiv \frac{\int_{\Theta^L}^{\infty} \left(\int_{\Theta_1^H}^{\Theta_2^H(\theta^L)} \mathrm{d}u_{\boldsymbol{\theta}} \mathrm{d}G^H(\theta^H|\theta^L) \right) \mathrm{d}G^L(\theta^L)}{\int_{\Theta^L}^{\infty} \left(\int_0^{\Theta_2^H(\theta^L)} \mathrm{d}u_{\boldsymbol{\theta}} \mathrm{d}G^H(\theta^H|\theta^L) \right) \mathrm{d}G^L(\theta^L)} \in [0,1].$$

Notice that the numerator gives the increase in rationing among individuals that prefer low-skilled employment over both high-skilled employment and unemployment (i.e., $\theta^L > \Theta^L$ and $\theta^H < \Theta_2^H(\theta^L)$), but prefer high-skilled employment over unemployment (i.e., $\theta^H > \Theta_1^H$). The denominator simply gives the total increase in rationing among all individuals that otherwise prefer low-skilled employment. The next Lemma formalizes the comparative statics.

Lemma D.1 *The general-equilibrium comparative statics of the NIN minimum-wage increase, as described by Lemma 1, are:*

(D.12)
$$\mathrm{d}V^U = \mathrm{d}V_{\theta^L}^L = \mathrm{d}V_{\theta^H}^H = \mathrm{d}\Theta^L = \mathrm{d}\Theta_1^H = \mathrm{d}\Theta_2^H(\theta^L) = 0.$$

$$\text{(D.13)} \qquad \mathrm{d}H = \int_{\Theta^L}^{\infty}\left(\int_{\Theta_1^H}^{\Theta_2^H(\theta^L)} \mathrm{d}u_{\boldsymbol{\theta}}\mathrm{d}G^H(\theta^H|\theta^L)\right)\mathrm{d}G^L(\theta^L) = \rho\alpha\varepsilon\frac{\mathrm{d}w^L}{w^L} \geq 0,$$

$$\text{(D.14)} \qquad \mathrm{d}L = -\int_{\Theta^L}^{\infty}\left(\int_{0}^{\Theta_2^H(\theta^L)} \mathrm{d}u_{\boldsymbol{\theta}}\mathrm{d}G^H(\theta^H|\theta^L)\right)\mathrm{d}G^L(\theta^L) = -\alpha\varepsilon\frac{\mathrm{d}w^L}{w^L} < 0,$$

$$\text{(D.15)} \qquad \mathrm{d}U = \int_{\Theta^L}^{\infty}\left(\int_{0}^{\Theta_1^H} \mathrm{d}u_{\boldsymbol{\theta}}\mathrm{d}G^H(\theta^H|\theta^L)\right)\mathrm{d}G^L(\theta^L) = (1-\rho)\alpha\varepsilon\frac{\mathrm{d}w^L}{w^L} \geq 0.$$

where $\varepsilon \equiv -F_L/(LF_{LL}) > 0$ *is the labor demand elasticity, and* $\alpha \equiv (1/L + \rho/H)^{-1} > 0$ *is a share parameter.*

Proof. By Lemma 1, the NIN minimum-wage increase features $\mathrm{d}(w^i - \tau^i) = 0$ for $i \in \{H, L, U\}$. Substitute this into the derivatives of eqs. (D.1)–(D.6) to obtain eq. (D.12). Take derivatives of equations (D.7)–(D.9) and substitute for eq. (D.12) to find the first equalities in eqs. (D.13)–(D.15). As in the Proof of Lemma 2, the total derivative of eq. (12) yields $\mathrm{d}H/H - \mathrm{d}L/L = \varepsilon \mathrm{d}w^L/w^L$, with $\varepsilon \equiv -F_L/(LF_{LL})$. Substitute for $\mathrm{d}H$ and $\mathrm{d}L$ from the first equalities in eqs. (D.13) and (D.14) and for ρ from eq. (D.11) to obtain $\int_{\Theta^L}^{\infty}\left(\int_0^{\Theta_2^H(\theta^L)} \mathrm{d}u_{\boldsymbol{\theta}}\mathrm{d}G^H(\theta^H|\theta^L)\right)\mathrm{d}G^L(\theta^L) = \alpha\varepsilon \mathrm{d}w^L/w^L$ with $\alpha \equiv (1/L + \rho/H)^{-1}$. Use the last result, along with the definition of ρ, to find the second equalities in eqs. (D.13)–(D.15). The final inequalities follow from $\alpha, \varepsilon > 0$ and $\rho \in [0, 1]$. ■

Besides the notational adjustments to account for two-dimensional heterogeneity, Lemma D.1 is identical to the corresponding Lemma 2 in the main text. The effects of the NIN minimum-wage increase are therefore the same. As net wages remain unaffected, so do utility and the critical ability levels. As a result, changes in unemployment, low-skilled employment, and high-skilled employment are purely determined by changes in rationing. The reduction in low-skilled employment equals the total increase in rationing. The increase in unemployment equals the increase in rationing among individuals that prefer unemployment over high-skilled employment. And the increase in high-skilled employment equals the increase in rationing among individuals that prefer high-skilled employment over unemployment. Summing up, compared to a distributionally equivalent tax change, a minimum wage increase reduces low-skilled employment and weakly increases both unemployment and high-skilled employment.

D.3 Welfare analysis

As in Proposition 1 of the main text, we define $\bar{V}_{12}^L$ ($\bar{V}_{23}^L$) as the average low-skilled utility of individuals that are rationed by the reform and prefer unemployment over high-skilled employment (high-skilled employment over unemployment), and $\bar{V}_{23}^H$ as the average high-skilled utility of the individuals that are rationed by the reform and prefer high-skilled employment over unemployment. Formally, we have:

$$\bar{V}^L_{12} \equiv \frac{\int_{\Theta^L}^{\infty}\left(\int_0^{\Theta_1^H} V^L_{\theta^L}\mathrm{d}u_{\boldsymbol{\theta}}\mathrm{d}G^H(\theta^H|\theta^L)\right)\mathrm{d}G^L(\theta^L)}{\int_{\Theta^L}^{\infty}\int_0^{\Theta_1^H}\mathrm{d}u_{\boldsymbol{\theta}}\mathrm{d}G^H(\theta^H|\theta^L)\mathrm{d}G^L(\theta^L)}, \tag{D.16}$$

$$\bar{V}^L_{23} \equiv \frac{\int_{\Theta^L}^{\infty}\left(\int_{\Theta_1^H}^{\Theta_2^H(\theta^L)} V^L_{\theta^L}\mathrm{d}u_{\boldsymbol{\theta}}\mathrm{d}G^H(\theta^H|\theta^L)\right)\mathrm{d}G^L(\theta^L)}{\int_{\Theta^L}^{\infty}\int_{\Theta_1^H}^{\Theta_2^H(\theta^L)}\mathrm{d}u_{\boldsymbol{\theta}}\mathrm{d}G^H(\theta^H|\theta^L)\mathrm{d}G^L(\theta^L)}, \tag{D.17}$$

$$\bar{V}^H_{23} \equiv \frac{\int_{\Theta^L}^{\infty}\left(\int_{\Theta_1^H}^{\Theta_2^H(\theta^L)} V^H_{\theta^H}\mathrm{d}u_{\boldsymbol{\theta}}\mathrm{d}G^H(\theta^H|\theta^L)\right)\mathrm{d}G^L(\theta^L)}{\int_{\Theta^L}^{\infty}\int_{\Theta_1^H}^{\Theta_2^H(\theta^L)}\mathrm{d}u_{\boldsymbol{\theta}}\mathrm{d}G^H(\theta^H|\theta^L)\mathrm{d}G^L(\theta^L)}. \tag{D.18}$$

With the help of this notation, we can now replicate Proposition 1 of the main text.

Proposition D.1 *A minimum-wage increase is more desirable than a distributionally equivalent change in the tax system if and only if a NIN minimum-wage increase raises social welfare, so that the following condition is satisfied:*

$$\rho(\tau^H-\tau^L)-(1-\rho)(\tau^L-\tau^U)\geq(1-\rho)\left(\frac{\bar{V}^L_{12}-V^U}{\lambda}\right)+\rho\left(\frac{\bar{V}^L_{23}-\bar{V}^H_{23}}{\lambda}\right), \tag{D.19}$$

where λ is the shadow value of public resources.

Proof. The total effect of the NIN minimum-wage increase on social welfare equals $\mathrm{d}\mathcal{W}/\lambda+\mathrm{d}\mathcal{B}$. $\mathrm{d}\mathcal{W}/\lambda$ is obtained by taking the total derivative of eq. (13):

$$\begin{aligned}\frac{\mathrm{d}\mathcal{W}}{\lambda} = &-\int_{\Theta^L}^{\infty}\left(\int_0^{\Theta_1^H}\left(\frac{V^L_{\theta^L}-V^U}{\lambda}\right)\mathrm{d}u_{\boldsymbol{\theta}}\mathrm{d}G^H(\theta^H|\theta^L)\right)\mathrm{d}G^L(\theta^L)\\ &-\int_{\Theta^L}^{\infty}\left(\int_{\Theta_1^H}^{\Theta_2^H(\theta^L)}\left(\frac{V^L_{\theta^L}-V^H_{\theta^H}}{\lambda}\right)\mathrm{d}u_{\boldsymbol{\theta}}\mathrm{d}G^H(\theta^H|\theta^L)\right)\mathrm{d}G^L(\theta^L).\end{aligned} \tag{D.20}$$

Multiply and divide the first term by $\int_{\Theta^L}^{\infty}\int_0^{\Theta_1^H}\mathrm{d}u_{\boldsymbol{\theta}}\mathrm{d}G^H(\theta^H|\theta^L)\mathrm{d}G^L(\theta^L)$ and the second term by $\int_{\Theta^L}^{\infty}\int_{\Theta_1^H}^{\Theta_2^H(\theta^L)}\mathrm{d}u_{\boldsymbol{\theta}}\mathrm{d}G^H(\theta^H|\theta^L)\mathrm{d}G^L(\theta^L)$, substitute for eqs. (D.13) and (D.15), and use the definitions of $\bar{V}^L_{12}$, $\bar{V}^L_{23}$, and $\bar{V}^H_{23}$ to obtain:

$$\frac{\mathrm{d}\mathcal{W}}{\lambda} = -\left[(1-\rho)\frac{(\bar{V}^L_{12}-V^U)}{\lambda}+\rho\frac{(\bar{V}^L_{23}-\bar{V}^H_{23})}{\lambda}\right]\alpha\varepsilon\frac{\mathrm{d}w^L}{w^L}. \tag{D.21}$$

The effect of the NIN minimum-wage increase on $\mathcal{B}$ is obtained by taking the total derivative of eq. (14):

$$\mathrm{d}\mathcal{B} = \tau^U\mathrm{d}U+\tau^L\mathrm{d}L+\tau^H\mathrm{d}H+L\mathrm{d}\tau^L+H\mathrm{d}\tau^H. \tag{D.22}$$

Substitute for $\mathrm{d}H$, $\mathrm{d}L$, and $\mathrm{d}U$ from eqs. (D.13)–(D.15) and for $\mathrm{d}\tau^H = -(L/H)\mathrm{d}\tau^L$ from Lemma 1 to obtain:

$$\text{(D.23)} \qquad \mathrm{d}\mathcal{B} = \left[\rho(\tau^H - \tau^L) - (1-\rho)(\tau^L - \tau^U)\right]\alpha\varepsilon\frac{\mathrm{d}w^L}{w^L}.$$

Finally, the NIN increase in the minimum wage raises social welfare if and only if $\mathrm{d}\mathcal{W}/\lambda + \mathrm{d}\mathcal{B} > 0$. Substitute for eqs. (D.21) and (D.23) to establish the proposition. ■

Proposition D.1 is identical to Proposition 1 in the main text. Consequently, Corollary 1 in the main text remains valid as well.

D.4 Optimal taxes without minimum wages

What does differ with the main text, is the structure of optimal taxes in the absence of a binding minimum wage. That is, unlike in the model of the main text, the tax wedges $\tau^H - \tau^L$ and $\tau^L - \tau^U$ are no longer necessarily positive in the tax optimum. To see this, we solve for the tax optimum in the absence of a minimum wage so that $u_{\boldsymbol{\theta}} = 0$ for all $\theta = (\theta^L, \theta^H)$. It is again easiest to solve for the tax optimum by using the primal approach. That is, we maximize social welfare subject to the economy's resource constraint with respect to net wages c^i. Substituting eqs. (D.1)–(D.3) and $\tau^i = w^i - c^i$ into the social welfare function in eq. (D.10) yields:

$$\begin{aligned}
\text{(D.24)} \qquad \mathcal{W} \equiv & \int_0^{\Theta^L}\left(\int_0^{\Theta_1^H} v(c^U)\mathrm{d}G^H(\theta^H|\theta^L)\right)\mathrm{d}G^L(\theta^L) \\
& + \int_{\Theta^L}^{\infty}\left(\int_0^{\Theta_2^H(\theta^L)}\left(v(c^L) - \frac{1}{\theta^L}\right)\mathrm{d}G^H(\theta^H|\theta^L)\right)\mathrm{d}G^L(\theta^L) \\
& + \int_0^{\Theta^L}\left(\int_{\Theta_1^H}^{\infty}\left(v(c^H) - \frac{1}{\theta^H}\right)\mathrm{d}G^H(\theta^H|\theta^L)\right)\mathrm{d}G^L(\theta^L) \\
& + \int_{\Theta^L}^{\infty}\left(\int_{\Theta_2^H(\theta^L)}^{\infty}\left(v(c^H) - \frac{1}{\theta^H}\right)\mathrm{d}G^H(\theta^H|\theta^L)\right)\mathrm{d}G^L(\theta^L),
\end{aligned}$$

where the first line aggregates utility of the unemployed, the second line aggregates utility of the low-skilled employed, and the third and fourth lines aggregate utility of the high-skilled employed. Recall that marginal changes in the critical ability levels do not affect social welfare.

Furthermore, notice that eqs. (D.7)–(D.9) allow us to write unemployment and high- and low-skilled employment as functions of net incomes. Thus, we can write $H = \tilde{H}(c^U, c^L, c^H)$ for high-skilled employment, $L = \tilde{L}(c^U, c^L, c^H)$ for low-skilled employment, and $U = \tilde{U}(c^U, c^L, c^H)$ for unemployment. Labor supply in occupation $i \in \{H, L, U\}$ is increasing in 'own' net income c^i, and decreasing in the net income c^j of other occupations

$j \neq i$. Substituting this back into the budget constraint of eq. (14) yields the economy's resource constraint:

(D.25) $$\mathcal{B} = -c^U \tilde{U}(c^U, c^L, c^H) + (w^L - c^L)\tilde{L}(c^U, c^L, c^H) + (w^H - c^H)\tilde{H}(c^U, c^L, c^H) - R.$$

Notice that marginal changes in gross wages do not directly affect the budget constraint. To see this, recall from Lemma 1 that $\mathrm{d}w^H = -(L/H)\mathrm{d}w^L$. Thus, taking the derivative of eq. (D.25) with respect to gross wages, leaving net wages and employment constant, yields $\mathrm{d}\mathcal{B} = L\mathrm{d}w^L + H\mathrm{d}w^H = 0$.

As usual, the Lagrangian for the government optimization problem can be written as $\mathcal{L} = \mathcal{W} + \lambda\mathcal{B}$. Taking derivatives with respect to c^H, c^L, and c^U, while substituting for $\mathrm{d}H + \mathrm{d}L + \mathrm{d}U = 0$ yields the following first-order conditions:

(D.26)
$$\frac{1}{\lambda}\frac{\partial \mathcal{L}}{\partial c^H} = \left(\frac{v'(c^H)}{\lambda} - 1\right) H - (w^H - c^H + c^U)\frac{\partial \tilde{U}}{\partial c^H} - (w^H - c^H - w^L + c^L)\frac{\partial \tilde{L}}{\partial c^H} = 0,$$

(D.27)
$$\frac{1}{\lambda}\frac{\partial \mathcal{L}}{\partial c^L} = \left(\frac{v'(c^L)}{\lambda} - 1\right) L - (w^L - c^L + c^U)\frac{\partial \tilde{U}}{\partial c^L} + (w^H - c^H - w^L + c^L)\frac{\partial \tilde{H}}{\partial c^L} = 0,$$

(D.28)
$$\frac{1}{\lambda}\frac{\partial \mathcal{L}}{\partial c^U} = \left(\frac{v'(c^U)}{\lambda} - 1\right) U + (w^L - c^L + c^U)\frac{\partial \tilde{L}}{\partial c^U} + (w^H - c^H + c^U)\frac{\partial \tilde{H}}{\partial c^U} = 0.$$

These first-order conditions can be rewritten in terms of welfare weights, tax wedges, and elasticities. We do so by substituting for net wages $c^i = w^i - \tau^i$, welfare weights $b^i \equiv v'(c^i)/\lambda$, and elasticities $\varepsilon^H_{\tau^L} \equiv -\frac{\partial \tilde{H}}{\partial c^L}\frac{c^H - c^L}{H} > 0$, $\varepsilon^H_{\tau^U} \equiv -\frac{\partial \tilde{H}}{\partial c^U}\frac{c^H - c^U}{H} > 0$, $\varepsilon^L_{\tau^H} \equiv -\frac{\partial \tilde{L}}{\partial c^H}\frac{c^H - c^L}{L} > 0$, $\varepsilon^L_{\tau^U} \equiv -\frac{\partial \tilde{L}}{\partial c^U}\frac{c^L - c^U}{L} > 0$, $\varepsilon^U_{\tau^H} \equiv -\frac{\partial \tilde{U}}{\partial c^H}\frac{c^H - c^U}{U} > 0$, and $\varepsilon^U_{\tau^L} \equiv -\frac{\partial \tilde{U}}{\partial c^L}\frac{c^L - c^U}{U} > 0$:

(D.29) $$1 - b^H = \left(\frac{\tau^H - \tau^U}{c^H - c^U}\right)\frac{U}{H}\varepsilon^U_{\tau^H} + \left(\frac{\tau^H - \tau^L}{c^H - c^L}\right)\frac{L}{H}\varepsilon^L_{\tau^H},$$

(D.30) $$1 - b^L = \left(\frac{\tau^L - \tau^U}{c^L - c^U}\right)\frac{U}{L}\varepsilon^U_{\tau^L} - \left(\frac{\tau^H - \tau^L}{c^H - c^L}\right)\frac{H}{L}\varepsilon^H_{\tau^L},$$

(D.31) $$1 - b^U = -\left(\frac{\tau^L - \tau^U}{c^L - c^U}\right)\frac{L}{U}\varepsilon^L_{\tau^U} - \left(\frac{\tau^H - \tau^U}{c^H - c^U}\right)\frac{H}{U}\varepsilon^H_{\tau^U}.$$

Contrary to the model in the main text, $\tau^H > \tau^U > \tau^L$ no longer leads to a contradiction. Indeed, as proven by Diamond (1980), Saez (2002), and Christiansen (2015), it might be optimal to subsidize low-skilled workers relative to the unemployed if the participation elasticity of the high-skilled ($\varepsilon^H_{\tau^U}$) is relatively large and the skill elasticity of the low-skilled ($\varepsilon^H_{\tau^L}$) is relatively small.