

CAHIER DE RECHERCHE #1801E
Département de science économique
Faculté des sciences sociales
Université d'Ottawa

WORKING PAPER #1801E
Department of Economics
Faculty of Social Sciences
University of Ottawa

Minimum Wage Policy with Optimal Taxes and Unemployment*

Adam M. Lavecchia[†]

January 2018

* Acknowledgements: I am grateful to Kory Kroft, Philip Oreopoulos and Michael Smart for their guidance on this project. Thanks to Craig Brett, Jessica Burley, David Cimon, Michael Gilraine, Jasjit Goraya, Ashique Habib, Wojciech Koopczuk, Erica Lavecchia, Uros Petronijevic, and participants at the CEA (Ottawa), SOLE (Seattle), NTA (Baltimore) and CPEG (Montreal) conferences whose helpful discussions improved the paper. Financial support from the Social Sciences and Humanities Research Council (SSHRC), Ontario Graduate Scholarship (OGS) and the H. Stanley Hunnisett Fund is gratefully acknowledged.

[†] Department of Economics, University of Ottawa, 120 University Private, Ottawa, Ontario, Canada, K1N 6N5; e-mail: alavecch@uottawa.ca.

Abstract

This paper sheds new light on the desirability of the minimum wage in the presence of an optimal non-linear income tax. Using a search-and-matching framework, I derive a novel condition that links the desirability of the minimum wage to three sufficient statistics: (1) the macro or general equilibrium labor force participation response to the minimum wage by low-skilled individuals; (2) the macro employment response to the minimum wage for low-skilled individuals; and (3) the welfare weight on low-skilled workers. This condition shows that the minimum wage is welfare improving if it pushes the labor market tightness – the ratio of the aggregate number of vacancies to low-skilled job seekers – closer to its efficient level. Guided by the theory, I estimate the first two sufficient statistics using an event study design, as well as state and federal minimum wage variation between 1979-2014. I estimate a macro participation elasticity of -0.24 and a macro employment elasticity of -0.32. The former represents new evidence on a previously overlooked margin of the minimum wage. With these estimates in hand, I simulate the total welfare gains from introducing a minimum wage beginning from the optimal income tax allocation. The simulations show that the minimum wage is welfare improving only if the government has very strong redistributive tastes.

Key words: Minimum wage; Sufficient statistics; Optimal policy; Labor force participation

JEL Classification: H21, H23, J21, J38, J64

1 Introduction

Stagnant real wages among low-skilled workers and rising income inequality are defining features of the U.S. labor market over the past three decades (Kopczuk, Saez, and Song (2010), Gould (2014)). As a response, policy makers have turned to minimum wage legislation as a way to boost incomes.¹ However, wage regulation may increase unemployment, creating a trade-off between higher after-tax incomes and fewer employment opportunities. Some argue that governments can avoid this trade-off by using the income tax system (e.g. the EITC) to increase the income from work (Neumark (2016)). Redistribution using the tax system, however, is limited by fiscal constraints. Policy makers must therefore consider the trade-offs and interactions between the minimum wage and income taxation when developing the optimal redistribution scheme.

This paper analyzes the desirability of minimum wage legislation in the presence of labor market frictions and optimal redistribution through the income tax system. Using a search-and-matching framework, I derive a novel condition that links the desirability of the minimum wage to its effect on low-skilled labor force participation and employment, as well as the welfare weight on low-skilled workers. The key insight from this condition is that the minimum wage can improve both labor market efficiency and redistribution even if income taxes are set optimally.² I then estimate the causal effect of the minimum wage on labor force participation and employment for U.S. low-skilled adults and discuss the policy implications.

I adopt a static search-and-matching framework in which individuals vary along two dimensions: ability and a fixed participation cost. This framework is relevant because search frictions are often used to justify the minimum wage (Flinn (2006), Dube, Lester, and Reich (2016)). The fixed participation cost varies both within and between skills, implying that two individuals with the same ability may make different labor force participation decisions. Labor markets are skill-specific, and the participation decisions of individuals and firms are not observed by the government. The government chooses a non-linear income tax schedule and possibly a minimum wage to maximize a social welfare function of individuals' expected utilities.

The labor market tightness – the ratio of the number of vacancies to job seekers – equalizes supply and demand in each of the low- and high-skilled labor markets. The tightness is important because it affects the probabilities that job seekers and firms match. In turn, these probabilities affect participation decisions, net output, and ultimately, the resources available for redistribution. However, the level of the tightness is usually not efficient, even when income taxes are set optimally.³ There are two reasons why this is the case. First, redistributive income taxation distorts

¹For example, 16 states and several U.S. cities introduced or increased their statutory minimum wage in 2016 (Whitten (2015)).

²This paper complements research in the labor economics literature that derives the conditions under which the minimum wage is desirable using search models without taxes, or in the presence of fixed (not necessarily) optimal tax instruments (Swinnerton (1996), Lang and Kahn (1998), Manning (2004), Flinn (2006), Gravrel (2015)). By allowing the government to set income taxes optimally, I am able to illustrate what the minimum wage does that an optimal non-linear income tax system cannot.

³Given that the focus of this paper is on the desirability of minimum wage legislation (which only affects the low-skilled labor market), all references to the tightness from now on refer to the low-skilled labor market unless otherwise

the labor market tightness, creating an equity-efficiency trade-off. Second, absent any government intervention, externalities caused by search frictions may push the tightness above or below its efficient level. This is because job seekers and firms do not internalize the effect of their participation decisions on the likelihood that other agents find a match (Hosios (1990)).

In the case where the labor market tightness is inefficiently high (absent any government intervention), the government can lower it by setting a positive tax liability or unemployment benefits for low-skilled workers, as in Landais, Michailat, and Saez (2016).⁴ Intuitively, a higher income tax liability reduces the number of vacancies firms are willing to create because it allows workers to negotiate a higher wage. The magnitude of the tax increase is limited by the fact that a higher tax rate lowers labor force participation. This behavioral response shrinks the tax base and limits the ability of the government to push the tightness all the way toward its efficient level. Thus, the distortions to the tightness caused by search frictions and redistribution leaves open the possibility that other policies – notably the minimum wage – can be welfare improving.

In this context, I analyze whether introducing a minimum wage improves total welfare beginning from the optimal income tax allocation. My theoretical contribution is to show that the total welfare gain or loss due to the introduction of a minimum wage can be expressed in terms of three reduced-form parameters⁵: (1) the macro or general equilibrium *labor force participation* response to the minimum wage by low-skilled individuals; (2) the macro *employment* response to the minimum wage for low-skilled individuals; and (3) the *welfare weight on low-skilled workers*. The macro participation and employment responses differ from micro or partial equilibrium responses in that the former incorporate the change to the equilibrium job finding rate caused by the minimum wage.

The intuition for my sufficient statistics condition is the following. First, the minimum wage alleviates the distortion to the tightness caused by search frictions if it increases equilibrium low-skilled labor force participation. Intuitively, the minimum wage increases the cost of low-skilled labor, thereby discouraging vacancy creation just like income taxes. However, unlike a higher employment tax, the effect of the minimum wage on the tax base is relatively small if it increases labor force participation (i.e. a positive macro participation response). As a result, the minimum wage pushes the labor market tightness downwards at a lower fiscal cost than income taxes.

The minimum wage also alleviates the distortion caused by redistributive income taxation if the welfare weight on low-skilled workers is greater than one. In the absence of externalities due to search frictions, the second-best income tax system features a negative employment tax rate for low-skilled workers if the government's welfare weight is greater than one (Saez (2002b), Jacquet, Lehmann, and Van Der Linden (2014), Kroft, Kucko, Lehmann, and Schmieder (2015)). However, this creates an equity-efficiency trade-off. Relative to the laissez-faire economy, a negative em-

stated. In general, the tightness in the high-skilled labor market may also be inefficient.

⁴The case where the low-skilled labor market tightness is inefficiently low (in the absence of any government intervention) is symmetric. In this case, increasing the tightness requires setting a negative employment tax for low-skilled workers. However, this both increases labor force participation and stimulates vacancy creation. The latter offsets the effect of more job seekers on the labor market tightness. In this situation, a wage ceiling restores efficiency.

⁵Using the terminology of Chetty (2009), I refer to these reduced-form parameters as 'sufficient statistics'.

ployment tax increases low-skilled labor force participation and induces firms to create more job openings. These demand and supply responses push the labor market tightness upwards and reduce net output because firms expend resources creating vacancies that go unfilled. The minimum wage alleviates this distortion by pushing the labor market tightness closer to its efficient level in the laissez-faire economy. The change in the tightness (and therefore the workers' job finding rate) due to the minimum wage is captured by the difference between the macro employment and participation responses. The social value of this change is equal to the change in the tightness multiplied by (one minus) the welfare weight on low-skilled workers.

Several papers have analyzed the combination of wage regulation and optimal income taxes under various assumptions about the micro-foundations of the labor market (Allen (1987), Guesnerie and Roberts (1987), Marceau and Boadway (1994), Boadway and Cuff (2001), Lee and Saez (2008, 2012), Hungerbühler and Lehmann (2009), Gerritsen and Jacobs (2016)). In each of these papers, the minimum wage is desirable only if it alleviates the equity-efficiency tradeoff caused by redistributive income taxation. The paper most closely related to my analysis is Hungerbühler and Lehmann (2009). Using a search-and-matching framework, they show that the minimum wage is welfare improving if it relaxes the incentive compatibility constraint in the government's optimal redistribution problem. Unlike Hungerbühler and Lehmann (2009), I model labor supply responses along the extensive margin. This difference is important because income taxes and externalities caused by search frictions distort the tightness in the same direction (upwards) when labor supply responses are along the intensive margin only. However, these distortions can push the tightness in opposite directions when labor supply responses are along the extensive margin. This allows me to distinguish between the distinct labor force participation and redistributive roles of wage regulation, unlike a model with intensive margin responses.⁶ I also improve upon earlier research by generalizing the basic desirability condition to allow for unemployment, wage spillovers and education responses in a tractable sufficient statistics formula.⁷

Guided by the theory, I estimate the macro labor force participation and macro employment responses to the minimum wage for unmarried individuals between the ages of 20 and 29 with less than a high school diploma. Nearly one in five workers in this group work at the minimum wage, making it a good proxy for low-skilled adults. My empirical contribution is to provide new evidence on the causal effect of the macro labor force participation response to the minimum wage, a margin mostly overlooked by the previous literature.⁸

Using pooled cross-sections from the Current Population Survey (CPS) and an event study design that exploits state and federal minimum wage increases between 1979 and 2014, I esti-

⁶Further, I show below that this modeling choice is empirically relevant given that the labor force participation rate of low-skilled workers responds significantly to the minimum wage.

⁷This paper contributes to the growing literature in public economics that derives the welfare implications of tax and transfer policies in terms of sufficient statistics (Saez (2001, 2002b), Chetty (2008), Landais, Michaillat, and Saez (2015), Kroft, Kucko, Lehmann, and Schmieder (2015)).

⁸To the best of my knowledge, the only causal estimates of the labor force participation response to the minimum wage in the U.S. are for teenagers (Wessels (2005), Luna-Alpizar (2015)). I discuss why focusing on low-skilled U.S. adults is more relevant for welfare analysis below.

mate a macro participation elasticity of -0.24. The negative participation response implies that the minimum wage increases over this period reduced the expected utility of low-skilled workers, on average. In line with several recent studies, I also estimate a macro employment elasticity of -0.32 (Clemens and Wither (2014), Clemens (2015), Monras (2017)). The macro participation and employment elasticities imply that the elasticity of the workers' job finding rate with respect to the minimum wage is -0.09. Therefore, approximately two thirds of the decline in employment due to the minimum wage is caused by lower labor force participation, with the remaining third caused by a lower equilibrium job finding rate.

With these estimates in hand, I use the sufficient statistics formula to calculate the marginal welfare gain from introducing a minimum wage for various values of the welfare weight on low-skilled workers. This approach improves upon the large 'reduced-form' empirical literature that estimates the effect of the minimum wage on labor market outcomes, but does not formally study the welfare consequences (Neumark and Wascher (1992), Card and Krueger (1994), Dube, Lester, and Reich (2010, 2016), Neumark, Salas, and Wascher (2014a,b), Belman and Wolfson (2014)). The sufficient statistics approach also allows me to make precise welfare statements about minimum wage increases without estimating deep structural parameters, such as individuals' utility functions or the matching function as in Flinn (2006) and Ahn, Arcidiacono, and Wessels (2011).⁹

Beginning from the *optimal* income tax allocation, my baseline estimates imply that the minimum wage is welfare improving if and only if the welfare weight for low-skilled workers is at least four times as large as the weight on the average individual. This is because the minimum wage reduces labor force participation, leading to lower net output. Consequently, wage regulation is only desirable if the redistributive role is large. Given that the decline in the tightness (i.e. the elasticity of the workers' job finding rate) is relatively small, the welfare weight on low-skilled workers must be large to justify the minimum wage.

This paper is organized as follows. Section 2 sets out the static search-and-matching model. In Section 3, I show how the minimum wage differs from income taxation and derive the sufficient statistics desirability condition. In Section 4, I consider several extensions of the baseline model and show how the desirability condition is modified. Section 5 describes the data and empirical strategy. The main empirical results and robustness checks are reported in Section 6. Section 7 shows how the estimated participation and employment responses can be used to calculate the marginal gain in welfare from introducing a minimum wage. Section 8 discusses the policy implications and promising avenues for future research.

⁹Flinn (2006) and Ahn, Arcidiacono, and Wessels (2011) study the welfare effects of the minimum wage when the government does not have access to other policy instruments.

2 Model

2.1 Individuals and labor markets

This section presents a static search-and-matching model (Diamond (1982); Mortensen and Pissarides (1999)), closely related to the models in several papers in the minimum wage and optimal income taxation literatures (Hungerbühler and Lehmann (2009); Lehmann, Parmentier, and Van Der Linden (2011); Jacquet, Lehmann, and Van Der Linden (2014); Kroft, Kucko, Lehmann, and Schmieder (2015); Landais, Michaillat, and Saez (2016)). Consider an economy with a unit measure of risk neutral individuals that are heterogeneous along two dimensions: ability, denoted by a , and a fixed participation (or search) cost, denoted by θ . For simplicity, I assume there are only two ability levels.¹⁰ An exogenous fraction $n_1 \in (0, 1]$ of agents have ability $a = a_1$ (type-1 agents), and the remaining fraction have ability $a = a_2$ (type-2 agents). I assume that $+\infty > a_2 > a_1 > 0$, so that type-1 agents can be thought of as being ‘low-skilled’ and type-2 agents as ‘high-skilled’. Individuals draw θ from a continuously differentiable and strictly positive conditional distribution $M(\theta|a)$ on the semi-open interval $[0, +\infty)$. Labor markets are perfectly segmented by skill. This means that type- a agents can only participate and work in the labor market that corresponds to their skill level.¹¹ Importantly, conditional on the fixed participation cost θ , all individuals with the same skill level are homogenous. Before turning to the labor supply decision of a type- (a, θ) individual, I describe the demand side of the economy.

2.2 Firms

Firms create vacancies v that require paying a fixed cost χ . This fixed cost may include the human resources required to hire a worker or the capital investment necessary for a worker to be productive. Moreover, a firm may be comprised of multiple vacancies.¹² Let $Y(a) = a$ denote the constant returns to scale production technology that maps skills into output. The profit from a filled vacancy is $a - w_a - \chi_a$, where w_a is the gross wage paid to a worker with skill level a . Due to search-and-matching frictions, a vacancy may be unfilled. Let $h_a = H_a(v_a, k_a)$ denote the matching function that determines how the number of vacancies and job-seekers k_a map into employment h_a . I assume that $H_a(0, k_a) = H_a(v_a, 0) = 0$, that $H_a(\cdot, \cdot)$ is increasing and continuously differentiable in both arguments and exhibits constant returns to scale. Let $\mu_a \in (0, 1)$ denote the elasticity of the matching function with respect to the number job-seekers.

A vacancy for a worker with skill level a matches with probability $q_a = \frac{H_a(v_a, k_a)}{v_a}$. The constant returns to scale assumption on the matching function implies that q_a can be written solely in terms of the labor market tightness v_a/k_a . The firm’s expected profit is $q_a(a - w_a) - \chi_a$. I assume free entry, so expected profits are equal to zero in equilibrium. The free-entry condition $q_a\left(\frac{v_a}{k_a}\right) = \frac{\chi_a}{a - w_a}$

¹⁰Extending the model to include more than two skills does not affect the results.

¹¹The perfect segmentation assumption is common in the literature. See, for example, Lee and Saez (2008) and Hungerbühler and Lehmann (2009).

¹²I implicitly assume that the firm’s aggregate production function is separable in the number of workers.

pins down the equilibrium labor market tightness as a function of the ability of workers and the gross wage.

2.3 Utility maximization and labor supply

Search frictions lead to involuntary unemployment; some workers willing to work at the market wage do not match with a firm. All job seekers with ability a find a match with the uniform probability $p_a(\frac{v_a}{k_a}) = \frac{H_a}{k_a} \in (0, 1)$. Given that p_a is also a function of the labor market tightness, it only depends on ability and the gross wage (i.e. $p_a = p(w_a; a)$).

Individuals take the job finding rate, tax policy and the gross wage as given and make a binary labor force participation decision. Job seekers that match with a firm receive the gross wage w_a , pay taxes $T_a = T(w_a)$ and consume $c_a = w_a - T(w_a)$. Those that are unsuccessful receive a non-employment benefit b .¹³ The *gross expected utility* from participating in the labor market for an individual with skill a is: $U_a = p_a c_a + (1 - p_a)b$; risk neutrality implies that gross expected utility is equal to expected income. The expected utility received by non-participants is $U_0 = b$. Let

$$U_a - U_0 = p_a(c_a - b) \quad (1)$$

denote the expected surplus, or the expected utility gain in from entering the labor market. Participating is optimal if $\theta \leq U_a - U_0$. The critical value for the participation cost for type- a agents is $\hat{\theta}_a = U_a - U_0$. Individuals with a participation cost below $\hat{\theta}_a$ find it optimal to search for a job, whereas those with $\theta > \hat{\theta}_a$ maximize their expected utility by remaining out of the labor force. The participation rate for type- a agents is $M(U_a - U_0|a)$, and the number of job-seekers can be expressed as a function of the expected surplus

$$k_a = n_a \cdot M(U_a - U_0|a) \quad (2)$$

The number of type- a individuals that are employed is $h_a = p_a k_a$.

2.4 Wage setting

Upon meeting, workers and firms bargain over the total surplus from a match: $a - T_a - b$. With an outside option of b , the workers' surplus from a match is $c_a - b = w_a - (T_a + b)$; the firm's surplus is $a - w_a$. Let $\beta_a \in (0, 1)$ denote the worker's (exogenous) bargaining power. Workers and the firms choose the gross wage that maximizes the Nash product $(c_a - b)^{\beta_a} (a - w_a)^{1-\beta_a}$. In the absence of a minimum wage, the equilibrium wage satisfies

$$a - w_a = \frac{1 - \beta_a}{\beta_1} (c_a - b) \quad \Rightarrow \quad w_a^* = \beta_a a + (1 - \beta_a)(T_a + b) \quad (3)$$

¹³Since the model is static, the non-employed do not differ in their job histories. Moreover, the government is assumed not to be able to observe whether an individual searches for a job. Therefore, the assumption that all the non-employed (non-participants and the unemployed) individuals receive the same benefit seems reasonable. This assumption is also made in [Hungerbühler and Lehmann \(2009\)](#) and [Lee and Saez \(2008, 2012\)](#).

It is well understood that wage bargaining does not ensure that the equilibrium allocation is efficient in search-and-matching models (i.e. the equilibrium allocation may not maximize aggregate output net of the fixed participation and vacancy costs).¹⁴ This is because each additional job seeker imposes a positive externality on firms by making it more likely that a vacancy is filled, and a negative externality on other workers by making it less likely that they find a match. Moreover, each additional vacancy makes it less likely that other firms meet a worker, but more likely that workers find a match.¹⁵ These externalities offset each other only if the relative importance of job-seekers to the matching process (given by μ_a) is equal to the fraction of the match surplus workers receive through the wage bargaining process (given by β_a). This is the well-known [Hosios \(1990\)](#) condition.

If $\beta_a < \mu_a$, then the ratio of the number of vacancies to the number of job-seekers v_a/k_a is too high. In this case, resources are wasted due to too many unfilled vacancies. Net output and the expected utility of workers can be increased by reducing the number of vacancies and increasing the number of job-seekers. The opposite is true if $\beta_a > \mu_a$. Externalities due to search-and-matching frictions provide one rationale for government intervention in the economy.

2.5 Government

The government's objective is to maximize an increasing and quasi-concave social welfare function of individuals' expected utilities. Specifically, I assume that the government chooses sets a non-linear income tax schedule (b, T_1, T_2) and possibly a minimum wage \underline{w} to maximize

$$SW(U_0, U_1, U_2) = \sum_{a=a_1}^{a_2} n_a \left\{ \int_0^{\hat{\theta}_a} \Phi(U_a - \theta) m(\theta|a) d\theta + (1 - M(\hat{\theta}_a|a)) \Phi(U_0) \right\} \quad (4)$$

where $\Phi'(\cdot) > 0$ and $\Phi''(\cdot) \geq 0$.¹⁶ The more concave the social welfare function, the more the government would like to redistribute income from type-2 to type-1 workers. The government maximizes $SW(\cdot)$ subject to the following budget constraint

$$\mathcal{B} = (T_1 + b)h_1 + (T_2 + b)h_2 - b - E = 0 \quad (5)$$

where E is an exogenous government spending requirement. The Lagrange multiplier on the government's budget constraint, λ , is the marginal cost of public funds. Let g_a denote the marginal social welfare weight for type- a workers.

¹⁴Given that individuals are risk neutral, an allocation that maximizes aggregate output net of vacancy costs also maximizes the sum of individuals' expected utilities.

¹⁵Put another way, given that wage bargaining takes place after firms and workers meet, the specific investments incurred by workers (the fixed participation cost θ) and firms (the fixed vacancy cost χ_a) may be held up if their investments are not appropriately rewarded at the bargaining stage.

¹⁶I implicitly assume that the optimal non-linear income tax does not generate pooling (i.e. $w_2 > w_1$). [Jacquet, Lehmann, and Van Der Linden \(2014\)](#) derive assumptions on the primitives of the model and the income tax system lead to this result.

$$g_a = \frac{n_a \int_0^{\hat{\theta}_a} \Phi'(\cdot) m(\theta|a) d\theta \cdot p_a}{\lambda h_a}$$

The marginal social welfare weight captures the increase in social welfare from a small increase in the consumption of type- a workers. This value is scaled by the marginal cost of public funds.

3 When does the minimum wage improve welfare?

In this section, I derive the conditions under which introducing a minimum wage is welfare improving when the government sets non-linear income taxes optimally. The analysis proceeds in several steps. First, I show that the sign of the effect of the minimum wage on the expected utility of low-skilled workers can be inferred if the labor force participation response to a higher wage floor is known. Second, I show that the increase in consumption from introducing a minimum wage can be replicated by reducing the income tax liability faced by type-1 workers appropriately. Moreover, I show that this policy increases expected utility by more than the minimum wage. As a result, the minimum wage will only be welfare improving if it is less fiscally costly than a similar transfer using the tax system. I derive conditions under which this is the case.

3.1 The expected utility of low-skilled workers and the minimum wage

The introduction of a binding minimum wage has two effects on the expected utility of type-1 workers. On one hand, a higher gross wage increases expected consumption by $p_1 dc_1 = p_1 d\bar{w}$. On the other hand, the higher gross wage lowers the demand for low-skilled labor. This affects the job finding rate faced by type-1 workers (through a decline in the labor market tightness). Holding the surplus from working $c_1 - b$ constant, a decline in the job-finding rate is costly for workers because it reduces their expected utility by $\frac{\partial p_1}{\partial \bar{w}}(c_1 - b)$. Using (3), the decline in the job finding rate due to the introduction of a minimum wage is¹⁷

$$\left. \frac{\partial p_1}{\partial \bar{w}} \right|_{\bar{w} = w_1^*} = -\frac{1 - \mu_1}{\mu_1} \cdot \frac{p_1}{a_1 - w_1^*} = -\frac{1 - \mu_1}{\mu_1(1 - \beta_1)} \cdot \frac{p_1}{a_1 - T_1 - b} < 0 \quad (6)$$

The minimum wage increases (resp. decreases) U_1 if the consumption effect $p_1 dc_1 = p_1 d\bar{w}$ is larger than (smaller than) the loss in expected surplus due to a lower job finding rate.¹⁸ Using equations (3) and (6), the change in U_1 due to the introduction of a small minimum wage is:

$$dU_1^{MW} \Big|_{\bar{w} = w_1^*} = p_1 \left[1 + \frac{\partial p_1}{\partial \bar{w}} \frac{c_1 - b}{p_1} \right] d\bar{w} = p_1 \left[\frac{\mu_1 - \beta_1}{\mu_1(1 - \beta_1)} \right] d\bar{w} \quad (7)$$

Equation (7) shows that a minimum wage increases the expected utility of type-1 workers if

¹⁷This derivation is due to [Jacquet, Lehmann, and Van Der Linden \(2014\)](#) (Appendix 1).

¹⁸This logic holds regardless of how wages are determined and whether or not the government has access to taxes (including whether or not taxes are set optimally).

and only if the bargaining power is below the [Hosios \(1990\)](#) condition for a laissez-faire economy. Intuitively, if the bargaining power is too low, the labor market tightness (and the job finding rate) is too high. Consequently, workers are willing to trade-off a lower job finding rate for a higher wage. If the worker's bargaining power is equal to the matching function elasticity, then the minimum wage has no effect on U_1 . If $\beta_1 > \mu_1$, imposing a wage ceiling increases U_1 .

Borrowing the terminology from [Landais, Michaillat, and Saez \(2016\)](#) and [Kroft, Kucko, Lehmann, and Schmieder \(2015\)](#), equation (7) should be interpreted as the *macroeconomic* (macro) or general equilibrium expected utility response to the minimum wage. The macro expected utility response differs from a *micro* or partial equilibrium response because it incorporates the equilibrium change to the labor market tightness, and therefore the job finding rate. Unlike the partial equilibrium response, a higher wage does not necessarily imply that macro expected utility response is positive.

If the minimum wage has a non-zero effect on expected utility, it also affects participation decisions. The macro (general equilibrium) participation response to the minimum wage also depends on whether the worker's bargaining power is below the [Hosios \(1990\)](#) condition. Differentiating $k_1 = n_1 M(U_1 - U_0|a_1)$ with respect to the minimum wage yields the macro participation response

$$\frac{dk_1}{d\bar{w}} = n_1 m(U_1 - U_0|a_1) \frac{dU_1}{d\bar{w}} = n_1 m(U_1 - U_0|a_1) p_1 \left[\frac{\mu_1 - \beta_1}{\mu_1(1 - \beta_1)} \right] \quad (8)$$

Intuitively, if the minimum wage increases the expected value of being in the labor force, the number of participants will increase.¹⁹ As shown in [Appendix A.1.2](#), the link between the signs of $\frac{dU_1}{d\bar{w}}$ and $\frac{dk_1}{d\bar{w}}$ does not depend on the assumptions of the static search framework or the level of taxes. This result, mostly overlooked by the empirical literature, is useful for policy makers determining whether minimum wage legislation increases the expected utility of low-skilled individuals ([Wessels \(2005\)](#); [Flinn \(2006\)](#)). Equation (8) shows that the macro participation response to the minimum wage – not the employment response – is proportional to the (expected) utility change.

Let $\left. \frac{\partial k_1}{\partial \bar{w}} \right|^{micro} = n_1 m(\cdot|a_1) p_1 > 0$ denote the *micro* (partial equilibrium) participation response to the minimum wage.²⁰ This response is the change in the number of low-skilled job-seekers due to an increase in the gross wage, holding the job finding rate constant. The following Lemma shows that the macro and micro participation responses are sufficient statistics for calculating the change in the social welfare function $SW(\cdot)$ due to the introduction of minimum wage legislation.

Lemma 1. *Beginning from any allocation, the change in social welfare due to the introduction of a minimum wage is*

¹⁹This insight is related to an observation by [Mincer \(1976\)](#), who noted that the flows into or out of sectors covered by minimum wage legislation are theoretically ambiguous. Also, papers in the labor and urban economics literatures show that whether new immigrants avoid states that recently increased their minimum wage is theoretically ambiguous ([Cadena \(2014\)](#); [Monras \(2017\)](#)).

²⁰The notation $\frac{dy}{d\bar{w}}$ refers to macro responses while $\left. \frac{\partial y}{\partial \bar{w}} \right|^{micro}$ refers to micro responses.

$$\frac{1}{\lambda} \cdot dSW^{MW} = g_1 h_1 \left[\frac{\mu_1 - \beta_1}{\mu_1(1 - \beta_1)} \right] d\underline{w} = g_1 h_1 \cdot \frac{dk_1}{d\underline{w}} \left[\frac{\partial k_1}{\partial \underline{w}} \Big|^{micro} \right]^{-1} \cdot d\underline{w} \quad (9)$$

In the special case when the government does not have access to taxes and transfers (i.e. $b = T_1 = T_2 = 0$), total welfare is given by $\frac{SW}{\lambda} + \mathcal{B} = \frac{SW}{\lambda}$. In this case, a minimum wage is desirable if and only if it increases U_1 and, therefore, $SW(\cdot)$. A sufficient condition for this is $\frac{dk_1}{d\underline{w}} > 0$.²¹

3.2 The minimum wage versus taxes as tools for redistribution

Opponents of minimum wage legislation argue that the income tax system is a better tool for redistribution. This is because the increase in consumption due to the minimum wage can be replicated by a tax reform that reduces the income tax liability faced by type-1 workers. I call this reform a “consumption-equivalent tax cut” (CETC). In the absence of a minimum wage, a CETC must take into account the fact that the gross wage is increasing in T_1 . From equation (3), $\frac{\partial w_1}{\partial T_1} = 1 - \beta_1$. This implies that the income tax liability must be reduced by *more* than one dollar to provide the same increase in consumption as a one dollar increase in the minimum wage. This will be important for understanding the differences between the fiscal costs of the two policies below. Formally, a CETC is defined as follows

$$\begin{aligned} dc_1^{CETC} &= \left[\frac{\partial w_1}{\partial T_1} - 1 \right] dT_1^{CETC} = d\underline{w} = dc_1^{MW} \\ \Rightarrow dT_1^{CETC} &= -\frac{1}{\beta_1} d\underline{w} \end{aligned}$$

The effect of a CETC on the expected utility of low-skilled workers differs from that of the minimum wage. Although both policies increase consumption by the same amount, a CETC reduces the gross wage, leading firms to create more vacancies (see equation (6)). This increases the labor market tightness and the job finding rate.²² Consequently, a CETC increases the expected utility of type-1 workers (and therefore $SW(\cdot)$) by more than a minimum wage. This is formalized by the following Lemmas.

Lemma 2. *Beginning from any allocation, the change in social welfare due to a CETC is*

$$\frac{1}{\lambda} \cdot dSW^{CETC} = g_1 h_1 \cdot \frac{1}{\mu_1} d\underline{w} > 0 \quad (10)$$

Moreover, a CETC unambiguously increases labor force participation.

²¹In a derivation available upon request, I show that the sign of the labor force participation response is sufficient to test the Hosios (1990) and the form of equation (14) is unchanged if job seekers choose search intensity as well as their participation status. This is because both the macro labor force participation response to the minimum wage and the macro search effort response to the minimum wage are proportional to the macro expected utility response. Given that labor market participation responses are easier to estimate than search intensity responses, it is reasonable to include the former in an empirically implementable formula for the welfare effect of minimum wage increases.

²²Totally differentiating p_1 leads to: $dp_1^{CETC} = \frac{\partial p_1}{\partial w_1} \frac{\partial w_1}{\partial T_1} dT_1^{CETC} = -\frac{(1-\mu_1)p_1}{\mu_1[a_1-T_1-b]} \cdot dT_1^{CETC} = \frac{(1-\mu_1)p_1}{\mu_1[a_1-T_1-b]} \frac{1}{\beta_1} d\underline{w} > 0$.

Lemma 3. *Beginning from any allocation, the minimum wage increases social welfare (and therefore the expected utility of type-1 workers) by less than a CETC. Formally,*

$$dSW_1^{MW} - dSW_1^{CETC} = -g_1 h_1 \left[\frac{1 - \mu_1}{\mu_1(1 - \beta_1)} \right] d\underline{w} < 0 \quad (11)$$

Equation (11) shows that taxes have better redistributive properties than the minimum wage. This result holds in any model in which a tax cut and the minimum wage have opposing effects on the job finding rate.²³ However, the government also faces a revenue constraint and cannot decrease the income tax liability of type-1 workers without regard for the tax revenue consequences. Given that taxes have better redistributive properties, wage regulation can only be part of the policy optimum if it achieves the same increase in consumption as a CETC at a lower fiscal cost.

3.3 The fiscal consequences of the minimum wage and taxes

3.3.1 Fixed taxes

Income taxes and the minimum wage have different fiscal costs. Before considering the case where income taxes are set optimally, I first describe how the two policies differ for any (fixed) tax system. This exercise shows how taxes interact with minimum wage legislation. Recall that the government's budget constraint is $\mathcal{B} = \sum_{a=a_1}^{a_2} (T_a + b)h_a + T_0 - E = 0$.

The change in net tax revenue due the minimum wage is $d\mathcal{B}^{MW} = (T_1 + b)dh_1^{MW}$. The fiscal cost depends on the sign and magnitude of the macro employment response dh_1^{MW} , as well as the sign and magnitude of the employment tax faced by low-skilled workers. Given that both the employment tax and dh_1^{MW} may be positive or negative, the sign of the net revenue effect of the minimum wage is ambiguous and is ultimately an empirical question.²⁴

Similarly, the change in net tax revenue due to a CETC may be positive or negative. Beginning from an allocation with no minimum wage, the fiscal cost of a CETC is $d\mathcal{B}^{CETC} = h_1 dT_1^{CETC} + (T_1 + b)dh_1^{CETC}$. The first term corresponds to the mechanical decrease in tax revenue; each type-1 worker is transferred an additional $dT_1^{CETC} = -\beta_1^{-1}d\underline{w}$ units of the consumption good. The second term corresponds to the behavioral (employment) response to taxes. Both labor supply and demand responses to a tax cut lead to higher employment.²⁵ However, the sign of the behavioral response (and ultimately $d\mathcal{B}^{CETC}$) is ambiguous because the initial employment tax faced by type-1 workers may be positive or negative.

²³To see this in a more general set-up, note that $dSW^{MW} - dSW^{CETC} = h_1 \left(\frac{c_1 - b}{p_1} \right) \left[\frac{\partial p_1}{\partial \underline{w}} - \frac{\partial p_1}{\partial T_1} \left(\frac{\partial w_1}{\partial T_1} - 1 \right) \right] d\underline{w}$. This expression is negative if a tax cut increases the job finding rate (i.e. $\frac{\partial p_1}{\partial T_1} \leq 0$), while a minimum wage reduces the job finding rate (i.e. $\frac{\partial p_1}{\partial \underline{w}} \leq 0$).

²⁴Employment may increase with the minimum wage if the macro participation response is positive and larger than the decline in the job finding rate.

²⁵To see this, note that $dh_1^{CETC} = \frac{dh_1}{dT_1} dT_1^{CETC} = \left[p_1 \frac{dk_1}{dT_1} + k_1 \frac{\partial p_1}{\partial T_1} \right] dT_1^{CETC}$. The term in square braces is negative because $\frac{dk_1}{dT_1} = -\frac{\beta_1}{\mu_1} n_1 m(\cdot | a_1) p_1 < 0$ and $\frac{\partial p_1}{\partial T_1} = \frac{\partial p_1}{\partial w_1} (1 - \beta_1) < 0$. Moreover, $dT_1^{CETC} = -\beta_1^{-1} d\underline{w} < 0$.

The difference between the fiscal costs of the two policies is $d\mathcal{B}^{MW} - d\mathcal{B}^{CETC}$. Beginning from any allocation where the employment tax faced by type-1 workers is non-positive $T_1 + b \leq 0$, the minimum wage increases c_1 at a lower revenue cost than a CETC. This is because the positive employment response to a CETC reduces the tax base. Since the fiscal cost of the minimum wage is always less than that of a CETC whenever $T_1 + b \leq 0$, there are also positive values of the employment tax for which $d\mathcal{B}^{MW} - d\mathcal{B}^{CETC} > 0$.

3.3.2 Optimal taxes

In the absence of a minimum wage, the government sets a non-linear income tax to redistribute income from type-2 to type-1 workers (redistribution goal), and to reduce the externalities associated with inefficient search and vacancy decisions (efficiency goal). To see why the minimum wage may be desirable, it is useful to check how well the income tax system achieves the redistribution and efficiency goals in the absence of wage regulation. The following Lemma, due to [Jacquet, Lehmann, and Van Der Linden \(2014\)](#) and [Kroft, Kucko, Lehmann, and Schmieder \(2015\)](#), shows that the optimal employment tax for type-1 workers depends on the bargaining power, the elasticity of the matching function, and the government's marginal social welfare weight.

Lemma 4. *The optimal employment tax liability for type-1 workers is given by*

$$T_1^* + b^* = -h_1 \left(\frac{\mu_1 - \beta_1 g_1}{\mu_1} \right) \left(\frac{dh_1}{dT_1} \right)^{-1} \quad (12)$$

Successively setting $g_1 = 1$ and $\beta_1 = \mu_1$ isolates the efficiency and redistribution roles of the optimal income tax system. If the social welfare function is utilitarian, then $g_1 = 1$, and the government's goal is to use the income tax system to move the labor market tightness closer to the level that maximizes individuals' expected utilities. Given that $\frac{dh_1}{dT_1} < 0$, the sign of the optimal employment tax depends only on $\mu_1 - \beta_1$. If the worker's bargaining power is below the value prescribed by the [Hosios \(1990\)](#) condition, the employment tax liability faced by type-1 workers is *positive*. Intuitively, the government sets a positive tax liability to increase the gross wage and lower the labor market tightness. However, in addition to reducing the number of vacancies, a higher employment tax discourages labor force participation. The lower level of participation reduces the tax base, thereby raising the fiscal cost of lowering the labor market tightness.

If $\beta_1 = \mu_1$, the tightness in the laissez-faire economy maximizes the sum of individuals' expected utilities. In this case, the government's goal is to redistribute income from type-2 to type-1 workers if $g_1 > 1$.²⁶ However, redistribution leads to an equity-efficiency trade-off. The larger is g_1 , the more negative the government would like to set the employment tax liability for type-1 workers ([Saez \(2002b\)](#)). However, a negative tax liability encourages labor force participation,

²⁶By rearranging the first-order condition for the non-employment benefit b , it is possible to show that $g_0 h_0 + g_1 h_1 + g_2 h_2 = 1$ (see [Saez \(2002b\)](#), [Jacquet, Lehmann, and Van Der Linden \(2014\)](#), and [Kroft, Kucko, Lehmann, and Schmieder \(2015\)](#)). In words, the weighted average of the welfare weights is one at the optimal income tax allocation.

increases the labor market tightness, and reduces net tax revenue.²⁷ Since redistributive income taxation distorts the labor market tightness (relative to the laissez-faire economy), policy instruments that alleviate this distortion may be desirable.

The following Lemma derives the sufficient conditions under which the minimum wage is less fiscally costly than a CETC, beginning from the optimal income tax allocation.

Lemma 5. *Beginning from the optimal income tax allocation without a minimum wage, the minimum wage is less fiscally costly than a CETC if $\frac{dk_1}{d\underline{w}} > 0$ and $g_1 > \mu_1$.*

Proof

Subtracting $d\mathcal{B}^{CETC}$ from $d\mathcal{B}^{MW}$ and using Lemma 4 leads to

$$\frac{d\mathcal{B}^{MW}}{d\underline{w}} - \frac{d\mathcal{B}^{CETC}}{d\underline{w}} = -h_1 \left(\frac{dh_1}{dT_1} \right)^{-1} \left[p_1 \frac{dk_1}{d\underline{w}} - p_1 \frac{dk_1}{dT_1} \frac{(1-\mu)g_1}{\mu_1(1-\beta_1)} + k_1 \frac{\partial p_1}{\partial \underline{w}} \frac{(\mu_1 - g_1)}{\mu_1} \right] \quad (13)$$

Since $\frac{dk_1}{dT_1} < 0$ and $\frac{dh_1}{dT_1} < 0$, (13) is positive whenever $\frac{dk_1}{d\underline{w}} > 0$ and $g_1 > \mu_1$. \square

To understand the first condition in Lemma 5, first consider the case where the government does not care about redistribution (i.e. assume a utilitarian social welfare function and $g_1 = 1$). In this case, the second and third terms in square brackets are positive, so the sign of (13) depends on the macro participation response to the minimum wage. Whenever $\beta_1 < \mu_1$, the macro participation response is positive. Intuitively, the increase in labor force participation mitigates the decline in employment (and the tax base) due to the minimum wage, so that the fiscal cost of wage regulation is relatively small. In contrast, introducing a CETC always reduces net tax revenue.

To understand the second condition, re-introduce redistribution and assume that the low-skilled labor market is efficient in the absence of taxation or wage regulation (set $\beta_1 = \mu_1$). In this case, the sign of (13) depends on the third term in square brackets. A larger value of g_1 means that the government values redistribution towards type-1 workers, pushing the employment tax downwards and the tightness in the low-skilled labor market upwards. In this case, the employment (behavioral) response to the minimum wage reduces the tax base by less than a CETC.

Lemma 3 shows that taxes are a better redistributive tool than the minimum wage in the absence of fiscal considerations. On the other hand, Lemma 5 shows that redistribution through the minimum wage can be less fiscally costly than taxes. Therefore, a minimum wage (resp. wage ceiling) will be part of the policy optimum whenever the fiscal savings are larger than (resp. smaller than) the utility effect.

²⁷In the case where $\beta_1 < \mu_1$ and $g_1 > 1$, there is a tension between the government's desire for redistribution and efficiency. On one hand, the government would like to set a positive employment tax to increase the gross wage and the lower the labor market tightness (efficiency). On the other hand, the government would like to lower the employment tax to transfer more income to type-1 workers; this leads to a tighter labor market. The case where $\beta_1 > \mu_1$ is symmetric.

3.4 Desirability condition for the minimum wage

I now combine the results from the previous two sections. Let \mathcal{L}^* denote the maximal value of the government's (constrained) objective function at the optimal income tax allocation in the absence of a minimum wage. The change in social welfare due to the introduction of a minimum wage is $\lambda^{-1}dSW^{MW} = g_1 h_1 \left[\frac{\mu_1 - \beta_1}{\mu_1(1 - \beta_1)} \right] d\underline{w}$. The net revenue effect of the minimum wage is $d\mathcal{B}^{MW} = (T_1^* + b^*)dh_1^{MW}$. Therefore, the change in total welfare is $\frac{1}{\lambda} \frac{d\mathcal{L}^*}{d\underline{w}} \Big|_{\underline{w} = w_1^*} = \frac{1}{\lambda} \frac{dSW^{MW}}{d\underline{w}} + \frac{d\mathcal{B}^{MW}}{d\underline{w}}$. If either of these two terms is non-zero, then wage regulation is part of the policy optimum. To check the conditions under which the sum of these terms is non-zero, I compare total welfare effects of the minimum wage and a CETC reform, beginning from the optimal income tax allocation.²⁸ Since the income tax is optimal, the CETC has no first-order effect on total welfare: $\frac{1}{\lambda} \frac{dSW^{CETC}}{d\underline{w}} + \frac{d\mathcal{B}^{CETC}}{d\underline{w}} = 0$. As a result, the change in total welfare due to the minimum wage can be expressed as $\frac{1}{\lambda} \frac{d\mathcal{L}^*}{d\underline{w}} \Big|_{\underline{w} = w_1^*} = \frac{1}{\lambda} \frac{dSW^{MW}}{d\underline{w}} - \frac{1}{\lambda} \frac{dSW^{CETC}}{d\underline{w}} + \frac{d\mathcal{B}^{MW}}{d\underline{w}} - \frac{d\mathcal{B}^{CETC}}{d\underline{w}}$. Subtracting (13) from (11) leads to the main theoretical result in the paper.

Proposition 1. *Beginning from the optimal income tax allocation, the change in total welfare from the introduction of a minimum wage is given by*

$$\frac{d\mathcal{L}^*}{d\underline{w}} \Big|_{\underline{w} = w_1^*} \propto \underbrace{p_1 \frac{dk_1}{d\underline{w}}}_{\text{Participation effect}} + \underbrace{k_1 \frac{\partial p_1}{\partial \underline{w}} (1 - g_1)}_{\text{Redistribution effect}} \quad (14)$$

Sufficient conditions for the minimum wage to be desirable are $\frac{dk_1}{d\underline{w}} > 0$ ($\beta_1 < \mu_1$) and $g_1 > 1$.

The proof is presented in Appendix A.1.3. Equation (14) shows that the minimum wage plays two distinct roles alongside income taxes as part of the policy optimum. I refer to the first term in (14) as the *Participation effect* of the minimum wage and the second term as the *Redistribution effect*. The former is positive if $\frac{dk_1}{d\underline{w}} > 0$, and the latter term is positive if $g_1 > 1$. Introducing a minimum wage (resp. wage ceiling) is desirable if the sum of these two terms is positive (resp. negative). Successively setting $g_1 = 1$ and $\beta_1 = \mu_1$ isolates the two roles of the minimum wage.

The first role is to increase net output by acting as a corrective tax on the tightness in the low-skilled labor market. This *Participation effect*. When $\beta_1 < \mu_1$, the tightness is above the level prescribed by the Hosios (1990) condition. In this case, the minimum wage induces firms to create fewer vacancies and encourages labor force participation. The increase in participation reinforces the decline in vacancies, thereby lowering the tightness and increasing net output. The expected increase in net output due to the minimum wage is proportional to $p_1 \frac{dk_1}{d\underline{w}}$. To contrast, lowering the labor market tightness using the income tax system is inefficient because taxes have opposing

²⁸Mathematically, this is equivalent to introducing a minimum wage while simultaneously increasing the tax liability for type-1 workers so that consumption (redistribution) is unchanged. This approach is adopted by Gerritsen and Jacobs (2016) for the case of the minimum wage in a model in which the labor market is competitive. This approach is also common in the literature on the desirability of non-uniform commodity taxation (Saez (2002a)). See also Kaplow (2012) on the desirability of Pigouvian taxation in the presence of externalities and income taxes.

effects on vacancies and participation. Although a higher employment tax also reduces the aggregate number of vacancies, it reduces labor force participation. The latter is fiscally costly because it leads to a smaller tax base.²⁹

The minimum wage also alleviates the distortion to the labor market tightness caused by redistributive income taxation; this is the *Redistribution effect*. To see this, suppose $\beta_1 = \mu_1$, so that the laissez-faire economy is efficient. In this case, if the government values redistribution towards low-skilled workers (i.e. $g_1 > 1$), then the employment tax rate faced by type-1 workers is negative. However, a negative employment tax distorts the labor market tightness upwards. Introducing a minimum wage alleviates this distortion by moving it closer to the level in the efficient laissez-faire economy. As a result, net output increases, thereby increasing the resources available for redistribution. The *Redistribution effect* is proportional to the effect of the minimum wage on the job-finding rate faced by low-skilled workers $k_1 \frac{\partial p_1}{\partial w}$.

Proposition 1 generalizes the results in Marceau and Boadway (1994) and Lee and Saez (2012) by deriving the sufficient conditions under which the minimum wage is welfare improving in an economy with search frictions. In Marceau and Boadway (1994) and Lee and Saez (2012) the labor market is perfectly competitive and labor supply responses are along the extensive margin. If the labor market is perfectly competitive, the minimum wage is welfare improving only if it undoes the distortion to the tightness in the low-skilled labor market caused by the second-best income tax system. This is the *Redistribution effect* described earlier.³⁰ Proposition 1 shows that relaxing the assumption of competitive labor markets leads to an efficiency role for the minimum wage that depends on the participation response $\frac{dk_1}{dw}$. This additional efficiency role implies that the minimum wage may be welfare improving even if the second-best employment tax rate faced by type-1 workers is positive, contrary to Proposition 3 in Lee and Saez (2012).

The two roles of the minimum wage can be understood as an application of the theory of the second best (Lipsey and Lancaster (1956)). An implication of this theory is that if the first-best allocation in an economy is not attainable, otherwise distortionary policies may be welfare improving. As discussed earlier, search frictions and redistributive income taxes imply that the laissez-faire equilibrium allocation may not maximize net output. Equation (14) shows that the minimum wage is welfare improving if it increases low-skilled labor force participation (thereby mitigating externalities due to search frictions), reduces the fiscal cost of redistributive income taxation, or both.

²⁹Conversely, a negative employment tax encourages vacancy creation and participation, which increases the tightness. In this case, higher labor force participation reduces net-tax revenue because of the positive behavioral (macro employment) response. It is also possible to show that beginning from an allocation with no taxes, introducing a minimum wage increases welfare by more than a tax cut (or tax increase) if $\frac{dk_1}{dw} > 0$ ($\beta_1 < \mu_1$).

³⁰In an economy with competitive labor markets, the *Redistribution effect* is positive if the optimal employment tax rate on the low-skilled is negative. Thus, the minimum wage and negative tax rates on the low-skilled are complementary policies. Proponents of the minimum wage argue that the two policies are complementary because a wage floor prevents the incidence effect of negative tax rates on firms (Rothstein (2010)). However, the analysis in this section shows that it is not the wage incidence effects of taxation per-se that make the two policies complements (changes to wages and other prices represent a transfer between firms and workers). Rather, it is the equity-efficiency trade-off generated by redistributive income taxation that distorts allocations that makes a policy tool that reduces this distortion desirable.

The fact that labor supply responses are modeled along the extensive margin explains why the minimum wage plays two distinct roles, unlike the main result in [Hungerbühler and Lehmann \(2009\)](#). They develop a search-and-matching model in which the primary response to income tax and minimum wage changes is along the intensive margin in the spirit of [Mirrlees \(1971\)](#). In their intensive margin model, making the tax system more progressive may induce high ability individuals to mimic the reported wage of low ability individuals. Consequently, at the optimal income tax allocation, high ability workers receive an information rent. [Hungerbühler and Lehmann \(2009\)](#) show that the minimum wage may allow the government to reduce this rent.

It is well-known that the optimal marginal income tax rate at the bottom of the wage distribution is positive when labor supply responses are along intensive margin only ([Mirrlees \(1971\)](#), [Saez \(2002b\)](#), [Hungerbühler and Lehmann \(2009\)](#)). Although a higher marginal tax rate increases progressivity, it also *reduces* the gross wage, increases the number of vacancies, and pushes the labor market tightness upwards ([Hungerbühler, Lehmann, Parmentier, and Van Der Linden \(2006\)](#), [Hungerbühler and Lehmann \(2009\)](#)).³¹ Therefore, the second-best tax system reinforces the upward distortion to the tightness in the low-skilled labor market when workers' bargaining power is too low when labor supply responds along the intensive margin only. Consequently, if the minimum wage improves labor market efficiency it also improves redistribution. However, if labor supply responses are along the extensive or participation margin only, as in this paper, the second-best income tax system distorts the tightness in the low-skilled labor market downwards if $g_1 < 1$. Therefore, the minimum wage can have opposing effects on efficiency and redistribution. For example, the minimum wage reduces (increases) efficiency but improves (reduces) redistribution if $\frac{dk_1}{dw} < 0$ and $g_1 > 1$ ($\frac{dk_1}{dw} > 0$ and $g_1 < 1$).

The *Participation effect* of the minimum wage can be replicated by other policies. For example, a subsidy to type-1 workers that is financed by a payroll tax on firms that hire these workers replicates this effect. This is the idea behind the 'proportional profit tax' in [Boone and Bovenberg \(2002\)](#). The key assumption required for implementing such a tax system is that the government can observe the fixed vacancy cost χ_a .³² In contrast, enforcing minimum wage legislation only requires information on individuals' earnings. This is the same information required for implementing the optimal non-linear income tax.

Given that $k_1 \frac{\partial p_1}{\partial w} = \frac{dh_1}{dw} - p_1 \frac{dk_1}{dw}$ and letting $e_1 = \frac{dk_1}{dw} \frac{w}{k_1}$ and $\eta_1 = \frac{dh_1}{dw} \frac{w}{h_1}$ denote the macro participation and employment elasticities with respect to the minimum wage, I can express the desirability condition in terms of the elasticities that have been the focus of the empirical literature.

Corollary 1. *Introducing a binding minimum wage at the optimal income tax allocation is desirable if and*

³¹This economic intuition is somewhat clouded by the mechanism design set-up of the optimal redistribution problem developed by [Hungerbühler and Lehmann \(2009\)](#). Viewing the problem through the extensive margin versus intensive margin lens highlights the economic forces underlying the desirability of the minimum wage.

³²Specifically, the government must ensure that firms do not claim a wage cost that is $\bar{w}_1 = w_1 - \chi_1$ to reduce their payroll tax liability. Moreover, if a firm is comprised of multiple vacancies for workers with different skill levels, the government must also be able to distinguish between the flow profits from each filled vacancy.

only if

$$e_1 + (\eta_1 - e_1)(1 - g_1) > 0 \quad (15)$$

The desirability condition (15) shows that both the macro participation and employment elasticities with respect to the minimum wage are important for welfare analysis. The participation elasticity is positive only if the policy increases the expected utility of the type-1 workers. It also captures the increase in expected output due to lowering the labor market tightness. Surprisingly, the participation response to the minimum wage has been mostly overlooked by the empirical literature. Below, I provide evidence that recent state and federal minimum wage increases lowered labor force participation for those at the bottom of the wage distribution.

4 Extensions

The static search-and-matching framework facilitates the intuition for why the minimum wage may be part of the policy optimum. This framework, however, relies on several simplifying assumptions. Explicit assumptions include: the lack of interactions or spillovers of the minimum wage on the wages and labor market outcomes of high-skilled workers, the fact that the only labor supply decision is along the extensive margin and that individuals are risk neutral. In this section, I discuss how relaxing these assumptions modifies (14), while retaining the intuition for the role of the minimum wage described earlier.³³

4.1 Wage spillovers

Several recent papers have argued that minimum wage legislation affects the wages and the labor market outcomes of those earning above the prevailing minimum (Autor, Manning, and Smith (2016), Dube, Giuliano, and Leonard (2015), Cengiz, Dube, Linder, and Zipperer (2017)). This section shows how (14) is modified when the baseline model is extended to allow for wage spillovers. In order to focus on the role of spillovers, I continue to assume that job-seekers make labor supply decisions along the extensive margin only. Unlike the baseline model, I allow low-skilled wages to arbitrarily affect the wages of high skilled workers in the absence of wage regulation, and vice versa (i.e. $\frac{\partial w_2}{\partial w_1} \leq 0$ and $\frac{\partial w_1}{\partial w_2} \leq 0$).

As with the model in Sections 2 and 3, I assume that the conditional job finding rate at each skill level, p_a , only depends (directly) on the gross wage w_a and is continuously differentiable. Taxes, the non-employment benefit and wages at other skill levels may affect the job finding rate, but only through changes to the gross wage. Specifically, I assume that the wage-setting mechanism is such that $w_a = w(T_a, b, w_{-a}; a)$ is a continuously differentiable function of T_a , b and w_{-a} . This

³³In Appendix A.1.6, I consider another extension in which I allow for a non-constant returns to scale matching function. Although the empirical literature supports the constant returns to scale approximation (Petrongolo and Pissarides (2001, 2006)), relaxing this assumption has important implications for the desirability of the minimum wage. In this case, I show that the desirability condition depends on three additional sufficient statistics if the matching function is not constant returns to scale. These are the macro participation and employment responses to *taxes* and the effect of taxes on the gross wage.

setup implies that raising (or lowering) taxes on workers at one skill level affects the gross wage and conditional job finding rates of workers at other skill levels. Wage spillovers also lead to labor supply responses at other skill levels. Similarly, changes to the minimum wage will affect the wages, conditional job finding rates, participation decisions and employment outcomes of high-skilled workers.

Proposition 2. *Suppose that taxes and minimum wage changes lead to spillovers at other skill levels. Beginning from the optimal income tax allocation, the change in total welfare from the introduction of a minimum wage is given by*

$$\frac{1}{\lambda} \frac{d\mathcal{L}}{d\underline{w}} \Big|_{\underline{w}=w_1^*} = \Lambda_1 \left[p_1 \frac{dk_1}{d\underline{w}} + k_1 \frac{\partial p_1}{\partial \underline{w}} (1 - g_1) \right] + \Lambda_2 \left[p_2 \frac{dk_2}{d\underline{w}} + k_2 \frac{\partial p_2}{\partial w_2} \frac{\partial w_2}{\partial \underline{w}} (1 - g_2) \right] \quad (16)$$

where $\Lambda_1 = -p_2 \frac{\partial k_2}{\partial T_2} \Big|_{w_1^*}^m h_1 \left[\frac{dh_1}{dT_1} \frac{dh_2}{dT_2} - \frac{dh_1}{dT_2} \frac{dh_2}{dT_1} \right]^{-1} > 0$ and $\Lambda_2 = p_1 \frac{\partial k_1}{\partial \underline{w}} \Big|_{w_1^*}^m h_2 \left[\frac{dh_1}{dT_1} \frac{dh_2}{dT_2} - \frac{dh_1}{dT_2} \frac{dh_2}{dT_1} \right]^{-1} > 0$.

The proof is presented in Appendix A.1.4. Proposition 2 shows that equation (14) is modified by two terms in the presence of spillovers. The third and fourth terms (second set of squared brackets) capture the efficiency and redistributive effects of the minimum wage for high-skilled workers. To isolate the efficiency effect, consider the case of a utilitarian social welfare function. This implies that $g_1 = g_2 = 1$ and (16) reduces to a weighted sum of the labor force participation responses of low- and high-skilled workers.

$$\frac{1}{\lambda} \frac{d\mathcal{L}}{d\underline{w}} \Big|_{\underline{w}=w_1^*} = \Lambda_1 p_1 \frac{dk_1}{d\underline{w}} + \Lambda_2 p_2 \frac{dk_2}{d\underline{w}}$$

The revealed preference intuition described in section 3.4 applies in the presence of wage spillovers. If the minimum wage increases the expected utility of low- and high-skilled workers, labor force participation by both groups will increase. To isolate the redistributive effect, assume that minimum wage legislation has no effect on the expected utility and labor supply of both low- and high-skilled workers (i.e. $\frac{dk_1}{d\underline{w}} = \frac{dk_2}{d\underline{w}} = 0$). In this case, (16) reduces to

$$\frac{1}{\lambda} \frac{d\mathcal{L}}{d\underline{w}} \Big|_{\underline{w}=w_1^*} = \Lambda_1 k_1 \frac{\partial p_1}{\partial \underline{w}} (1 - g_1) + \Lambda_2 k_2 \frac{\partial p_2}{\partial w_2} \frac{\partial w_2}{\partial \underline{w}} (1 - g_2)$$

and the effect of the minimum wage on total welfare depends on the unemployment responses $-\frac{\partial p_1}{\partial \underline{w}}$ and $-\frac{\partial p_2}{\partial w_2} \frac{\partial w_2}{\partial \underline{w}}$ and the welfare weights. The minimum wage is welfare improving only if increasing unemployment improves the equity-efficiency tradeoff of the tax schedule. If the participation and unemployment responses to the minimum wage of high-skilled workers are small, then the third and fourth terms in equation (16) are quantitatively unimportant and (14) is a good approximation for evaluating the welfare effects of minimum wage legislation. In an economy with more than two skill levels, equation (16) is modified only by the extra terms that capture the spillover effects at each skill level.

4.2 Endogenous ability distribution

Appendix A.1.5 shows how (14) is modified when the model is extended to allow individuals to vary their education level and labor market participation. In particular, I assume that searching for a job in the high-skilled labor market requires paying a fixed education cost in addition to the fixed participation cost.³⁴ Participating in the low-skilled labor market only requires paying the fixed participation cost. The labor demand side of the economy, as well as the government's policy tools are identical to the model described earlier. This framework generalizes the model in Gerritsen and Jacobs (2016) in that it allows for search frictions as well as changes in education choices due to the minimum wage.³⁵

In this case, the form of (14) is the same except for an additional term that captures the tax revenue gains or losses due to changes in skill decisions because of the minimum wage. Allowing labor supply to respond along the intensive or education choice margin reduces the welfare gains from introducing a minimum wage for two reasons. First, if the minimum wage raises low-skilled labor force participation, it also induces some individuals that would otherwise have been high-skilled to forgo their investment in education. This distortion reduces tax revenue from high-skilled workers, lowering welfare. Second, if labor supply responses are stronger along the intensive margin than the participation margin, the optimal employment tax rate faced by low-skilled workers is more likely to be positive (Saez (2002b), Kroft, Kucko, Lehmann, and Schmieder (2015)). A positive (negative) employment tax rate for low-skilled workers pushes the tightness in the low-skilled labor market downwards (upwards).³⁶ Since the minimum wage lowers the tightness, a positive employment tax implies that the *Redistribution effect* is negative.

The total welfare gain or loss from the education choice distortion is proportional to the sensitivity of skill choices to policy changes. If the effect of the minimum wage on education choices is small (large), then the welfare cost of distorting skill decisions is quantitatively unimportant (important). In a recent review of the literature, Belman and Wolfson (2014) (chapter 6) conclude that the effect of the minimum wage on human capital decisions is likely small. I also find that minimum wage increases have no statistically significant effect on the probability of graduating from high school for most demographic subgroups (see Appendix Table ??). This suggests that abstracting from the education choice margin in the baseline model described earlier is reasonable for calculating the welfare consequences of minimum wage increases.

³⁴I do not make assumptions about the correlation between the fixed participation and education costs.

³⁵Another important difference between Gerritsen and Jacobs (2016) and the model described here is that the former also consider the welfare consequences of the minimum wage for an arbitrary rationing schedule. I assume that that job rationing due to the minimum wage is uniform and that all workers in the same labor market are identical after having paid their fixed (sunk) participation and education costs.

³⁶If the optimal employment tax rate is negative (positive), redistributive concerns distort the tightness upwards (downwards).

4.3 Fixed (not necessarily optimal) income tax rates

The sufficient statistics approach can also be used to evaluate the welfare gains or losses from increasing the minimum wage beginning from any fixed (not necessarily optimal) tax system. This is policy-relevant because the current federal and state income tax schedules may not be second-best optimal. In the absence of spillovers, the gain in total welfare from increasing a minimum wage is given by: $g_1 h_1 \frac{dk_1}{dw} \left[\frac{\partial k_1}{\partial w} \Big|^{micro} \right]^{-1} + (T_1 + b) \frac{dh_1}{dw}$ for any fixed tax schedule $\{T_1, T_2, b\}$. The first term captures the social welfare effect of a minimum wage increase and has the same sign as the macro participation response $\frac{dk_1}{dw}$. The second term captures the net tax revenue gains or losses due to introducing a minimum wage. Given that the participation and employment responses to the minimum wage are endogenous to the tax system and labor market institutions, they vary across jurisdictions and over time. Consequently, minimum wage legislation may be desirable in one country or state but not another.

4.4 Risk aversion

Allowing for risk aversion does not affect the form of (14). If individuals are risk averse, the allocation that maximizes aggregate output net of fixed participation and vacancy costs is not the same as the allocation that maximizes individuals' expected utilities. Intuitively, individuals care about smoothing consumption between the unemployment and employment states, and are willing to trade-off lower output (and therefore a lower after-tax wage) for a higher job finding rate. In this case, income taxation plays a social insurance role in addition to the redistributive and efficiency roles.³⁷ If the minimum wage increases labor force participation or reduces the efficiency cost of taxation, then it leads to higher net output. Alleviating distortions to the tightness caused by search frictions and redistribution does not (directly) compete with the social insurance role of the tax and transfer system.³⁸

5 Data and empirical strategy

5.1 Data

The sufficient statistics to be estimated are: (1) the macro participation, and (2) the macro employment elasticities with respect to the minimum wage. Annual state-level minimum wage data come from [Dube, Lester, and Reich \(2016\)](#). Individual-level information on labor market outcomes are from the monthly extracts of the Current Population Survey (CPS), downloaded from the Minnesota Population Center web site.³⁹ The advantage of using the CPS is that it provides

³⁷Note, however, that the social insurance role of taxation is a substitute for a missing market for (private) unemployment insurance ([Chetty \(2008\)](#)). In the presence of perfect insurance markets, workers could enter into insurance contracts prior to the realization of their employment outcome.

³⁸Risk aversion does indirectly affect the macro labor force participation response by increasing the utility cost of a lower job finding rate. It also affects the sign of the optimal employment tax liability faced by low-skilled workers.

³⁹[Flood, King, Ruggles, and Warren \(2015\)](#).

information on labor force participation as well as employment. I pool the CPS monthly cross sections for the 1979 to 2014 period (inclusive). In addition to labor market outcomes, the CPS has rich information on individuals' demographic characteristics and (reported) hourly wages; the latter is available only for individuals in the monthly outgoing rotation group (MORG). The information on wages is primarily used to identify individuals who expect to earn near the minimum wage. Throughout the analysis, I use the sample weights provided by the CPS.⁴⁰ All dollar amounts are deflated to 2014 dollars using the BLS Urban CPI.

Each month CPS respondents are asked about their labor force activities for the week preceding the interview, known as the 'reference week'. I classify an individual as employed if they report being at work, or report being employed but temporarily absent from work (i.e. due to an illness) during the reference week. A respondent is classified as being a labor force participant if they report being employed, or if they report being unemployed and actively searching for work.

In the theoretical model, the government's objective is to maximize a social welfare function by setting tax policy and a minimum wage in an economy comprised of single-person households. To better relate the empirical analysis to the theory, I restrict the sample unmarried adults between the ages of 20 and 65 that report not being in school full-time. This restriction excludes teenagers, a group that is the focus of much of the empirical literature on the minimum wage.⁴¹ I exclude this group because most teenagers are not heads of household, and because teenagers are not the intended recipients of tax and transfer policies.⁴²

Among unmarried adults, younger, less-educated workers are more likely to earn the minimum wage than older educated workers (conditional on working). Using information on reported hourly wages, I construct an indicator variable equal to one if an individual in state s and year t reports a wage is within \$0.25 of the prevailing real minimum wage for their state in the following year. Formally, for each wage earner in state s in year t in the MORG sample I construct:

$$\text{MW bite}_{i(st)} = 1\{\text{real wage}_{i(st)} - \text{real MW}_{st+1} \leq 0.25\}$$

Table 1 reports the fraction of workers that satisfy this condition for 16 age-education groups. Unsurprisingly, the minimum wage is most binding for adults between 20 and 29 years of age with less than a high school diploma. Table 1 also shows that the bindingness of the minimum wage monotonically decreases in both age and educational attainment. Using this information, the proxy for type-1 workers in the empirical analysis below is individuals between the ages of 20 and 29 with less than a high school education. As a robustness check, I estimate the effect of the minimum wage on labor force participation and employment using variation in wage bites across age-education groups.

⁴⁰In particular, when estimating the participation and employment responses to the minimum wage, I use the variable *wtfinl* to weight individuals. When estimating the effect of the minimum wage on hourly wages, I use the variable *earnwt*, as recommended by the BLS.

⁴¹Married individuals are also excluded because policy changes may lead to different labor supply responses for dual-earner households.

⁴²Nevertheless, Online Appendix Table B3 report estimates of the effect of the minimum wage on labor force and employment rates of teenagers.

Table 2 reports additional summary statistics for CPS respondents between the ages of 20 and 29 for all education categories. A few important statistics stand out. First, educated workers are significantly more likely to participate in the labor force and be employed. Second, and perhaps more importantly, the unemployment rate is much higher for those with less education. For example, the unemployment rate for high school dropouts is $\frac{0.71-0.56}{0.71} = 0.21$, compared to $\frac{0.86-0.75}{0.86} = 0.13$ for high school graduates, $\frac{0.89-0.82}{0.89} = 0.08$ for those with some college (or an associate’s degree), and $\frac{0.93-0.89}{0.93} = 0.04$ for those with a bachelors degree or higher. This suggests that labor market frictions may be more important for workers with lower levels of education.

5.2 Empirical strategy

5.2.1 Event study design

The macro participation and employment responses to the minimum wage are identified using state and federal minimum wage changes between 1979 and 2014. I assume that labor markets are defined by an individual’s age, education, state and year as in Table ???. Letting i denote individuals and using equation (2), the basic empirical relationship of interest takes the form:

$$k_{i(a,e,s,t)} = \mathcal{K}_{i(a,e,s,t)}(\underline{w}; a, e, s, t)$$

There are two challenges with identifying the causal effect of the minimum wage on labor market outcomes. Both stem from the fact that there is no variation in the minimum wage across individuals in the *same* labor market. The first challenge is that minimum wage increases may be correlated with labor market-specific shocks to labor supply and demand (Dube, Lester, and Reich (2010, 2016)). For example, states tend to increase their minimum wage near the top of the business cycle when employment growth is strong (Baskaya and Rubinstein (2012)). In this case, comparing labor market outcomes of individuals in states that increase their minimum wage relative to those that do not may underestimate the effect of the minimum wage.

The second challenge with identifying the causal effect of the minimum wage on labor market outcomes arises because the ‘treatment status’ of states changes over time (Meer and West (2015)). To illustrate, suppose state A increases its minimum wage in period t and state B increases its minimum wage in period $t + 1$. For simplicity, assume that these are the only two states and that the two states have identical labor market outcomes period to period t . Because state B switches from a “control state” to a “treated state” at period $t + 1$, the effect of the minimum wage is identified by the difference between the labor market outcomes of individuals in state A and state B between t and $t + 1$. If the minimum wage affects the growth rate of outcomes (rather than changes in levels), or if the effect of the minimum wage plays out over time (i.e. labor demand responses may be delayed if firm’s production plans are fixed in the short run), then the difference in outcomes between state A and state B between t and $t + 1$ will understate the effect of the minimum wage.

A solution to these challenges is to adopt an event-study research design around minimum

wage increases (Card and Krueger (1994); Clemens and Wither (2014); Clemens (2015); Monras (2017)). This addresses the issue of states' time-varying treatment status because treated and control units are compared in event time. Event study designs also permit a graphical presentation of the data. This allows researchers to assess the importance of differential pre-trends and delayed responses to the minimum wage. For these reasons, I adopt an event-study research design to estimate the macro participation and employment responses to the minimum wage.⁴³

I pool all federal and state minimum wage events between 1979 and 2014. Over this period, there were 768 state-years in which the statutory *nominal* minimum wage increased from one year to the next. 614 of these events were due to federal minimum wage increases that were binding, and 154 were due to state minimum wage increases. Appendix Figures A1, A2, and A3 report the frequency of minimum wage events by state and year.

The event time variable, $event_{st}$, denotes time (in years) relative to a minimum wage increase for state s . This variable is coded as follows. First, I set $event_{st} = 0$ for all state-years when the prevailing *nominal* state minimum wage increases between $t - 1$ and t . Since many events are due to legislated minimum wage increases that come into effect over two or three years, it is common for $event_{st}$ to be equal to zero for multiple consecutive years.⁴⁴ Second, state-years that are τ years before an event are coded as $event_{st} = -\tau$. Similarly, state-years that are τ years after an event are coded as $event_{st} = \tau$. All state-years more than four years before (resp. four years after) a minimum wage event are coded as $event_{st} = -4$ (resp. $event_{st} = 4$).

Table 3 lists an example of how the variable $event_{st}$ is coded for Wisconsin for the years 1986-1996. Wisconsin increased its minimum wage in 1989 prior to three federal minimum wage increases between 1990 and 1992. As a result, the variable $event_{st}$ is equal to zero for all four years between 1989 and 1992. Table 3 also illustrates how the $event_{st}$ variable 're-sets' at every minimum wage event. While 1993 is coded as $event_{st} = 1$ (following the 1992 federal minimum wage increase), 1994 is coded as $event_{st} = -2$ because it precedes another federal minimum wage increase in 1996.

Figure 1 shows the variation in the $event_{st}$ variable across states in 2003 and 2004. A clear pattern in this figure is that in any given year, states in close proximity to each other are at similar points in event time. For example, states in the Pacific census division (Alaska, California, Hawaii, Oregon and Washington) are all either in event period 0 or 1 in 2003, and in event period 0, 1 or 2 in 2004. These states all experienced at least one event between 2002 and 2004.

5.2.2 Estimating equation

Let $post_{st} = 1\{event_{st} \geq 0\}$. The estimating equation for outcome $y_{i(a,e,s,t)}$ is:

⁴³In Appendix B.3, I replicate the findings in Allegretto, Dube, Reich, and Zipperer (2016) using the log of the real minimum wage as the independent variable of interest. Unlike the estimates reported in Tables 4 and 5, the estimates using this approach are generally small and statistically significant, especially once commonly used geographic controls are used.

⁴⁴I discuss how this affects the interpretation of the graphical and regression results below. Moreover, I show in Appendix Tables B1 and B2 that estimates of the effect of the minimum wage on participation and employment are similar if only the first year of a series of state minimum wage increases is labelled as $event_{st} = 0$.

$$y_{i(a,e,s,t)} = \alpha_0 + \alpha_1 \cdot \text{post}_{st} + \alpha_2 \cdot \text{pre*event}_{st} + X'_{i(a,e,s,t)}\alpha_3 + \delta_s + \delta_t + e_{i(a,e,s,t)} \quad (17)$$

where $\text{pre*event}_{st} = (1 - \text{post}_{st}) \times \text{event}_{st}$, $X_{i(a,e,s,t)}$ is a vector of demographic characteristics: a quartic polynomial in age, and dummy variables for race, gender, and immigrant status. δ_s and δ_t are state and year fixed effects, respectively. The former controls for all time-invariant characteristics that are common to individuals in the same state; the latter for common shocks to the U.S. economy that affect labor market outcomes. The coefficient α_2 is an estimate of the (linear) trend in participation and employment in the years leading up to a minimum wage increase.

The coefficient of interest, α_1 , is an estimate of the macro participation (or employment) response to the minimum wage. This estimate is unbiased if, in the absence of a minimum wage, labor force participation would have evolved following the trend implied by the pre-event years (i.e. the α_2 coefficient). This identification assumption is violated if the *path* of labor market outcomes change beginning in the year of an event for reasons other than minimum wage increases. The estimate of α_1 should be interpreted as a macro or general equilibrium response because it is identified using variation in minimum wage increases across labor markets. It captures the net effect of the minimum wage on outcomes, allowing for changes in the equilibrium job finding rate faced by job-seekers.

The variable pre*event_{st} allows states to have different linear trends before a minimum wage increase. Conceptually, allowing for a (linear) trend in outcomes before an event addresses the concern that minimum wage increases may be correlated with pre-existing employment trends (Dube, Lester, and Reich (2010)). I show in the next section that labor force participation and employment are trending upwards in the years leading up to an event. Moreover, I experiment with a quadratic pre-trend specification and find that the results are quantitatively similar to the baseline linear specification. I also report estimates of the short-run (one-to-two years) and medium-run (three to four years) effects of the minimum wage on participation and employment. Sorkin (2015) provides a theoretical justification for this specification when firms face adjustment costs for labor. Although imprecise, the medium-run estimates are larger in absolute value than the short-run estimates, suggesting that after accounting for pre-trends, labor force participation and employment respond gradually to minimum wage changes.

6 Results

6.1 Graphical evidence

This section presents graphical evidence of the effects of minimum wage legislation on low-skilled labor force participation and employment. I construct figures that show how labor market outcomes evolve before and after a minimum wage increase by plotting the a^j coefficients from the following regression

$$y_{i(a,e,s,t)} = a + \sum_{j=-4}^{-2} a^j \cdot 1\{\text{event}_{st} = j\} + \sum_{j=0}^4 a^j \cdot 1\{\text{event}_{st} = j\} + X'_{i(a,e,s,t)}\lambda + \gamma_s + \gamma_t + u_{i(a,e,s,t)} \quad (18)$$

where $1\{\text{event}_{st} = j\}$ is an indicator variable equal to one if the state-year st is in event time j , and equal to zero otherwise. This specification is a non-parametric version of equation (17).

Figure 2 plots the a^j coefficients when the dependent variable is the real minimum wage. The real minimum wage is stable during the years leading up to an event, and then increases by approximately \$0.68 or 9.1%, on average. Within three years of an event, the real minimum wage in ‘control’ states converges to that in ‘treatment’ states. In particular, the value of the real minimum wage is only \$0.09 higher in treatment states than control states by the third year following an event. This shows that the window over which to estimate minimum wage effects is small.

Figures 3a and 3b plot the a^j coefficients when the dependent variable is labor force participation and employment, respectively. Both outcomes are trending upwards in the years preceding a minimum wage increase. This is especially true for employment. The employment rate of low-skilled adults between the ages of 20 and 29 increases by nearly two percentage points in the three years leading up to a minimum wage event. Relative to a mean of 56 percent over this period (see Table 2), this represents a 3.5 percent increase. Participation and employment drop immediately following an event (from period -1 to 0) by more than a percentage point. Also, the upward trend in participation and employment reverses following a minimum wage increase. This suggests that the minimum wage may have an effect on both the level and growth rate of labor market outcomes.

The trends in participation and employment growth before a minimum wage event support the inclusion of the $\text{pre} \cdot \text{event}_{st}$ variable in equation (17). Figure 4 provides an illustrative example of how the inclusion of this variable affects the estimate of α_1 . The solid blue line plots hypothetical a^j coefficients; the values were chosen so that the figure would resemble the patterns in Figures 3a and 3b. The dashed line plots the counterfactual estimates for the outcome variable that would be observed had the trend in the years preceding the minimum wage increase continued. In any given period, the gap between the two lines is the effect of the minimum wage in that period. The causal effect of the minimum wage on labor market outcomes, α_1 , is the average of these gaps over the post-treatment years.

Figures 5a and 5b plot the a^j coefficients after fitting and removing a linear trend for the years before a minimum wage event.⁴⁵ After removing the pre-trend, the patterns from Figures 3a and 3b become clear. The labor force participation of 20-29 year old high school dropouts falls by a statistically significant 2 percentage points immediately following a minimum wage increase.

⁴⁵The figures are constructed as follows. I first regress the outcome of interest on a linear trend ($\text{pre} \cdot \text{event}_{st}$) and state and year fixed effects for the years preceding a minimum wage event. The fitted values from this regression are saved. For each observation, I subtract the fitted value from the outcome variable to create a new ‘de-trended’ outcome variable. This ‘de-trended’ outcome is used as the dependent variable for the estimation of equation (18) and the construction of Figures 5a and 5b.

Relative to a mean of 70 percent over the sample period, this represents an economically significant 3 percent decline. The participation rate falls by an additional 2 percentage points over the following three years, suggesting an even larger long-run response to the minimum wage. The pattern for employment is similar, except the decline is larger. This is not surprising because the general equilibrium employment response to the minimum wage is the sum of the general equilibrium participation and job finding rate responses. The estimates in the following section quantify the patterns in Figure 5.

6.2 Main estimates

Table 4 reports the coefficient estimates from the estimation of (17). In panel A, the dependent variable is the labor force participation indicator, and in panel B it is employment. The specification in column 1 omits the $pre*event_{st}$ variable. The coefficient on the $post_{st}$ variable implies that the average minimum wage event is associated with a 0.5 percentage point decline in the labor force participation rate of age 20-29 high school dropouts. This estimate is not statistically significant at conventional levels (p-value 0.15). The specification in column 2 adds the $pre*event_{st}$ variable. As suggested by Figure 3a, the coefficient estimate for the former variable is positive; participation is increasing by 0.5 percentage points per-year prior to a minimum wage increase (p-value 0.13). The estimate for α_1 implies that the average minimum wage event leads to a 1.6 percentage point decline in the labor force participation rate.

To better understand the economic magnitude of this decline, I convert the estimate for α_1 into an elasticity as follows. The average labor force participation rate for the sample of adults age 20-29 without a high school diploma is 70 percent. This implies that the average minimum wage increase lowers the labor force participation rate by 1.12 percent ($-0.016/0.7 = -0.012$). Since the average event increases the real minimum wage by \$0.68 or 9.1 percent ($0.091 = 0.68/7.49$), this implies a macro participation elasticity of $e_1 = -0.24$.⁴⁶

The baseline specification in column 2 allows states to have different linear trends before a minimum wage increase. In column 3, I check whether the estimate for α_1 changes if I allow states to have different linear and quadratic trends before a minimum wage increase. The coefficient estimate for α_1 , -0.018, is not statistically different from column 2 (p-value 0.33). This suggests that the linear approximation is reasonable; all remaining robustness checks adopt this specification. Columns 4 and 5 report estimates of a specification that augments (17) by checking whether the short and medium run effects of the minimum wage differ. In particular, these specifications replace the $post_{st}$ variable with $post*short_{st}$ and $post*medium_{st}$ variables. The variable $post*short_{st}$ is a dummy variable equal to one for the year of and the year following the most recent minimum wage increase (i.e. $event_{st} = 0$ or $event_{st} = 1$), and zero otherwise. The variable $post*medium_{st}$ is a dummy variable equal to one for the third and fourth years following a minimum wage increase

⁴⁶This elasticity is calculated using the *nlcom* command in Stata. This allows me to calculate standard errors for the participation and employment elasticities. The standard error for the participation elasticity is 0.093; the 95% confidence interval is (-0.422, -0.058).

(i.e. $event_{st} = 2$ or $event_{st} = 3$). In column 5, the estimates of $post*short_{st}$ and $post*medium_{st}$ suggest that the medium-run participation response to the minimum wage is more negative than the short-run response (-0.017 versus -0.015). However, the difference between these estimates is not statistically significant (p-value 0.85).

The dependent variable in panel B is the employment indicator. In the specification that excludes the $pre*event_{st}$ variable, the estimate for α_1 implies that the minimum wage has no effect on employment (column 1). Unsurprisingly, the estimate for α_1 decreases to a statistically significant -0.015 in the baseline specification (p-value 0.06). The macro employment elasticity implied by this estimate is $\eta_1 = -0.32$ (p-value 0.03).⁴⁷

The estimate for α_2 is 0.008 and is statistically significant at the 5 percent level. Similar to evidence in [Baskaya and Rubinstein \(2012\)](#) and [Monras \(2017\)](#), this implies that, prior to an event, employment is growing by almost one percentage point per year faster in states that experience a minimum wage increase than states that do not. Similar to panel A, allowing states to have different quadratic employment trends does not have a statistically significant effect on the estimate for α_1 .⁴⁸ Estimates of the short- and medium-run effects of the minimum wage are similar.

Appendix Table [A1](#) reports specifications identical to those above when the dependent variable is the (real) hourly wage reported by CPS MORG respondents. Minimum wage events increase the hourly wage of young high school dropouts by a \$0.22, on average (implied elasticity of 0.25). Together with [Figures 3-5](#) and [Table 4](#), this suggests that minimum wage events are binding on the wage distribution of unmarried adults 20-29 without a high school diploma. Furthermore, labor force participation and employment decline following minimum wage increases; the decline in the former implies that the expected utility of this group falls.

6.3 Robustness

[Table 5](#) reports estimates for α_1 when equation (17) is augmented to include commonly used geographic controls. In Panel A, the dependent variable is the labor force participation indicator; in Panel B, the dependent variable is the employment indicator. Column 1 reports the baseline estimates from [Table 4](#) as a benchmark. The specification in column 2 adds state-specific linear time trends to equation (17). The estimate for the macro participation response increases slightly in magnitude from -0.016 to -0.012, however, this difference is not statistically significant (p-value 0.37). In column 3, state-specific linear and quadratic time trends are added to equation (17). The estimate for α_1 increases slightly but is less precise. I am unable to reject the null hypothesis that the estimate in column 3 is statistically different from the baseline estimate in column 1. Thus, after allowing states to have different (linear) trends in the years leading up to a minimum wage increase, adding state specific time trends does not significantly affect the estimate for the macro participation response. This contrasts sharply with standard two-way fixed effects regres-

⁴⁷The 95% confidence interval for the macro employment elasticity is (-0.601, -0.041).

⁴⁸The difference between the estimates for α_1 in columns 2 and 3 of panel B, [Table 4](#) is 0.006 with a standard error of 0.010 (p-value 0.55).

sion models, where the inclusion of state specific time trends typically leads to minimum wage effects that are small in magnitude and statistically insignificant (Allegretto, Dube, , Reich, and Zipperer (2016)).

Columns 4 and 5 augment equation (17) by including census region by year and census division by year fixed effects, respectively. In these specifications, minimum wage effects are identified by comparing states that experience a minimum wage increase to other states in the same census region (or division). Relative to the baseline estimates in column 1, adding census region-by-year fixed effects leads to a less negative macro participation response. The estimate for α_1 implies that the average minimum wage event lowers participation by -0.7 percentage points. However, precision falls significantly and I cannot reject the null hypothesis that this estimate is the same as the baseline estimate in column 1. The same is true for the specification in column 5.

In Panel B, the dependent variable is employment. The patterns are largely similar to those in panel A. The one exception is that the inclusion of census region-by-time and census division-by-time fixed effects significantly affects the estimate for α_1 . For example, in column 4, the coefficient estimate falls (in absolute value) from -0.015 to -0.000 (standard error (0.008)).

Comparing the labor market outcomes of individuals in neighboring states could lead to biased estimates of the macro participation and employment responses if the migration decisions of low-skilled workers respond to minimum wage legislation. Consistent with evidence in Cadena (2014) and Monras (2017), I find that immigrants avoid states that experience minimum wage increases. In particular, the fraction of age 20-29 high school dropouts in a state that are immigrants declines by 2.2 percentage points following a minimum wage increase (see Appendix Table A2).⁴⁹ Relative to a sample mean of 59 percent, this represents a statistically significant 3.7 percent decrease. However, the migration response to the minimum wage is even stronger when comparing states within the same census division; minimum wage increases lead to a 4.8 percentage point or 8.1 percent decline in the fraction of immigrants (see Appendix Table A3).

The decline in the fraction of immigrants affects the interpretation of the macro participation and employment responses from specifications with census region or census division-by-year fixed effects. In particular, a large inflow of low-skilled workers from a state that increases its minimum wage may lower the job finding rate and depress labor market outcomes in a neighboring state, at least in the short run. This general equilibrium response attenuates macro estimates of the labor force participation and employment responses to the minimum wage.

Comparing groups of individuals in the same state that are differentially affected by minimum wage increases is another way to check whether state-specific economic shocks are driving the negative baseline estimates in Table 4. Table 6 reports estimates from a regression that augments equation (17) by comparing individuals in different age-education groups within the same state that are differentially affected by minimum wage increases. For example, nearly 17 percent of unmarried high school dropouts age 20-29 earn within \$0.25 of the minimum wage, compared

⁴⁹The participation and employment rates of low-skilled immigrants are higher than those of low-skilled natives (see Appendix Table A3). This implies that the migration response to the minimum wage explains approximately 9 percent of the macro participation and 14 percent of the macro employment responses to the minimum wage, respectively.

to fewer than 3 percent of college graduates over the age of 50 (Table 1). A similar identification strategy is used by Clemens (2015) to study the employment effects of federal minimum wage increases during the Great Recession.⁵⁰

Let g denote an age-education group (high school dropouts age 20-29, high school graduates age 20-29, ..., college graduates age 50+), and Bite_g denote the fraction of workers (across all states and years) in group g that earn within \$0.25 of the prevailing minimum wage. The estimating equation for this difference-in-difference-differences (DDD) strategy is

$$y_{a,e,s,t} = \tilde{\alpha}_0 + \tilde{\alpha}_1 \cdot \text{Bite}_g \times \text{post}_{st} + \tilde{\alpha}_2 \cdot \text{Bite}_g \times \text{pre*event}_{st} + \delta_{st} + \delta_{sg} + \delta_{tg} + e_{a,e,s,t} \quad (19)$$

where δ_{st} , δ_{sg} , and δ_{tg} denote the state-year, state-group, and year by group fixed effects, respectively. The coefficient $\tilde{\alpha}_1$ captures the (average) differential effect of a minimum wage increase on the outcomes of individuals in different age-education groups. It will be negative if the minimum wage lowers labor force participation (or employment) more for individuals in age-education groups where the minimum wage is more binding. The $\text{Bite}_g \times \text{pre*event}_{st}$ variable allows labor force participation and employment to vary across age-education groups before an event. For computational ease, I collapse the data to age-group-education group-state-year means before estimating equation (19).⁵¹

Column 1 of Table 6 reports the coefficient estimate for $\tilde{\alpha}_1$ for a simplified version of equation (19) that omits the $\text{Bite}_g \times \text{pre*event}_{st}$ variable. Similar to Column 1 of Table 4, the estimate for $\tilde{\alpha}_1$ from this specification suggests that the minimum wage leads to a small, statistically insignificant decline in the labor force participation rate. After including these variables, the estimate for $\tilde{\alpha}_1$ falls to -0.08 and is statistically significant at the one percent level (column 2). This estimate implies that increasing the fraction of workers in an age-education group that work at the minimum wage from zero to 100 percent will lower the labor force participation rate for that group by 8 percentage points. The estimate for $\tilde{\alpha}_1$ cannot be directly compared with the baseline estimate for α_1 in Table 4 because the minimum wage is not binding on 100 percent on unmarried dropouts age 20-29. However, using the variation in the ‘bite’ of the minimum wage in the data, I can re-scale the estimate in column 2 of Table 6 as follows. From Table 1, the minimum wage is most binding for unmarried dropouts between the ages of 20-29, and least binding for college graduates age 40-49. The difference in the minimum wage ‘bite’ between these groups is 14 percentage points ($0.166 - 0.023 = 0.143$). Multiplying 0.143 by -0.080 yields -0.011, close to the estimate for α_1 from column 2, Table 4. The patterns for the effect of the minimum wage on employment is similar (panel B).

Table 7 and Appendix Tables A3-A8 explore additional robustness checks. I briefly summarize these here. In Table 7, I interact the post_{st} variable with the magnitude of an event, ‘min wage

⁵⁰Estimates of the macro participation and employment elasticities are qualitatively similar when using the 2007-2009 federal minimum wage changes as policy variation and correcting for pre-trends (equation (17)). These estimates are available upon request.

⁵¹In the regressions, each cell is weighted by the number of CPS respondents in that cell.

increase $_{st}'$. Minimum wages are increased the most in jurisdictions with initially low wage floors (column 1). Moreover, larger minimum wage increases have more negative effects on low-skilled labor force participation (column 2) and employment (column 3). Minimum wages do not have a statistically significant effect on the labor market outcomes of higher skilled workers (Appendix Table A4), suggesting that any spillover effects are small in magnitude. Also, the minimum wage has more negative effects on the participation and employment of women and immigrants. This is not surprising because these groups are more likely to be minimum wage workers. More recent (post-1997) minimum wage changes had a less negative impact, although the fraction of the population with less than a high school diploma has been steadily declining over time. Finally, the baseline estimates in Table 4 are robust to not using the sample weights and to restricting the sample to the MORG respondents for whom information on hourly wages is available. Based on the evidence in Tables 4-7 and the appendices, I conclude that the estimates for the macro labor force participation response to the minimum wage are reasonably robust to the inclusion of geographic controls. Estimates of the macro employment response are less robust.

7 Numerical simulations

Numerically implementing the sufficient statistics desirability condition for the minimum wage (equation (15)) requires knowledge of three parameters. The first is the macro labor force participation elasticity e_1 . This parameter captures the *Participation effect* – the change in total welfare due to labor force entry (or exit) due to the minimum wage. The second parameter is the change in the job finding rate due to the minimum wage $\eta_1 - e_1$. The third parameter is one minus the government’s marginal social welfare weight for low-skilled workers, $1 - g_1$. Multiplying the second and third parameters gives the *Redistribution effect*. Using the estimated macro participation and employment elasticities from the previous section, I calculate the marginal gain in total welfare from introducing a minimum wage as a function of g_1 ⁵².

Figure 6 plots the gain in total welfare from introducing a minimum wage at the optimal income tax allocation. The vertical axis measures the change in total welfare per individual in units of the consumption good. The horizontal axis measures the government’s welfare weight for low-skilled workers, g_1 . The dashed vertical line corresponds to the case where the social welfare function is utilitarian. Values to the right of the vertical line represent stronger redistributive tastes: values to the left represent regressive tastes. Figure 6 shows that the minimum wage is only desirable for values of g_1 greater than 4. In other words, the government would have to be indifferent between giving \$1 to type-1 workers and \$4 of public funds (to be distributed to other agents) in order for the minimum wage to be desirable.

There are two reasons why the estimates from the previous section imply that redistributive tastes must be very strong to justify the minimum wage. The first is because the minimum wage

⁵²Figure A6 in the Appendix shows the results from a second policy experiment: calculating the welfare gains from introducing a minimum wage beginning from any fixed income tax system.

reduces labor force participation and, therefore, output net of participation and vacancy costs. The second reason is because the reduction in the job finding rate, and the labor market tightness, implied by the difference $\eta_1 - e_1$ is relatively small. Consequently, the second-best correction is also small.

It is important to remember that the desirability condition (15) assumes that: workers of a similar ability are homogenous after paying the fixed participation cost θ , and that small minimum wage changes do not lead to spillovers on the wages (or job finding rates) of high-skilled workers. Relaxing either of these assumptions would affect the desirability of the minimum wage for any (e_1, η_1, g_1) triple. For example, if spillovers increased the expected utility of high-skilled workers (net of fiscal effects), then the curve plotting the total welfare gain due to the minimum wage would shift upwards. While I find that the minimum wage raises the hourly wages of workers higher in the skill distribution, its effect on labor force participation and employment are small and statistically insignificant (Appendix Table A4). Proposition 2 shows that only the latter are relevant for welfare analysis, suggesting that using equation (15) is a reasonable approximation of the welfare effect of the minimum wage.

Figures 7a and 7b show how the desirability of the minimum wage varies with the macro participation and employment elasticities, respectively. Figure 7a shows how the total welfare gain due to the minimum wage is affected by varying the macro participation elasticity. Importantly, the employment elasticity is held constant. A larger macro participation response increases both the *Participation effect* and the *Redistribution effect* of the minimum wage. The *Redistribution effect* increases because a more positive participation response implies a larger decline in the tightness (and therefore job finding rate).

Figure 7b shows how the total welfare gain due to the minimum wage varies with the macro employment elasticity. Holding the participation elasticity constant, more negative values of η_1 actually increase the desirability of the minimum wage. The intuition for this somewhat surprising result is as follows. Holding e_1 constant implies that the expected utility of low-skilled workers is unchanged. A more negative employment response implies that the decline in the labor market tightness ($\frac{\partial p_1}{\partial w}$) is larger, increasing the *Redistribution effect*. In other words, the more negative employment elasticity (holding the participation elasticity constant and assuming $g_1 > 1$) the more the minimum wage alleviates the distortion induced by redistributive income taxation. This increases the resources available for redistribution. For values of $g_1 < 1$, reducing the labor market tightness lowers welfare so a more negative employment elasticity does not increase the desirability of the minimum wage.

8 Conclusion

This paper contributes to the normative and empirical literatures on the minimum wage. I derive a condition under which the minimum wage complements optimal income taxation in terms of three sufficient statistics. This condition highlights two distinct roles for the minimum wage.

First, the minimum wage acts as a corrective tax on the tightness in the low-skilled labor market. This is the case if the minimum wage increases the labor force participation rate of low-skilled individuals. Second, wage regulation alleviates the distortion to the labor market tightness caused by redistributive income taxation. My theoretical analysis shows that both the macro labor force participation and employment elasticities with respect to the minimum wage are important for welfare analysis. Surprisingly, the former has mostly been overlooked by the empirical literature.

I also provide new evidence on the labor force participation response to the minimum wage using U.S. state and federal minimum wage variation from 1979-2014. Using an event study design, I estimate a macro participation elasticity of -0.24. This implies that the average minimum wage increase lowered the net output and the expected utility of low-skilled workers. My baseline estimate of the macro employment elasticity is -0.32. These estimates imply that approximately two-thirds of the decline in employment caused by minimum wage increases is due to lower labor force participation, and the remaining third is due to a lower equilibrium job finding rate.

My results have important implications for the current debate about the desirability of minimum wage legislation. First, evaluating the effect of minimum wage increases on the well-being (expected utility) of low-skilled individuals should be determined by estimating the macro labor force participation response. Furthermore, using estimates of the macro employment response significantly understates the effect of the minimum wage on well-being. For example, my baseline elasticity estimates imply that using the macro employment response to approximate the change in expected utility of U.S. low-skilled adults would overstate the negative impact of the minimum wage by approximately one third.

Second, the approach in this paper can be extended to incorporate other arguments in the debate about the desirability of the minimum wage. For example, in the model described earlier the minimum wage may be welfare-improving even though it increases unemployment among low-skilled individuals. Agents in the model only care about the pecuniary benefits and costs of working and being unemployed. In reality, however, unemployment may have non-pecuniary costs or decrease the likelihood of finding a job in the future. Incorporating these costs is important for understanding the welfare effects of wage regulation.

Opponents of minimum wage legislation also argue that it targets low-income households poorly compared to other policy instruments. This is because many minimum wage workers do not belong to low-income households. Therefore, studying the conditions under which the minimum wage is desirable with multi-person households in which some low-wage workers belong to affluent households is another interesting extension. I leave this question for future research.

References

AHN, T., P. ARCIDIACONO, AND W. WESSELS (2011): "The Distributional Impacts of Minimum Wage Increases When Both Labor Supply and Labor Demand are Endogenous," *Journal of Business and Economic Statistics*, 29(1).

- ALLEGRETTO, S., A. DUBE, , M. REICH, AND B. ZIPPERER (2016): "Credible Research Designs for Minimum Wage Studies: A Response to Neumark, Salas and Wascher," *Industrial and Labor Relations Review*, forthcoming.
- ALLEN, S. P. (1987): "Taxes, redistribution, and the minimum wage: a theoretical analysis," *Quarterly Journal of Economics*, 102(3), 477–490.
- AUTOR, D., A. MANNING, AND C. L. SMITH (2016): "The Contribution of the Minimum Wage to U.S. Wage Inequality over Three Decades: A Reassessment," *American Economic Journal: Applied Economics*, 8(1), 58–99.
- BASKAYA, Y. S., AND Y. RUBINSTEIN (2012): "Using Federal Minimum Wages to Identify the Impact of Minimum Wages on Employment and Earnings across the U.S. States," .
- BELMAN, D., AND P. J. WOLFSON (2014): *What Does the Minimum Wage Do?* W.E. Upjohn Institute for Employment Research, Kalamazoo, MI.
- BOADWAY, R., AND K. CUFF (2001): "A minimum wage can be welfare-improving and employment-enhancing," *European Economic Review*, 45(3), 553–576.
- BOONE, J., AND L. BOVENBERG (2002): "Optimal labour taxation and search," *Journal of Public Economics*, 85, 53–97.
- CADENA, B. C. (2014): "Recent Immigrants as Labor Market Arbitrageurs: Evidence from the Minimum Wage," *Journal of Urban Economics*, 80, 1–12.
- CARD, D., AND A. KRUEGER (1994): "Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania," *American Economic Review*, 84(4), 772–793.
- CENGIZ, D., A. DUBE, A. LINDER, AND B. ZIPPERER (2017): "The effect of minimum wages on the total number of jobs: Evidence from the United States using a bunching estimator," .
- CHETTY, R. (2008): "Moral Hazard vs. Liquidity and Optimal Unemployment Insurance," *Journal of Political Economy*, 116(2), 174–234.
- (2009): "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 1, 451–488.
- CLEMENS, J. (2015): "The Minimum Wage and the Great Recession: Evidence from the Current Population Survey," Working Paper 21830, National Bureau of Economic Research.
- CLEMENS, J., AND M. WITHER (2014): "The Minimum Wage and the Great Recession: Evidence of Effects on the Employment and Income Trajectories of Low-Skilled Workers," Working Paper 20724, National Bureau of Economic Research.
- DIAMOND, P. A. (1982): "Wage Determination and Efficiency in Search Equilibrium," *Review of Economic Studies*, 49(2), 217–227.

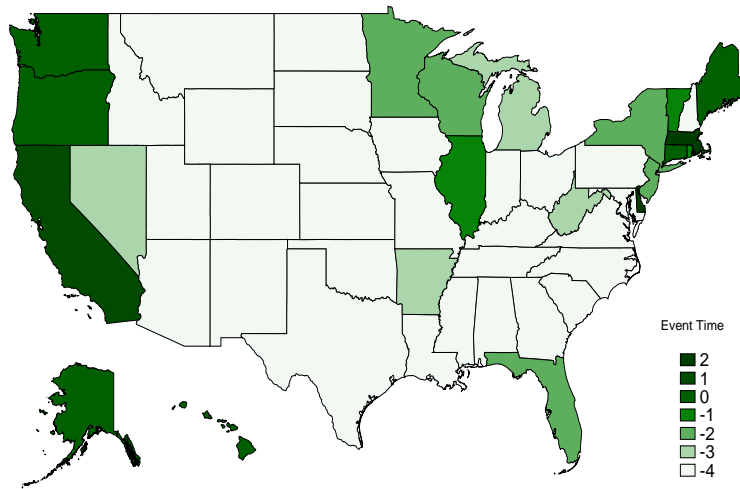
- DUBE, A., L. GIULIANO, AND J. LEONARD (2015): "Fairness and Frictions: The Impact of Unequal Raises on Quit Behavior," Discussion Paper Series 9149, Institute for the Study of Labor (IZA).
- DUBE, A., T. W. LESTER, AND M. REICH (2010): "Minimum Wage Shocks, Employment Flows and Labor Market Frictions," *Review of Economics and Statistics*, 92(4), 945–964.
- (2016): "Minimum Wage Shocks, Employment Flows and Labor Market Frictions," *Journal of Labor Economics*, 34(3), 663–704.
- FLINN, C. J. (2006): "Minimum wage effects on labor market outcomes under search, bargaining and endogenous contract rates," *Econometrica*, 74(4), 1013–1062.
- FLOOD, S., M. KING, S. RUGGLES, AND J. R. WARREN (2015): "Integrated Public Use Microdata Series, Current Population Survey: Version 4.0," [Machine-readable database]. Minneapolis: University of Minnesota.
- GERRITSEN, A., AND B. JACOBS (2016): "Is the Minimum Wage an Appropriate Tool for Redistribution," .
- GOULD, E. (2014): "Why America's Workers Need Faster Wage Growth – And What We Can Do About It," Briefing Paper 382, Economic Policy Institute.
- GRAVEL, F. (2015): "Participation, Recruitment Selection, and the Minimum Wage," *Scandinavian Journal of Economics*, 117(4), 1281–1305.
- GUESNERIE, R., AND K. ROBERTS (1987): "Minimum wage legislation as a second best policy," *European Economic Review*, 31(1-2), 490–498.
- HOSIOS, A. J. (1990): "On the efficiency of matching and related models of search and unemployment," *Review of Economic Studies*, 57(2), 279–298.
- HUNGERBÜHLER, M., AND E. LEHMANN (2009): "On the optimality of a minimum wage: New insights from optimal tax theory," *Journal of Public Economics*, 93(3-4), 464–481.
- HUNGERBÜHLER, M., E. LEHMANN, A. PARMENTIER, AND B. VAN DER LINDEN (2006): "Optimal Redistributive Taxation in a Search Equilibrium Model," *Review of Economic Studies*, 73(3), 743–767.
- JACQUET, L., E. LEHMANN, AND B. VAN DER LINDEN (2014): "Optimal income taxation with Kalai wage bargaining and endogenous participation," *Social Choice Welfare*, 42(2), 381–402.
- KAPLOW, L. (2012): "Optimal Control of Externalities in the Presence of Income Taxation," *International Economic Review*, 53(2), 487–509.
- KOPCZUK, W., E. SAEZ, AND J. SONG (2010): "Earnings Inequality and Mobility in the United States: Evidence from Social Security Data Since 1937," *Quarterly Journal of Economics*, 125(1), 91–128.

- KROFT, K., K. KUCKO, E. LEHMANN, AND J. SCHMIEDER (2015): "Optimal Income Taxation with Unemployment and Wage Responses: A Sufficient Statistics Approach," Working Paper 21757, National Bureau of Economic Research.
- LANDAIS, C., P. MICHAILLAT, AND E. SAEZ (2015): "A Macroeconomic Theory of Optimal Unemployment Insurance," Working Paper 16526, National Bureau of Economic Research.
- (2016): "A Macroeconomic Theory of Optimal Unemployment Insurance," *American Economic Journal: Economic Policy*, forthcoming.
- LANG, K., AND S. KAHN (1998): "The effect of minimum-wage laws on the distribution of employment: theory and evidence," *Journal of Public Economics*, 69(1), 67–82.
- LEE, D., AND E. SAEZ (2008): "Optimal minimum wage policy in competitive labor markets," Working Paper 14320, National Bureau of Economic Research.
- (2012): "Optimal minimum wage policy in competitive labor markets," *Journal of Public Economics*, 96(9-10), 739–749.
- LEHMANN, E., A. PARMENTIER, AND B. VAN DER LINDEN (2011): "Optimal income taxation with endogenous participation and search unemployment," *Journal of Public Economics*, 95(11-12), 1523–1537.
- LIPSEY, R., AND K. LANCASTER (1956): "The General Theory of Second Best," *Review of Economic Studies*, 24(1), 11–32.
- LUNA-ALPIZAR, J. L. (2015): "Worker Heterogeneity and the Asymmetric Effects of Minimum Wages," Job Market Paper, UC Irvine.
- MANNING, A. (2004): "Monopsony and the efficiency of labour market interventions," *Labour Economics*, 11(2), 145–163.
- MARCEAU, N., AND R. BOADWAY (1994): "Minimum Wage Legislation and Unemployment Insurance as Instruments for Redistribution," *Scandinavian Journal of Economics*, 96(1), 67–81.
- MEER, J., AND J. WEST (2015): "Effects of the Minimum Wage on Employment Dynamics," Working Paper 19162, National Bureau of Economic Research.
- MINCER, J. (1976): "Unemployment Effects of Minimum Wages," *Journal of Political Economy*, 84(4), S87–S104.
- MIRRLIENS, J. A. (1971): "An exploration in the theory of optimum income taxation," *Review of Economic Studies*, 38(2), 175–208.
- MONRAS, J. (2017): "Minimum Wages and Spatial Equilibrium," *Journal of Labor Economics*, forthcoming.

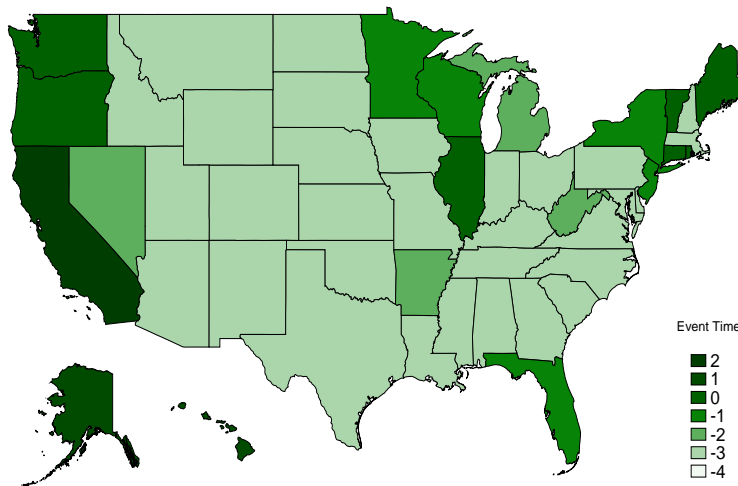
- MORTENSEN, D. T., AND C. A. PISSARIDES (1999): "New developments in models of search in the labor market," in *Handbook of Labor Economics*, ed. by O. C. Ashenfelter, and D. Card, vol. 3, chap. 39, pp. 2567–2627. Elsevier.
- NEUMARK, D. (2016): "Policy levers to increase jobs and increase income from work after the Great Recession," *IZA Journal of Labor Policy*, 5(6).
- NEUMARK, D., I. I. SALAS, AND W. WASCHER (2014a): "Revisiting the Minimum Wage-Employment Debate: Throwing Out the Baby with the Bathwater," *Industrial and Labor Relations Review*, 67(2.5), 608–648.
- NEUMARK, D., J. I. SALAS, AND W. WASCHER (2014b): "More on Recent Evidence on the Effects of Minimum Wages in the United States," Working Paper 20619, National Bureau of Economic Research.
- NEUMARK, D., AND W. WASCHER (1992): "Evidence on Employment Effects of Minimum and Subminimum Wage: Panel Data on State Minimum Laws," *Industrial and Labor Relations Review*, 46(1), 55–81.
- PETRONGOLO, B., AND C. A. PISSARIDES (2001): "Looking into the black box: a survey of the matching function," *Journal of Economic Literature*, 38, 391–431.
- (2006): "Scale effects in markets with search," *Economic Journal*, 116, 21–44.
- ROTHSTEIN, J. (2010): "Is the EITC as good as the NIT? Conditional Cash Transfers and Tax Incidence," *American Economic Journal: Economic Policy*, 2(1), 177–208.
- SAEZ, E. (2001): "Using elasticities to derive optimal income tax rates," *Review of Economic Studies*, 68(1), 205–229.
- (2002a): "The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes," *Journal of Public Economics*, 83, 217–230.
- (2002b): "Optimal income transfer programs: Intensive versus extensive labor supply responses," *Quarterly Journal of Economics*, 117(3), 1039–1073.
- SORKIN, I. (2015): "Are there long-run effects of the minimum wage?," *Review of Economic Dynamics*, 18(2), 306–333.
- SWINNERTON, K. A. (1996): "Minimum Wages in an Equilibrium Search Model with Diminishing Returns to Labor in Production," *Journal of Labor Economics*, 14(2), 340–355.
- WESSELS, W. J. (2005): "Does the Minimum Wage Drive Teenagers Out of the Labor Force?," *Journal of Labor Research*, 26(1), 169–176.
- WHITTEN, S. (2015): "16 states to raise minimum wage in 2016," Online article, CNBC, <http://www.cnbc.com/2015/12/24/13-states-to-raise-minimum-wage-in-2016.html>.

Figure 1: Event Time Variation Across States

(a) 2003

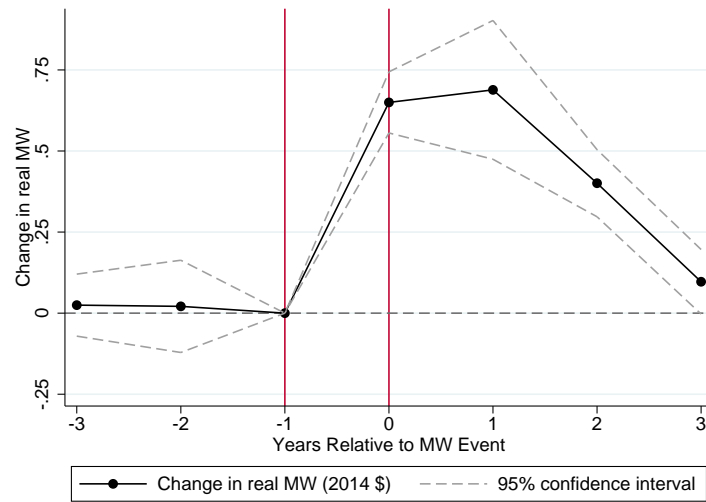


(b) 2004



Notes: This figure shows the spatial variation of the event time variable for 2003 and 2004. Lighter shades correspond to years before a minimum wage event, and darker shades correspond to years after an event.

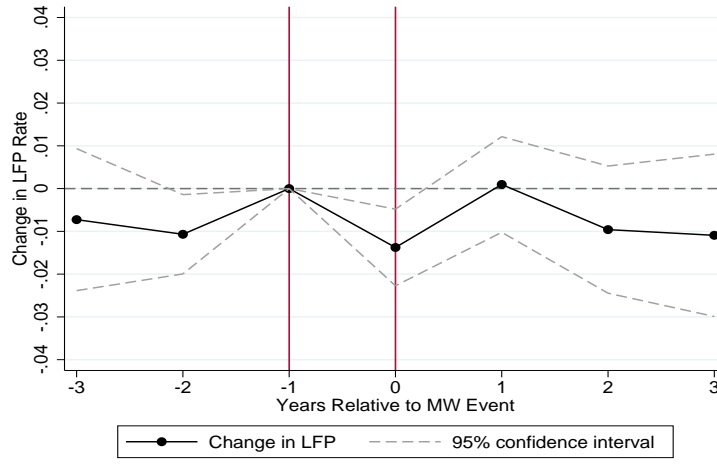
Figure 2: Evolution of the Real Minimum Wage



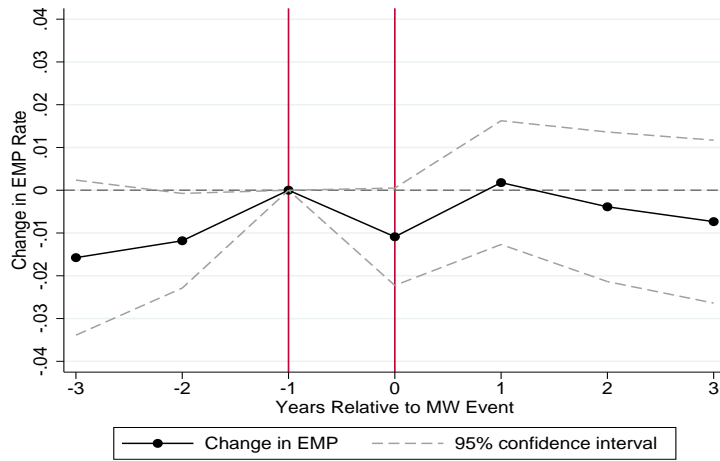
Notes: This figure reports estimates of the effect of nominal minimum wage increases on the real minimum wage. The vertical lines indicate the year before a minimum wage increase and the year of a minimum wage increase, respectively. Each marker is an estimate of the a^j coefficient from equation (18). All dollar amounts are deflated to 2014 dollars using the BLS Urban CPI.

Figure 3: The Effect of the Minimum Wage on Labor Market Outcomes

(a) Labor Force Participation

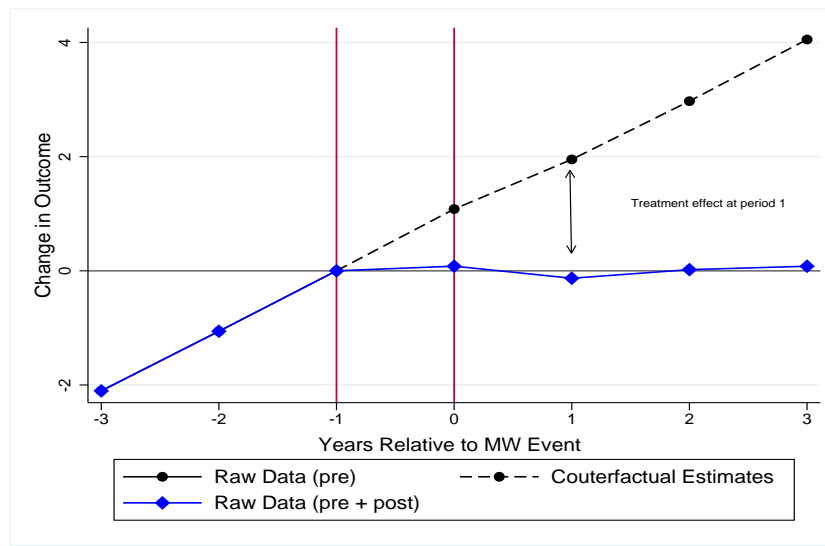


(b) Employment



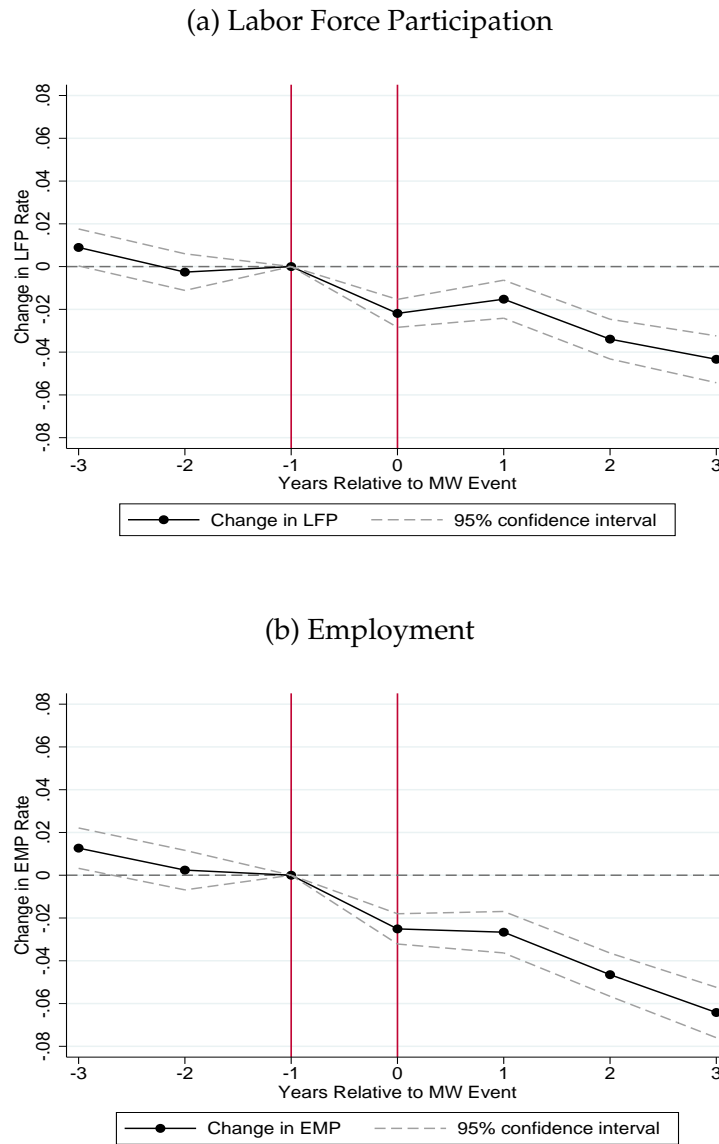
Notes: This figure reports estimates of the effect of nominal minimum wage increases on the labor market outcomes of age 20-29 unmarried high school dropouts. In Figure 3a, the dependent variable is a labor force participation status indicator; in Figure 3b, the dependent variable is an employment status indicator. The vertical lines indicate the year before a minimum wage increase and the year of a minimum wage increase, respectively. Each marker is an estimate of the a^j coefficient from equation (18).

Figure 4: Intuition for Allowing States to Have Different Pre and Post-Event Trends



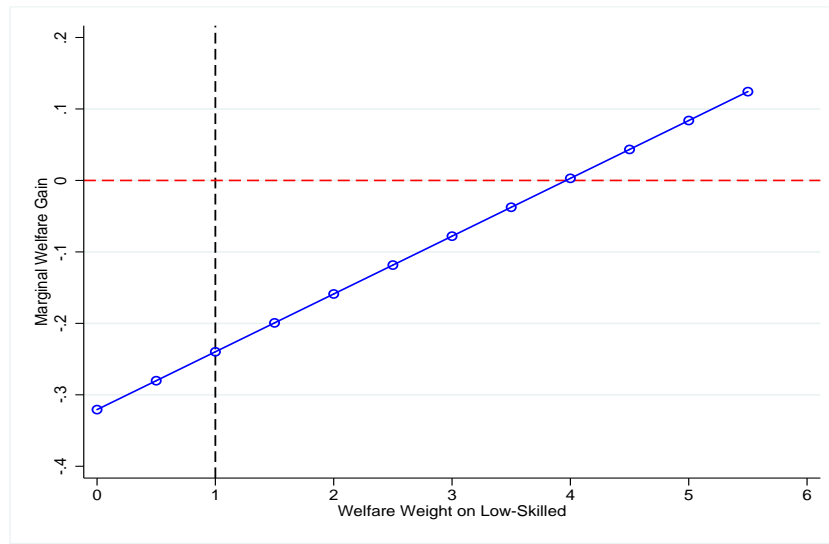
Notes: This figure illustrates the intuition for the inclusion of the $event_{st}$ and $event_{st} \times post_{st}$ variables in equation (17). The blue solid line and blue markers represent hypothetical realization of the data for the case where the outcome variable is trending upwards in states that experience a minimum wage increase. The dashed line corresponds to the counterfactual estimate for the evolution of the outcome variable under the assumption that the pre-existing trend continued.

Figure 5: De-trended Participation and Employment Responses to Minimum Wage Increases



Notes: This figure reports de-trended estimates of the effect of nominal minimum wage increases on the labor market outcomes of age 20-29 unmarried high school dropouts. In Figure 5a, the dependent variable is a labor force participation status indicator; in Figure 5b, the dependent variable is an employment status indicator. The vertical lines indicate the year before a minimum wage increase and the year of a minimum wage increase, respectively. Each marker is an estimate of the a^j coefficient from equation (18) after fitting and removing a linear trend for the years before a minimum wage event.

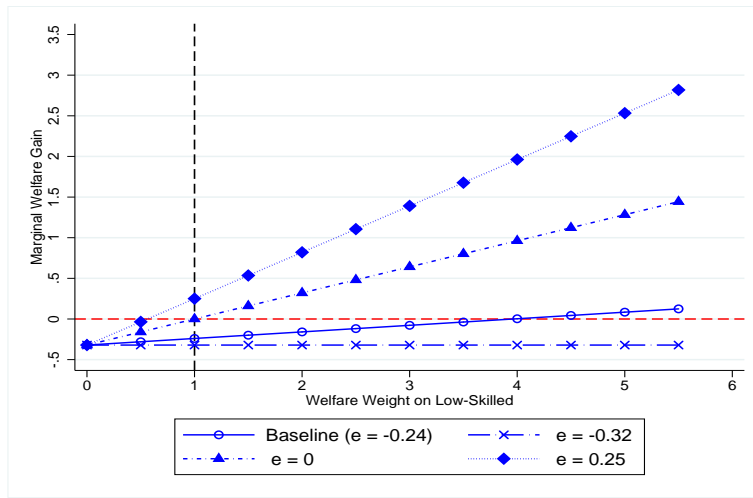
Figure 6: Marginal Welfare Gains from Introducing a Minimum Wage



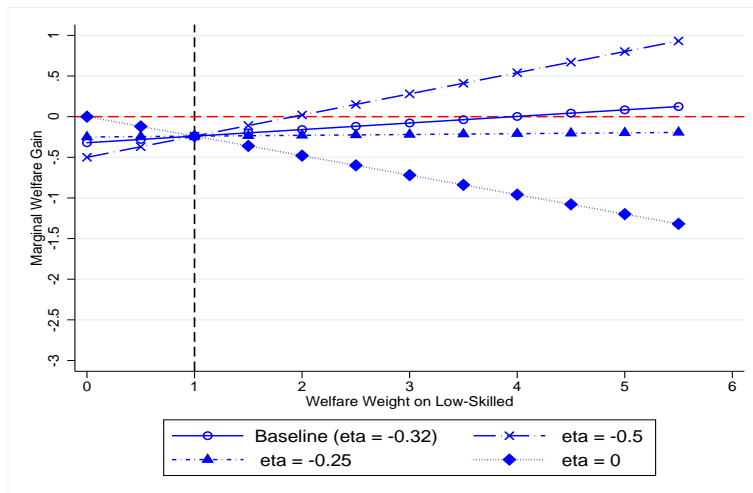
Notes: This figure plots the marginal gain in total welfare from introducing a minimum wage at the optimal income tax allocation using the baseline estimates for the macro participation and employment elasticities. The vertical axis measures the gain in total welfare per individual in units of the consumption good. The horizontal axis measures the government's marginal social welfare weight for low-skilled workers (relative to the marginal cost of public funds). The dashed vertical line indicates the case where the government's welfare function is utilitarian. Values to the right of the dashed line correspond to a government with more redistributive tastes; values to the left of the dashed line correspond to regressive tastes.

Figure 7: Marginal Welfare Gains from a Minimum Wage: Sensitivity Checks

(a) Changing the macro participation elasticity e_1



(b) Changing the macro employment elasticity η_1



Notes: This figure shows how the marginal welfare gain from introducing a minimum wage varies for different values of the macro participation and employment elasticities. Figure 7a checks plots marginal gains in total welfare for various values of the macro participation elasticity; Figure 7b does the same for various values of the macro employment elasticity. The vertical axis measures the gain in total welfare per individual in units of the consumption good. The horizontal axis measures the government's marginal social welfare weight for low-skilled workers (relative to the marginal cost of public funds). The dashed vertical line indicates the case where the government's welfare function is utilitarian. Values to the right of the dashed line correspond to a government with more redistributive tastes; values to the left of the dashed line correspond to regressive tastes.

Table 1: Fraction within \$0.25 of Next Year's Minimum Wage

Age/Ed. Group	Education Group				Total
	HS Dropouts	HS Graduates	Some College	College	
20-29	0.166	0.105	0.106	0.064	0.105
30-39	0.132	0.067	0.051	0.033	0.062
40-49	0.127	0.050	0.036	0.023	0.051
50+	0.124	0.050	0.034	0.024	0.055
Total	0.140	0.075	0.068	0.040	0.075

Notes: This table reports the fraction of workers in each age-education group cell that report a nominal hourly wage within \$0.25 of the following year's nominal minimum wage. The sample is all unmarried individuals in the CPS MORG samples between the ages of 20 and 65 that report not being in school full time. Each observation is weighted by its inverse sampling probability provided in the CPS.

Table 2: Summary Statistics: Monthly CPS (1979-2014)

	(1) All	(2) HS Dropouts	(3) HS Graduates	(4) Some College	(5) College
Labor Force Participant	0.86	0.71	0.86	0.89	0.93
Employed	0.78	0.56	0.75	0.82	0.89
Real Hourly Wage	12.92	10.79	12.39	12.71	16.81
Real MW	7.46	7.43	7.45	7.47	7.49
Age	24.21	23.85	23.86	23.89	25.50
Female	0.45	0.40	0.42	0.48	0.51
White	0.77	0.72	0.76	0.78	0.82
Black	0.17	0.23	0.20	0.16	0.09
Immigrant	0.47	0.59	0.50	0.42	0.42

Notes: This table reports summary statistics for individuals surveyed in the Current Population Survey (CPS) monthly files between 1979 and 2014. The sample is restricted to all unmarried individuals between the ages of 20 and 29 that report not being in school full time. Each observation is weighted by its inverse sampling probability provided in the CPS. All dollar amounts are deflated to 2014 dollars using the BLS Urban Consumer Price Index.

Table 3: Example of Coding the Event Time Variable

State	Year	Avg. Nominal MW	Event Time
WI	1986	3.35	-3
WI	1987	3.35	-2
WI	1988	3.35	-1
WI	1989	3.50	0
WI	1990	3.76	0
WI	1991	4.14	0
WI	1992	4.25	0
WI	1993	4.25	1
WI	1994	4.25	-2
WI	1995	4.25	-1
WI	1996	4.38	0

Table 4: The Effect of Minimum Wages on LFP and EMP

	(1)	(2)	(3)	(4)	(5)
A. LFP (Mean: 0.71)					
post _{st}	-0.005 (0.003)	-0.016** (0.006)	-0.018* (0.010)		
pre*event _{st}		0.005* (0.003)	0.008 (0.009)		0.005* (0.003)
pre*event _{st} ²			0.001 (0.002)		
post*short _{st}				-0.005 (0.004)	-0.015** (0.006)
post*medium _{st}				-0.006 (0.006)	-0.017* (0.010)
B. EMP (Mean: 0.56)					
post _{st}	0.001 (0.004)	-0.015* (0.008)	-0.021* (0.011)		
pre*event _{st}		0.008** (0.003)	0.015 (0.011)		0.008** (0.004)
pre*event _{st} ²			0.001 (0.002)		
post*short _{st}				0.000 (0.006)	-0.015* (0.008)
post*medium _{st}				0.001 (0.007)	-0.015 (0.011)
N	562,283	562,283	562,283	562,283	562,283

Notes: All columns include the following controls: dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status, state fixed effects and year fixed effects. The sample is all unmarried high school dropouts between the ages of 20 and 29 that report not being in school full time. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table 5: Robustness Check: Spatial Controls

	(1)	(2)	(3)	(4)	(5)
A. LFP (Mean: 0.71)					
post _{st}	-0.016*** (0.006)	-0.012** (0.006)	-0.009 (0.006)	-0.007 (0.008)	-0.006 (0.008)
pre*event _{st}	0.005* (0.003)	0.003 (0.002)	0.003 (0.002)	0.003 (0.003)	0.002 (0.003)
B. EMP (Mean: 0.56)					
post _{st}	-0.015* (0.008)	-0.013 (0.009)	-0.009 (0.008)	-0.000 (0.009)	0.005 (0.008)
pre*event _{st}	0.008** (0.003)	0.006* (0.003)	0.006** (0.003)	0.003 (0.003)	-0.000 (0.003)
State-Linear Trend	N	Y	Y	N	N
State-Quadratic Trend	N	N	Y	N	N
Region-Year FE	N	N	N	Y	N
Div.-Year FE	N	N	N	N	Y

Notes: N = 562,283. All columns include the following controls: dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status and state fixed effects. The sample is all unmarried high school dropouts between the ages of 20 and 29 that report not being in school full time. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table 6: Triple Differences Specification

	(1)	(2)	(3)
A. LFP			
$\text{Bite}_g \times \text{post}_{st}$	-0.015 (0.021)	-0.083*** (0.027)	
$\text{Bite}_g \times \text{pre}^* \text{event}_{st}$		0.035** (0.014)	0.036** (0.015)
$\text{Bite}_g \times \text{post}^* \text{short}_{st}$			-0.081*** (0.024)
$\text{Bite}_g \times \text{post}^* \text{medium}_{st}$			-0.090* (0.049)
B. EMP			
$\text{Bite}_g \times \text{post}_{st}$	-0.007 (0.028)	-0.090*** (0.039)	
$\text{Bite}_g \times \text{pre}^* \text{event}_{st}$		0.043*** (0.015)	0.045*** (0.016)
$\text{Bite}_g \times \text{post}^* \text{short}_{st}$			-0.079** (0.031)
$\text{Bite}_g \times \text{post}^* \text{medium}_{st}$			-0.124** (0.054)
N	11,533,091	11,533,091	11,533,091

Notes: The sample is all unmarried CPS respondents between the ages of 20 and 65 that report not being in school full time. All columns include the following controls: dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status, state-year fixed effects, state-age-education group fixed effects, and year-age-education-group fixed effects. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level.

* $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table 7: Large versus Small Minimum Wage Events

	(1)	(2)	(3)
	Real MW	LFP	EMP
$post_{st}$	0.021 (0.082)	0.001 (0.009)	0.005 (0.013)
$post*min\ wage\ increase_{st}$	0.411*** (0.045)	-0.008* (0.005)	-0.010* (0.005)
$min\ wage\ increase_{st}$	-0.427*** (0.096)	-0.002 (0.005)	-0.002 (0.006)
$pre*event_{st}$	-0.127*** (0.025)	0.004 (0.003)	0.007* (0.004)
N	562,283	562,283	562,283

Notes: All columns include the following controls: dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status and state fixed effects. The sample is all unmarried high school dropouts between the ages of 20 and 29 that report not being in school full time. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

A Appendix (not for publication)

A.1 Theoretical appendix

A.1.1 Government

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

$$\frac{1}{\lambda} \mathcal{L} = \frac{1}{\lambda} SW(U_0, U_1, U_2) + \mathcal{B} \quad (\text{A1})$$

A.1.2 Proof of the claim on page 10

Consider a more general model with involuntary unemployment. In particular, assume that the gross wage for workers with ability a is given by $w_a = w(T_a, b; a)$, where $w(\cdot; a)$ is a continuously differentiable function. Similarly the job finding rate is given by $p_a = p(w_a, T_a, b; a)$, where $p(\cdot; a)$ is also continuously differentiable in its arguments. Assume that workers are risk averse, with consumption utility given by $u(c_a)$. Non-participants and the unemployed receive utility $u(b)$. Therefore, labor force participants receive the expected utility $U_a = p_a u(c_a) + (1 - p_a)u(b)$. This framework corresponds to the 'no cross-effects' model in Section II.2 of [Kroft, Kucko, Lehmann, and Schmieder \(2015\)](#). Suppose that the government introduces a minimum wage just above the equilibrium wage for type-1 workers, $\underline{w} = w_1 + d\underline{w}$. The change in U_1 is given by

$$dU_1 = p_1 u'(c_1) \left[1 + \frac{\partial p_1}{\partial \underline{w}} \frac{u(c_1) - u(b)}{p_1 u'(c_1)} \right] d\underline{w}$$

Since the participation decision depends on $U_1 - U_0 = p_1(u(c_1) - u(b))$, the macro participation response to the minimum wage is

$$\frac{dk_1}{d\underline{w}} = n_1 m(U_1 - U_0 | a_1) p_1 u'(c_1) \frac{dU_1}{d\underline{w}} = \left[1 + \frac{\partial p_1}{\partial \underline{w}} \frac{u(c_1) - u(b)}{p_1 u'(c_1)} \right] \frac{\partial k_1}{\partial \underline{w}} \Big|_{\text{micro}}$$

This proves the claim on page 10.

A.1.3 Proof of Proposition 1

Subtracting (13) from (11) leads to

$$\frac{1}{\lambda} \frac{d\mathcal{L}^{MW}}{d\underline{w}} \Big|_{\underline{w} = w_1^*} = -g_1 h_1 \left[\frac{1 - \mu_1}{\mu_1(1 - \beta_1)} \right] - \frac{h_1}{\mu_1} \left(\frac{dh_1}{dT_1} \right)^{-1} \left[p_1 \mu_1 \frac{dk_1}{d\underline{w}} - p_1 \frac{dk_1}{dT_1} \frac{(1 - \mu_1)g_1}{(1 - \beta_1)} + k_1 \frac{\partial p_1}{\partial \underline{w}} (\mu_1 - g_1) \right]$$

Next, factor out $-h_1 \left(\frac{dh_1}{dT_1} \right)^{-1} > 0$ to get

$$\frac{1}{\lambda} \frac{d\mathcal{L}^{MW}}{d\underline{w}} \Big|_{\underline{w}=w_1^*} = -h_1 \left(\frac{dh_1}{dT_1} \right)^{-1} \left[\frac{dh_1}{dT_1} g_1 \left(\frac{1-\mu_1}{\mu_1(1-\beta_1)} \right) + p_1 \frac{dk_1}{d\underline{w}} - p_1 \frac{dk_1}{dT_1} g_1 \left(\frac{1-\mu_1}{\mu_1(1-\beta_1)} \right) + k_1 \frac{\partial p_1}{\partial \underline{w}} \frac{(\mu_1 - g_1)}{\mu_1} \right]$$

Substituting $\frac{dh_1}{dT_1} = p_1 \frac{dk_1}{dT_1} + k_1 \frac{\partial p_1}{\partial T_1} = -p_1 \frac{\beta_1}{\mu_1} \frac{\partial k_1}{\partial \underline{w}} \Big|^{micro} + k_1 \frac{\partial p_1}{\partial T_1}$ and $\frac{\partial p_1}{\partial T_1} = \frac{\partial p_1}{\partial w_1} (1 - \beta_1)$, rearranging the expression above leads to equation (14).

$$\begin{aligned} -\frac{dh_1}{dT_1} \cdot \frac{1}{\lambda h_1} \frac{d\mathcal{L}^{MW}}{d\underline{w}} \Big|_{\underline{w}=w_1^*} &= p_1 \frac{dk_1}{d\underline{w}} + k_1 \frac{\partial p_1}{\partial \underline{w}} \frac{(\mu_1 - g_1)}{\mu_1} + k_1 \frac{\partial p_1}{\partial w_1} g_1 \frac{(1 - \mu_1)}{\mu_1} \\ &= p_1 \frac{dk_1}{d\underline{w}} + k_1 \frac{\partial p_1}{\partial \underline{w}} (1 - g_1) \end{aligned}$$

A.1.4 Wage Spillovers

Several recent papers have argued that minimum wage legislation affects the wages (and also the labor market outcomes) of those above the prevailing minimum (Autor, Manning, and Smith (2016), Dube, Giuliano, and Leonard (2015), Cengiz, Dube, Linder, and Zipperer (2017)). This section shows how (14) is modified when the baseline model is extended to allow for wage spillovers. I continue to assume that job-seekers make labor supply decisions along the extensive margin only. Unlike the baseline model, I relax the assumption that wages are determined by Nash bargaining. Instead, I allow low-skilled wages to arbitrarily affect the wages of high skilled workers, and vice versa (i.e. $\frac{\partial w_2}{\partial w_1} \leq 0$ and $\frac{\partial w_1}{\partial w_2} \leq 0$).

As in the model in Sections 2 and 3, I assume that the conditional job finding rate at each skill level, p_a , only depends (directly) on the gross wage w_a and is continuously differentiable. Taxes, the non-employment benefit and wages at other skill levels may affect the job finding rate, but only through changes to the gross wage. Specifically, I assume that the wage-setting mechanism is such that $w_a = w(T_a, b, w_{-a}; a)$ is a continuously differentiable function of T_a , b and w_{-a} .

Let $\frac{dk_2}{dT_2}$ and $\frac{\partial k_2}{\partial T_2} \Big|^{micro}$ denote the macro and micro labor force participation responses to taxes for high-skilled workers. Similar to the micro participation response to the minimum wage, $\frac{\partial k_2}{\partial T_2} \Big|^{micro}$ captures the labor force participation response to income tax changes by high-skilled workers if the gross wage w_2 and the job finding rate p_2 are held constant. On the other hand, $\frac{dk_2}{dT_2}$ is the macro labor force participation response to increases in the income tax liability faced by high-skilled workers that takes into account general equilibrium changes to w_2 and p_2 . Beginning from the optimal tax system, introducing a binding minimum wage for low skilled workers leads to $SW(\cdot)$ and \mathcal{B} :

$$\begin{aligned}
\frac{1}{\lambda} \frac{dSW}{d\underline{w}} &= \frac{n_1}{\lambda} \int_0^{\hat{\theta}_1} \Phi'(U_1 - \theta) m(\theta|a_1) d\theta \times \frac{dU_1}{d\underline{w}} + \frac{1 - n_1}{\lambda} \int_0^{\hat{\theta}_2} \Phi'(U_2 - \theta) m(\theta|a_2) d\theta \times \frac{dU_2}{d\underline{w}} \\
&= g_1 h_1 \left[1 + \frac{\partial p_1}{\partial \underline{w}} \frac{u(c_1) - u(b)}{p_1 u'(c_1)} \right] + g_2 h_2 \left[1 + \frac{\partial p_2}{\partial w_2} \frac{u(c_2) - u(b)}{p_2 u'(c_2)} \right] \frac{\partial w_2}{\partial \underline{w}} \\
&= g_1 h_1 \frac{dk_1}{d\underline{w}} \left[\frac{\partial k_1}{\partial \underline{w}} \Big|^{micro} \right]^{-1} - g_2 h_2 \frac{dk_2}{d\underline{w}} \left[\frac{\partial k_2}{\partial T_2} \Big|^{micro} \right]^{-1} \tag{A2}
\end{aligned}$$

$$\frac{d\mathcal{B}}{d\underline{w}} = (T_1^* + b^*) \frac{dh_1}{d\underline{w}} + (T_2^* + b^*) \frac{dh_2}{d\underline{w}} \tag{A3}$$

where

$$T_1^* + b^* = \frac{-h_1 \left(1 - \frac{dk_1}{dT_1} \left[\frac{\partial k_1}{\partial T_1} \Big|^{micro} \right]^{-1} g_1 \right) \frac{dh_2}{dT_2} - h_1 \frac{dk_1}{dT_2} \left[\frac{\partial k_1}{\partial T_1} \Big|^{micro} \right]^{-1} g_1 \frac{dh_2}{dT_1} + h_2 \left(1 - \frac{dk_2}{dT_2} \left[\frac{\partial k_2}{\partial T_2} \Big|^{micro} \right]^{-1} g_2 \right) \frac{dh_2}{dT_1} + h_2 \frac{dk_2}{dT_1} \left[\frac{\partial k_2}{\partial T_2} \Big|^{micro} \right]^{-1} g_2 \frac{dh_2}{dT_2}}{\frac{dh_1}{dT_1} \frac{dh_2}{dT_2} - \frac{dh_1}{dT_2} \frac{dh_2}{dT_1}} \tag{A4}$$

and

$$T_2^* + b^* = \frac{-h_2 \left(1 - \frac{dk_2}{dT_2} \left[\frac{\partial k_2}{\partial T_2} \Big|^{micro} \right]^{-1} g_2 \right) \frac{dh_1}{dT_1} - h_2 \frac{dk_2}{dT_1} \left[\frac{\partial k_2}{\partial T_2} \Big|^{micro} \right]^{-1} g_2 \frac{dh_1}{dT_2} + h_1 \left(1 - \frac{dk_1}{dT_1} \left[\frac{\partial k_1}{\partial T_1} \Big|^{micro} \right]^{-1} g_1 \right) \frac{dh_1}{dT_2} + h_1 \frac{dk_1}{dT_2} \left[\frac{\partial k_1}{\partial T_1} \Big|^{micro} \right]^{-1} g_1 \frac{dh_1}{dT_1}}{\frac{dh_1}{dT_1} \frac{dh_2}{dT_2} - \frac{dh_1}{dT_2} \frac{dh_2}{dT_1}} \tag{A5}$$

Note that

- $\frac{dw_2}{dT_1} = \frac{\partial w_2}{\partial w_1} \frac{\partial w_1}{\partial T_1}$ and $\frac{dw_1}{dT_2} = \frac{\partial w_1}{\partial w_2} \frac{\partial w_2}{\partial T_2}$
- $\frac{\partial k_1}{\partial T_1} \Big|^{micro} = - \frac{\partial k_1}{\partial \underline{w}} \Big|^{micro}$
- $\frac{dk_1}{dT_1} = \frac{dk_1}{d\underline{w}} \frac{\partial w_1}{\partial T_1} - \frac{\partial k_1}{\partial \underline{w}} \Big|^{micro}$ and $\frac{dh_1}{dT_1} = \frac{dh_1}{d\underline{w}} \frac{\partial w_1}{\partial T_1} - p_1 \frac{\partial k_1}{\partial \underline{w}} \Big|^{micro}$
- $\frac{dk_1}{dT_2} = \frac{dk_1}{d\underline{w}} \frac{\partial w_1}{\partial w_2} \frac{\partial w_2}{\partial T_2}$ and $\frac{dh_1}{dT_2} = \frac{dh_1}{d\underline{w}} \frac{\partial w_1}{\partial w_2} \frac{\partial w_2}{\partial T_2}$
- $\frac{dk_2}{dT_2} = \frac{dk_2}{d\underline{w}} \frac{\partial w_1}{\partial w_2} \frac{\partial w_2}{\partial T_2} + \frac{\partial k_2}{\partial T_2} \Big|^{micro}$ and $\frac{dh_2}{dT_2} = \frac{dh_2}{d\underline{w}} \frac{\partial w_1}{\partial w_2} \frac{\partial w_2}{\partial T_2} + p_2 \frac{\partial k_2}{\partial T_2} \Big|^{micro}$
- $\frac{dk_2}{dT_1} = \frac{dk_2}{d\underline{w}} \frac{\partial w_1}{\partial T_1}$ and $\frac{dh_2}{dT_1} = \frac{dh_2}{d\underline{w}} \frac{\partial w_1}{\partial T_1}$
- $[\cdot] = \frac{dh_1}{dT_1} \frac{dh_2}{dT_2} - \frac{dh_1}{dT_2} \frac{dh_2}{dT_1} = p_2 \frac{\partial k_2}{\partial T_2} \Big|^{micro} \left[\frac{dh_1}{d\underline{w}} \frac{\partial w_1}{\partial T_1} - p_1 \frac{\partial k_1}{\partial \underline{w}} \Big|^{micro} \right] - p_1 \frac{\partial k_1}{\partial \underline{w}} \Big|^{micro} \frac{dh_2}{d\underline{w}} \frac{\partial w_1}{\partial w_2} \frac{\partial w_2}{\partial T_2}$

Substituting A4 and A5 into A3 and simplifying yields

$$\begin{aligned}
& \frac{1}{\lambda} \frac{d\mathcal{L}}{d\bar{w}} \Big|_{\bar{w}=w_1^*} - h_1 [\cdot]^{-1} \left\{ -g_1 \frac{dk_1}{d\bar{w}} \left[\frac{\partial k_1}{\partial \bar{w}} \right]^m \right\}^{-1} p_2 \frac{\partial k_2}{\partial T_2} \Big|_{\bar{w}}^m \frac{dh_1}{d\bar{w}} \frac{\partial w_1}{\partial T_1} + g_1 \frac{dk_1}{d\bar{w}} \left[\frac{\partial k_1}{\partial \bar{w}} \right]^m \right\}^{-1} p_2 \frac{\partial k_2}{\partial T_2} \Big|_{\bar{w}}^m p_1 \frac{\partial k_1}{\partial \bar{w}} \Big|_{\bar{w}}^m \\
& + g_1 \frac{dk_1}{d\bar{w}} \left[\frac{\partial k_1}{\partial \bar{w}} \right]^m \right\}^{-1} p_1 \frac{\partial k_1}{\partial \bar{w}} \Big|_{\bar{w}}^m \frac{dh_2}{d\bar{w}} \frac{\partial w_1}{\partial w_2} \frac{\partial w_2}{\partial T_2} + (1-g_1) p_2 \frac{\partial k_2}{\partial T_2} \Big|_{\bar{w}}^m \frac{dh_1}{d\bar{w}} + g_1 \frac{dk_1}{d\bar{w}} \left[\frac{\partial k_1}{\partial \bar{w}} \right]^m \right\}^{-1} p_2 \frac{\partial k_2}{\partial T_2} \Big|_{\bar{w}}^m \frac{dh_1}{d\bar{w}} \frac{\partial w_1}{\partial T_1} - p_1 \frac{dk_1}{d\bar{w}} \frac{\partial w_1}{\partial w_2} \frac{\partial w_2}{\partial T_2} g_1 \frac{dh_2}{d\bar{w}} \Big\} \\
& + h_2 [\cdot]^{-1} \left\{ -g_2 p_2 \frac{dk_2}{d\bar{w}} \frac{dh_1}{d\bar{w}} \frac{\partial w_1}{\partial T_1} + g_2 p_2 \frac{dk_2}{d\bar{w}} p_1 \frac{\partial k_1}{\partial \bar{w}} \Big|_{\bar{w}}^m + g_2 \frac{dk_2}{d\bar{w}} \left[\frac{\partial k_2}{\partial T_2} \right]^m \right\}^{-1} p_1 \frac{\partial k_1}{\partial \bar{w}} \Big|_{\bar{w}}^m \frac{dh_2}{d\bar{w}} \frac{\partial w_1}{\partial w_2} \frac{\partial w_2}{\partial T_2} \\
& + g_2 p_2 \frac{dk_2}{d\bar{w}} \frac{dh_1}{d\bar{w}} \frac{\partial w_1}{\partial T_1} + (1-g_2) p_1 \frac{\partial k_1}{\partial \bar{w}} \Big|_{\bar{w}}^m \frac{dh_2}{d\bar{w}} - \frac{dk_2}{d\bar{w}} \frac{\partial w_1}{\partial w_2} \frac{\partial w_2}{\partial T_2} \left[\frac{\partial k_2}{\partial T_2} \right]^m \right\}^{-1} g_2 p_1 \frac{\partial k_1}{\partial \bar{w}} \Big|_{\bar{w}}^m \frac{dh_2}{d\bar{w}} \Big\}
\end{aligned}$$

The following terms in curly braces in the expression above cancel out: the first and fifth terms and the seventh and tenth terms. Simplifying and using the elasticity definitions yields:

$$\begin{aligned}
\frac{1}{\lambda} \frac{d\mathcal{L}}{d\bar{w}} \Big|_{\bar{w}=w_1^*} &= \Lambda_1 \left[p_1 \frac{dk_1}{d\bar{w}} + k_1 \frac{\partial p_1}{\partial \bar{w}} (1-g_1) \right] + \Lambda_2 \left[p_2 \frac{dk_2}{d\bar{w}} + k_2 \frac{\partial p_2}{\partial w_2} \frac{\partial w_2}{\partial \bar{w}} (1-g_2) \right] \\
\frac{1}{\lambda} \frac{d\mathcal{L}}{d\bar{w}} \Big|_{\bar{w}=w_1^*} &= \Lambda_1 \frac{h_1}{\bar{w}} \left[e_1 + (\eta_1 - e_1)(1-g_1) \right] + \Lambda_2 \frac{h_2}{\bar{w}} \left[e_2 + (\eta_2 - e_2)(1-g_2) \right] \quad (\text{A7})
\end{aligned}$$

where $\Lambda_1 = -p_2 \frac{\partial k_2}{\partial T_2} \Big|_{\bar{w}}^{\text{micro}} h_1 [\cdot]^{-1} > 0$ and $\Lambda_2 = p_1 \frac{\partial k_1}{\partial \bar{w}} \Big|_{\bar{w}}^{\text{micro}} h_2 [\cdot]^{-1}$.

Imposing the restriction that there are no wages spillovers (i.e. $\frac{\partial w_2}{\partial \bar{w}} = 0$) implies that the participation and employment responses to the minimum wage by high-skilled workers are equal to zero (i.e. $\frac{dk_2}{d\bar{w}} = \frac{dh_2}{d\bar{w}} = 0$). In this case, the second term in (A7) is equal to zero and the expression collapses to (14).

A.1.5 Endogenous Ability Distribution

The theoretical model set out in Section 2 assumes that the ability distribution is fixed and that the only labor supply response to taxes and the minimum wage is along the extensive (labor force participation) margin. In this section, I relax these assumptions. In particular, I assume that participating in the high-skilled labor market requires obtaining an education and that the education choice is binary. I also assume that individuals are heterogeneous along two dimensions: a fixed participation (or search) cost θ and a fixed cost of obtaining schooling ψ . Individuals draw the fixed participation and schooling costs from a distribution with joint density $\tilde{m}(\theta, \psi)$; the support of both θ and ψ is $[0, +\infty)$.

The timing of the model differs slightly from Section 2. Individuals realize their values of (θ, ψ) and make a joint decision about whether to obtain schooling and participate in the labor market, taking the policy variables, gross wages and the job finding rates as given.⁵³ An individual with

⁵³I assume that $U_2 > U_1 > U_0$.

a participation schooling cost pair (θ, ψ) obtains schooling if $U_2 - \theta - \psi \geq \max\{U_1 - \theta, U_0\}$. The number of individuals that obtain schooling and participate in the high-skilled labor market is

$$n_2 = k_2 = \int_0^{U_2-U_1} \int_0^{U_1-U_0} \tilde{m}(\theta, \psi) d\theta d\psi + \int_0^{U_2-U_0-\theta} \int_{U_1-U_0}^{U_2-U_0} \tilde{m}(\theta, \psi) d\theta d\psi \quad (\text{A8})$$

The number of individuals that do not obtain an education (low-skilled) is $n_1 = 1 - n_2$. The number of individuals that participate in the low-skilled labor market is

$$k_1 = \int_{U_2-U_1}^{+\infty} \int_0^{U_1-U_0} \tilde{m}(\theta, \psi) d\theta d\psi \quad (\text{A9})$$

The labor demand side of the economy is the same as in Section 2. The micro (partial equilibrium) and macro (general equilibrium) participation responses to the minimum wage by low-skilled individuals are respectively

$$\left. \frac{\partial k_1}{\partial \underline{w}} \right|^{\text{micro}} = \left[\int_{U_2-U_1}^{+\infty} \tilde{m}(U_1 - U_0, \psi) d\psi + \int_0^{U_1-U_0} \tilde{m}(\theta, U_2 - U_1) d\theta \right] p_1 \quad (\text{A10})$$

$$\frac{\partial k_1}{\partial \underline{w}} = \left[\int_{U_2-U_1}^{+\infty} \tilde{m}(U_1 - U_0, \psi) d\psi + \int_0^{U_1-U_0} \tilde{m}(\theta, U_2 - U_1) d\theta \right] p_1 \left(1 + \frac{\partial p_1}{\partial \underline{w}} \frac{c_1 - b}{p_1} \right) \quad (\text{A11})$$

Policy changes affect the gains from obtaining an education as well as the value of being in the labor force. The first term in square brackets in equations A10 and A11 represents the change in low-skilled labor force participation due to entry (or exit) by individuals without an education. The second term represents the change in low-skilled labor force participation due changes in fraction of individuals that obtain schooling.

The Lagrangian function for the government's optimization problem writes

$$\mathcal{L} = SW(U_0, U_1, U_2) - \lambda(b + E - (T_1 + b)h_1 - (T_2 + b)h_2)$$

where

$$\begin{aligned} SW(U_0, U_1, U_2) = & \int_{U_2-U_1}^{+\infty} \int_0^{U_1-U_0} \Phi(U_1 - \theta) \tilde{m}(\theta, \psi) d\theta d\psi + \int_{U_2-U_1}^{+\infty} \int_{U_1-U_0}^{+\infty} \Phi(U_0) \tilde{m}(\theta, \psi) d\theta d\psi \\ & + \int_0^{U_2-U_1} \int_0^{U_1-U_0} \Phi(U_2 - \theta - \psi) \tilde{m}(\theta, \psi) d\theta d\psi + \int_0^{U_2-U_0-\theta} \int_{U_1-U_0}^{U_2-U_0} \Phi(U_2 - \theta - \psi) \tilde{m}(\theta, \psi) d\theta d\psi \\ & + \int_{U_2-U_0-\theta}^{U_2-U_1} \Phi(U_0) \tilde{m}(\theta, \psi) d\theta d\psi + \int_0^{U_2-U_1} \int_{U_2-U_0}^{+\infty} \Phi(U_0) \tilde{m}(\theta, \psi) d\theta d\psi \end{aligned}$$

Individuals can be partitioned into six groups depending on their (privately optimal) education and labor market participation decisions. Each term in the $SW(\cdot)$ function corresponds to the contribution to the government's social welfare objective by each of these six groups. For example,

the first term is the contribution to social welfare from low-skilled individuals that participate in the labor market. The second, fifth and sixth terms are the contribution to social welfare from low-skilled individuals that do not participate in the labor market. The third and fourth terms capture the contribution to social welfare from high-skilled individuals that participate in the high-skilled labor market.

Lemma 6. *With an endogenous ability distribution, the level of the optimal employment tax liability for the low- and high-skilled labor workers are, respectively,*

$$T_1^* + b^* = \frac{-h_1 \left(\frac{\mu_1 - \beta_1 \tilde{g}_1}{\mu_1} \right) \frac{dh_2}{dT_2} + h_2 \left(\frac{\mu_2 - \beta_2 \tilde{g}_2}{\mu_2} \right) \frac{dh_2}{dT_1}}{\frac{dh_1}{dT_1} \frac{dh_2}{dT_2} + \frac{dh_1}{dT_2} \frac{dh_2}{dT_1}} \quad (\text{A12})$$

$$T_2^* + b^* = \frac{h_1 \left(\frac{\mu_1 - \beta_1 \tilde{g}_1}{\mu_1} \right) \frac{dh_1}{dT_2} - h_2 \left(\frac{\mu_2 - \beta_2 \tilde{g}_2}{\mu_2} \right) \frac{dh_1}{dT_1}}{\frac{dh_1}{dT_1} \frac{dh_2}{dT_2} + \frac{dh_1}{dT_2} \frac{dh_2}{dT_1}} \quad (\text{A13})$$

where

$$\tilde{g}_1 = \frac{\int_{U_2 - U_1}^{+\infty} \int_0^{U_1 - U_0} \Phi'(U_1 - \theta) \tilde{m}(\theta, \psi) d\theta d\psi \cdot p_1}{\lambda h_1}$$

and

$$\tilde{g}_2 = \frac{\left\{ \int_0^{U_2 - U_1} \int_0^{U_1 - U_0} \Phi'(U_2 - \theta - \psi) \tilde{m}(\theta, \psi) d\theta d\psi + \int_0^{U_2 - U_0 - \theta} \int_{U_1 - U_0}^{U_2 - U_0} \Phi'(U_2 - \theta - \psi) \tilde{m}(\theta, \psi) d\theta d\psi \right\} \cdot p_2}{\lambda h_2}$$

Lemma 6 shows that the optimal employment tax liabilities for low- and high-skilled workers differ from Lemma 4.⁵⁴ It is well-known that the possibility of labor supply responses along the skill choice or intensive margin increases the employment tax rate at the bottom of the wage distribution and decreases the employment tax rate at the top of the wage distribution (Saez (2002b), Kroft, Kucko, Lehmann, and Schmieder (2015)). This is because reducing the employment tax rate faced by low-skilled workers induces some that would otherwise have obtained an education and participated in the high-skilled labor market to change their education choice. The possibility of distorting education choices pushes the $T_1^* + b^*$ upwards.

As in the main text, the government considers whether to introduce a minimum wage that is binding in the low-skilled labor market after setting income taxes optimally. To illustrate how the minimum wage differs from income taxes, I compare the social welfare and fiscal effects of a minimum wage with a CETC. The following lemma shows that the effects of the minimum wage

⁵⁴If the cross-effect terms are set to zero ($\frac{dh_1}{dT_2} = \frac{dh_2}{dT_1} = 0$), equations (A11) and (A12) reduce to the optimal employment tax formula in the main text (equation (12)).

and a CETC on the social welfare function $SW(U_0, U_1, U_2)$ are identical to those from the model in Section 2.

Lemma 7. *With an endogenous ability distribution, the changes in social welfare due to the introduction of a minimum wage and a CETC are, respectively,*

$$dSW^{MW} = \tilde{g}_1 h_1 \left[\frac{\mu_1 - \beta_1}{\mu_1(1 - \beta_1)} \right] d\bar{w} = \tilde{g}_1 h_1 \frac{dk_1}{d\bar{w}} \left[\frac{\partial k_1}{\partial \bar{w}} \Big|^{micro} \right]^{-1} d\bar{w} \quad (A14)$$

and

$$dSW^{CETC} = \tilde{g}_1 h_1 \frac{1}{\mu_1} d\bar{w} \quad (A15)$$

Intuitively, those that are induced to switch their education choice due to the minimum wage or a CETC are initially indifferent between participating in either labor market. Consequently, skill changes have no first order effect on $SW(\cdot)$ (by the envelope theorem). However, policy-induced skill changes do have first-order effects on the net tax revenue collected by the government.

Lemma 8. *With an endogenous ability distribution, the change in net tax revenue \mathcal{B} due the introduction of a minimum wage and CETC are, respectively,*

$$d\mathcal{B}^{MW} = (T_1 + b)dh_1^{MW} + (T_2 + b)dh_2^{MW} \quad (A16)$$

$$d\mathcal{B}^{CETC} = h_1 dT_1^{CETC} + (T_1 + b)dh_1^{CETC} + (T_2 + b)dh_2^{CETC} \quad (A17)$$

Combining Lemmas 6, 7, and 8 and substituting the optimal employment tax formulas (equations (A12) and (A13)) leads to the formula for the change in total welfare from the introduction of a minimum wage.

Proposition 3. *Suppose the ability distribution is endogenous in the way described above. Beginning from the optimal income tax allocation, the change in total welfare from the introduction of a minimum is*

$$\begin{aligned} \frac{1}{\lambda h_1} \frac{d\mathcal{L}^*}{d\bar{w}} \Big|_{\bar{w} = w_1^*} &= \Psi_1 \left[p_1 \frac{dk_1}{d\bar{w}} + k_1 \frac{\partial p_1}{\partial \bar{w}} (1 - \tilde{g}_1) \right] + \Psi_2 \left[k_1 \frac{\partial p_1}{\partial \bar{w}} \left(1 - \frac{dk_2}{dT_2} \left(\frac{\partial k_2}{\partial T_2} \Big|^{micro} \right)^{-1} \tilde{g}_2 \right) \frac{h_2}{h_1} \right. \\ &\quad \left. - \frac{dh_1}{dT_2} \frac{dk_1}{d\bar{w}} \left(\frac{\partial k_1}{\partial \bar{w}} \Big|^{micro} \right)^{-1} \right] \end{aligned} \quad (A18)$$

where $\left[\cdot \right] = \frac{dh_1}{dT_1} \frac{dh_2}{dT_2} + \frac{dh_2}{dT_1} \frac{dh_1}{dT_2} > 0$ and $\Psi_1 = -\frac{dh_2}{dT_2} \left[\cdot \right]^{-1} > 0$ and $\Psi_2 = -\left[\cdot \right]^{-1} \frac{dh_2}{d\bar{w}} \left[\frac{dk_1}{d\bar{w}} \right]^{-1} \frac{\partial k_1}{\partial \bar{w}} \Big|^{micro} > 0$.

Imposing the restriction that the ability distribution is exogenous implies that $\frac{dk_1}{dT_2} = \frac{dk_2}{dT_1} = \frac{dk_2}{dw} = 0$. In this case, (A18) reduces to (14). The change in total welfare from the effect of the minimum wage on education choices is captured by the third term in (A18). Importantly, the third term in (A18) is proportional to $\frac{dh_2}{dw}$ (through Ψ_2). Therefore, if skill decisions are not responsive to the minimum wage, the third term in (A18) will be small. As in the main text, the intuition for this third term can be understood by considering the distortions to tightness in the low- and high-skilled labor markets in the absence of the minimum wage. The following two special cases illustrate how the minimum wage may alleviate or exacerbate distortions to the labor market tightness beginning from the optimal income tax allocation.

Special Case 1 (Utilitarian SWF and Efficient High-Skilled Labor Market)

Suppose that the social welfare function $SW(\cdot)$ is utilitarian and that the tightness in the high-skilled labor market is efficient. The former assumption implies that $\tilde{g}_0 = \tilde{g}_1 = \tilde{g}_2 = 1$ and that social welfare is maximized if output net of fixed vacancy, participation and schooling costs is maximized. The latter assumption implies that $\beta_2 = \mu_2$ and $\frac{dk_2}{dT_2} = \frac{\partial k_2}{\partial T_2} \Big|_{\text{micro}}$. In this case, the optimal tax formulas reduce to

$$T_1^* + b^* = \frac{-h_1 \left(\frac{\mu_1 - \beta_1}{\mu_1} \right) \frac{dh_2}{dT_2}}{\frac{dh_1}{dT_1} \frac{dh_2}{dT_2} - \frac{dh_1}{dT_2} \frac{dh_2}{dT_1}} \quad (\text{A19})$$

and

$$T_2^* + b^* = \frac{h_1 \left(\frac{\mu_1 - \beta_1}{\mu_1} \right) \frac{dh_1}{dT_2}}{\frac{dh_1}{dT_1} \frac{dh_2}{dT_2} - \frac{dh_1}{dT_2} \frac{dh_2}{dT_1}} \quad (\text{A20})$$

Thus, $T_1^* + b^* > 0$ and $T_2^* + b^* > 0$ if the laissez-faire tightness in the low-skilled labor market is inefficiently high (i.e. $\beta_1 < \mu_1$).⁵⁵ As in the main text, if the tightness in the low-skilled labor market is inefficiently high in the absence of government intervention, the government sets a positive employment tax liability for low-skilled workers. This reduces the number of vacancies that firms are willing to create, as well as low-skilled labor market participation. However, unlike the model with only extensive margin labor supply responses, the optimal policy also involves a positive employment tax for high-skilled workers. A positive employment tax for high-skilled workers discourages individuals from switching to the high-skilled labor market. However, this shrinks the tax base by reducing the supply and demand for high-skilled workers.

In this situation, (A18) reduces to

⁵⁵Conversely, if $\beta_1 > \mu_1$, then $T_1^* + b^* < 0$ and $T_2^* + b^* < 0$. Moreover, note that $\frac{dh_1}{dT_2} > 0$, $\frac{dh_2}{dT_1} > 0$ and $\frac{dh_1}{dT_1} \frac{dh_2}{dT_2} - \frac{dh_1}{dT_2} \frac{dh_2}{dT_1} > 0$ by assumption.

$$\frac{1}{\lambda h_1} \frac{d\mathcal{L}^*}{d\bar{w}} \Big|_{\bar{w} = w_1^*} = \Psi_1 p_1 \frac{dk_1}{d\bar{w}} - \Psi_2 \frac{dh_1}{dT_2} \frac{dk_1}{d\bar{w}} \left(\frac{\partial k_1}{\partial \bar{w}} \Big|_{\text{micro}} \right)^{-1} = - \left[\cdot \right]^{-1} \frac{dh_2}{dT_2} \left\{ p_1 \frac{dk_1}{d\bar{w}} - \frac{dh_2}{d\bar{w}} \frac{dh_1}{dT_2} \left[\frac{dh_2}{dT_2} \right]^{-1} \right\} \quad (\text{A21})$$

The first term in (A21) is the *Participation effect* from Section 3.4. Since $\frac{dh_1}{dT_2} > 0$, $\frac{dh_2}{dT_2} < 0$ and $\frac{dh_2}{d\bar{w}}$ has the opposite sign as $\frac{dk_1}{d\bar{w}}$, the second term in (A21) is positive whenever the *Participation effect* is positive. Intuitively, the second term in (A21) captures the fact that the raising the value of participating in the low-skilled labor market also distorts education decisions downward. This exacerbates the decline in the tax base due to the positive employment tax rate on high-skilled workers. Conversely, suppose that the tightness in the low-skilled labor market is inefficiently low in the absence of any government intervention. In this case, although the minimum wage reduces low-skilled labor force participation, it leads to higher tax revenue from high-skilled workers. This is because the minimum wage alleviates the disincentive to obtain an education caused by the a negative employment tax for low-skilled workers.

Special Case 2 (Efficient Low- and High-Skilled Labor Markets)

If both the low- high-skilled labor markets are efficient in the absence of any government intervention then: $\beta_2 = \mu_2$, $\frac{dk_2}{dT_2} = \frac{\partial k_2}{\partial T_2} \Big|_{\text{micro}}$, $\beta_1 = \mu_1$, $\frac{dk_1}{dT_1} = - \frac{\partial k_1}{\partial \bar{w}} \Big|_{\text{micro}}$, and $\frac{dk_1}{d\bar{w}} = \frac{dk_2}{d\bar{w}} = \frac{dh_2}{d\bar{w}} = 0$. Since the low- and high-skilled labor markets are efficient (i.e. output net of fixed participation and schooling costs is maximized), the government's only role is to use the income tax system and the minimum wage to redistribute income. In this case, the optimal tax formulas reduce to

$$T_1^* + b^* = \frac{-h_1(1 - \tilde{g}_1) \frac{dh_2}{dT_2} + h_2(1 - \tilde{g}_2) \frac{dh_1}{dT_2}}{\frac{dh_1}{dT_1} \frac{dh_2}{dT_2} - \frac{dh_1}{dT_2} \frac{dh_2}{dT_1}} \quad (\text{A22})$$

and

$$T_2^* + b^* = \frac{-h_2(1 - \tilde{g}_2) \frac{dh_1}{dT_1} + h_1(1 - \tilde{g}_1) \frac{dh_1}{dT_2}}{\frac{dh_1}{dT_2} \frac{dh_2}{dT_2} - \frac{dh_1}{dT_2} \frac{dh_2}{dT_1}} \quad (\text{A23})$$

If \tilde{g}_1 is large enough, the government wants to set a negative employment tax liability for low-skilled workers. Relative to the baseline model with no education decision, however, it's ability to do so is limited by the fact that this would induce some individuals to switch to the low-skilled labor market (second term in the numerator of (A22)). Thus, the optimal employment tax rate at the bottom (top) of the wage distribution is higher (lower) than it would be without an "intensive margin" response. In this situation, (A18) reduces to

$$\frac{1}{\lambda h_1} \left. \frac{d\mathcal{L}^*}{d\bar{w}} \right|_{\bar{w} = \bar{w}_1^*} = \left[\frac{dh_1}{dT_1} \frac{dh_2}{dT_2} - \frac{dh_1}{dT_2} \frac{dh_2}{dT_1} \right]^{-1} k_1 \frac{\partial p_1}{\partial \bar{w}} \left(-h_1(1 - \tilde{g}_1) \frac{dh_2}{dT_2} + h_2(1 - \tilde{g}_2) \frac{dh_1}{dT_2} \right) \quad (\text{A24})$$

The term in round braces is the numerator of (A22). This term captures the sign of the employment tax liability faced by low-skilled workers. If it is negative, then the optimal income tax system features a subsidy for low-skilled workers. Given that the low- and high skilled labor markets are efficient in the absence of government intervention, a negative employment tax for low-skilled workers distorts the tightness upwards through the familiar equity-efficiency trade-off. In this case, because $\frac{\partial p_1}{\partial \bar{w}} < 0$, the minimum wage is welfare improving because it lowers the tightness in the low-skilled labor market. On the other hand, if the optimal income tax features a positive employment tax at the bottom, then the term in round braces is positive and optimal tax system pushes the tightness in the low-skilled labor market further below its efficient (first-best) level. In that case, introducing a minimum wage exacerbates the distortion to the tightness caused by the income tax system.

In general, the tightness in the low-skilled labor market is not efficient because of externalities caused by search frictions and distortions caused by the the income tax system. Corollary 2 gives the sufficient conditions under which the minimum wage is desirable in terms of estimable labor force participation and employment responses to the minimum wage and taxes.

Corollary 2. *Suppose the ability distribution is endogenous in the way described above. Sufficient conditions for the minimum wage to be welfare improving are: (i) $\frac{dk_1}{d\bar{w}} > 0$, (ii) $\tilde{g}_1 > 1$, and*

$$(iii) \left(\frac{dh_1}{d\bar{w}} - p_1 \frac{dk_1}{d\bar{w}} \right) \left(1 - \frac{dk_2}{dT_2} \left[\frac{\partial k_2}{\partial T_2} \right]^{micro} \right)^{-1} \tilde{g}_2 \frac{h_2}{h_1} > \frac{dh_1}{dT_2} \frac{dk_1}{d\bar{w}} \left[\frac{\partial k_1}{\partial \bar{w}} \right]^{micro} \right)^{-1}.$$

A.1.6 General matching technology and wage setting

This section shows how (14) is modified when the baseline model is extended to allow for a non-constant returns to scale matching technology. Although the empirical literature supports the constant returns to scale approximation (Petrongolo and Pissarides (2001, 2006)), relaxing this assumption has important implications for the desirability of the minimum wage. In particular, I assume that the only restriction on p_a is that it is continuously differentiable in the gross wage w_a , the tax liability T_a , and the non-employment benefit b faced by type- a workers. Moreover, assume that the wage-setting mechanism is such that $w_a = w(\cdot; a)$ is a continuously differentiable function of T_a and b . This framework corresponds to the ‘no cross-effect’ model in Kroft, Kucko, Lehmann, and Schmieder (2015).⁵⁶

The equivalents of the changes in $SW(\cdot)$ and \mathcal{B} are:

⁵⁶This model only assumes that the labor supply response is on the participation margin only and that the minimum wage and taxes have no spillover effects on the wages and job finding rates in other occupations (separable production function).

$$\frac{1}{\lambda}dSW^{MW} = \frac{n_1}{\lambda} \int_0^{\hat{\theta}_1} \Phi'(U_1 - \theta)m(\theta|a_1)d\theta dU_1^{MW} = g_1 h_1 \left[1 + \frac{\partial p_1}{\partial \underline{w}} \frac{u(c_1) - u(b)}{p_1 u'(c_1)} \right] d\underline{w} \quad (A25)$$

$$d\mathcal{B}^{SW} = (T_1^* + b^*)dh_1^{MW} = -h_1 \left\{ 1 + \left[\frac{\partial w_1}{\partial T_1} - 1 + \frac{\partial p_1}{\partial T_1} \cdot \frac{u(c_1) - u(b)}{p_1 u'(c_1)} \right] g_1 \right\} \left(\frac{dh_1}{dT_1} \right)^{-1} \frac{\partial h_1}{\partial \underline{w}} d\underline{w} \quad (A26)$$

$$\frac{1}{\lambda}dSW^{CETC} = g_1 h_1 \left[1 + \frac{\partial p_1}{\partial T_1} \left(\frac{\partial w_1}{\partial T_1} - 1 \right) \frac{u(c_1) - u(b)}{p_1 u'(c_1)} \right] d\underline{w} \quad (A27)$$

$$d\mathcal{B}^{CETC} = h_1 dT_1^{CETC} + (T_1^* + b^*)dh_1^{CETC} = -g_1 h_1 \left[1 + \frac{\partial p_1}{\partial T_1} \left(\frac{\partial w_1}{\partial T_1} - 1 \right) \frac{u(c_1) - u(b)}{p_1 u'(c_1)} \right] d\underline{w} \quad (A28)$$

Combining (A25)-(A28) and evaluating at the optimal income tax allocation leads to the expression for the total welfare effect due to the minimum wage:

$$-\frac{dh_1}{dT_1} \cdot \frac{1}{\lambda h_1} \cdot \frac{d\mathcal{L}^{MW}}{d\underline{w}} \Big|_{\underline{w}=w_1^*} = p_1 \frac{dk_1}{d\underline{w}} + k_1 \frac{\partial p_1}{\partial \underline{w}} (1 - g_1) + k_1 \left[\frac{\partial p_1}{\partial \underline{w}} \frac{\partial w_1}{\partial T_1} - \frac{\partial p_1}{\partial T_1} \right] g_1 \quad (A29)$$

Given that $k_1 \frac{\partial p_1}{\partial \underline{w}} = \frac{dh_1}{d\underline{w}} - p_1 \frac{dk_1}{d\underline{w}}$ and $k_1 \frac{\partial p_1}{\partial T_1} = \frac{dh_1}{dT_1} - p_1 \frac{dk_1}{dT_1}$, (A29) can be expressed in terms of the macro participation and employment responses to the minimum wage and taxes.

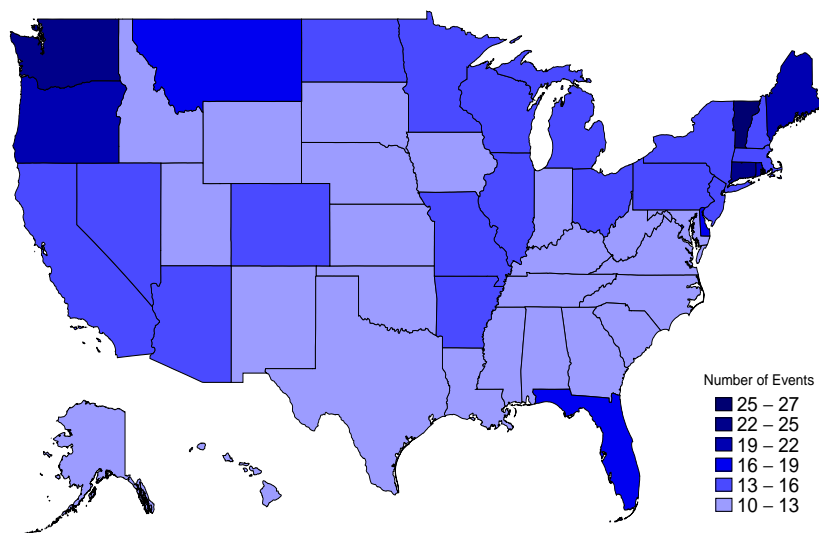
$$-\frac{dh_1}{dT_1} \cdot \frac{1}{\lambda h_1} \cdot \frac{d\mathcal{L}^{MW}}{d\underline{w}} \Big|_{\underline{w}=w_1^*} = p_1 \frac{dk_1}{d\underline{w}} + \left[\frac{dh_1}{d\underline{w}} - p_1 \frac{dk_1}{d\underline{w}} \right] (1 - g_1) + \left[\left(\frac{dh_1}{d\underline{w}} - p_1 \frac{dk_1}{d\underline{w}} \right) \frac{\partial w_1}{\partial T_1} - \left(\frac{dh_1}{dT_1} - p_1 \frac{dk_1}{dT_1} \right) \right] g_1$$

In this case, the form of (14) is the same except for an additional term that captures the tax revenue gains due to the change in the job finding rate. To illustrate, suppose the matching function exhibits decreasing returns to scale. This implies that the job finding rate is decreasing in the number of job-seekers and increasing in the tightness. If the minimum wage increases labor force participation, the job finding rate falls more than in the constant returns to scale case. As a result, some matches that would have taken place do not occur. This has a negative fiscal cost proportional to the government's welfare weight g_1 . In contrast, a CETC has a less negative impact on the worker's job finding rate because it encourages vacancy creation. With increasing returns to scale, the job finding rate may increase with the minimum wage. In this situation, the *Redistribution effect* may be negative if $\frac{dk_1}{d\underline{w}} > 0$. Consequently, the additional term is positive because employment is unambiguously increases.

A.2 Empirical appendix

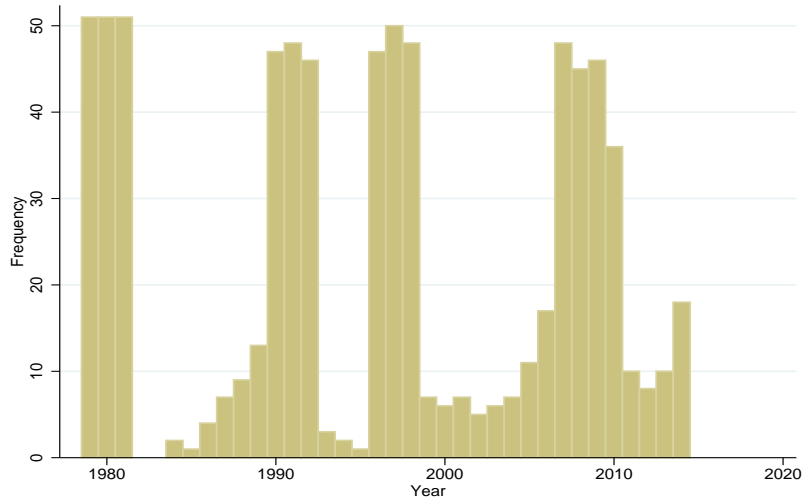
A.2.1 Appendix figures

Figure A1: Number of Minimum Wage Events by State



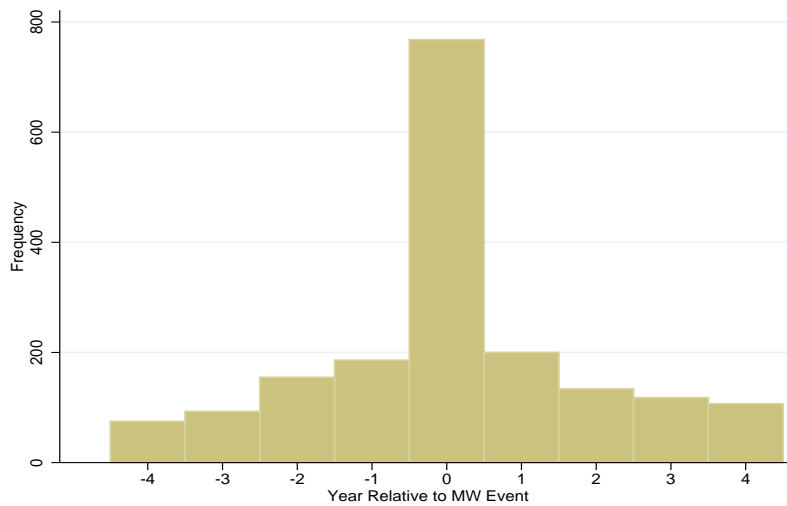
Notes: This figure reports the number of minimum wage increases by state, from 1979 to 2014.

Figure A2: Distribution of Events by Year



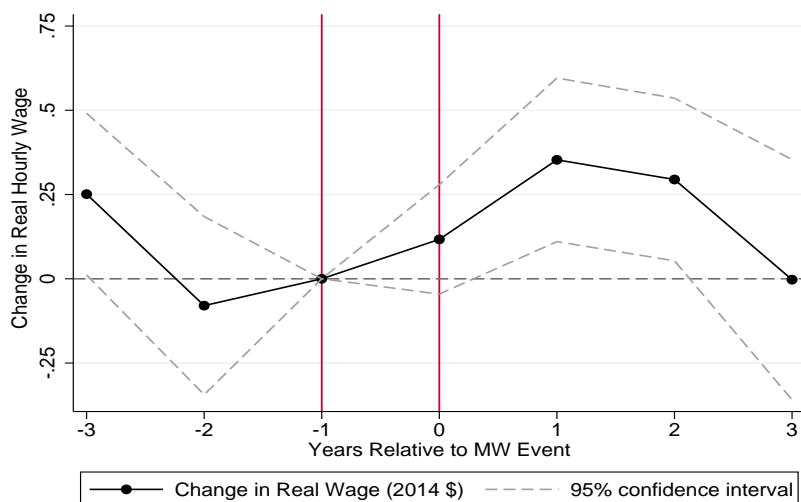
Notes: This figure reports a frequency count of minimum wage increases by year, from 1979 to 2014.

Figure A3: Distribution of Event Time



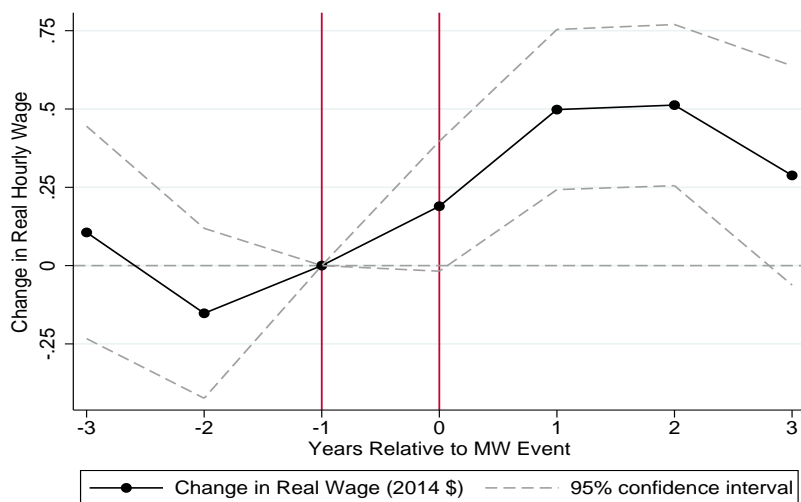
Notes: This histogram plots the frequency count for values of the $event_{st}$ (years relative to minimum wage increase) variable.

Figure A4: The Effect of the Minimum Wage on Real Hourly Wages



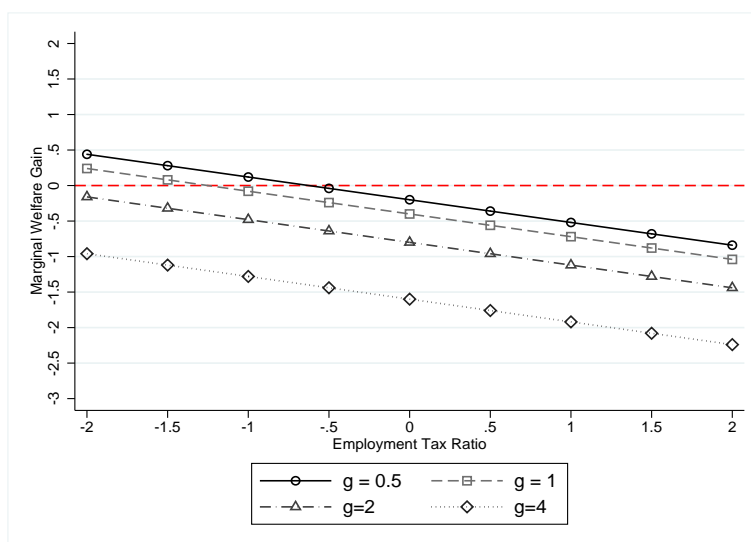
Notes: This figure reports estimates of the effect of nominal minimum wage increases on reported real hourly wage of CPS MORG respondents. The vertical lines indicate the year before a minimum wage increase and the year of a minimum wage increase, respectively. Each marker is an estimate of the a^i coefficient from equation (18). All dollar amounts are deflated to 2014 dollars using the BLS Urban CPI.

Figure A5: De-trended Real Hourly Wage Response to the Minimum Wage



Notes: This figure reports de-trended estimates of the effect of nominal minimum wage increases on the reported real hourly wage of unmarried age 20-29 CPS MORG respondents with less than a high school diploma. The vertical lines indicate the year before a minimum wage increase and the year of a minimum wage increase, respectively. Each marker is an estimate of the a^i coefficient from equation (18) after fitting and removing a linear trend for the years before a minimum wage event. All dollar amounts are deflated to 2014 dollars using the BLS Urban CPI.

Figure A6: Marginal Welfare Gain from Introducing a Minimum Wage: Fixed Income Tax System



Notes: This figure shows the results from a second policy experiment: calculating the welfare gains from introducing a minimum wage beginning from any fixed income tax system. To do so, I calibrate the sufficient statistics formula $g_1 \frac{e_1}{e_1^m} + \frac{T_1+b}{w} \eta_1 \propto g_1 h_1 \frac{dk_1}{dw} \left[\frac{\partial k_1}{\partial w} \Big|^{micro} \right]^{-1} + (T_1 + b) \frac{dh_1}{dw}$ introduced in Section 4.4. In addition to the macro participation and employment elasticities and the welfare weight for low-skilled workers, this formula also depends on the *micro* participation response to the minimum wage. Given that the micro participation response to the minimum wage is equal to -1 times the micro participation response to income taxes, I use the estimate $e_1^m = 0.6$ from Kroft, Kucko, Lehmann, and Schmieder (2015).

A.2.2 Appendix tables

Table A1: The Effect of Minimum Wages on Real Hourly Wages

	(1)	(2)	(3)	(4)	(5)
post _{st}	0.131** (0.061)	0.219 (0.140)	0.258 (0.240)		
pre*event _{st}		-0.046 (0.061)	-0.089 (0.271)		-0.046 (0.061)
pre*event _{st} ²			-0.010 (0.066)		
post*short _{st}				0.134** (0.061)	0.220 (0.138)
post*medium _{st}				0.125 (0.090)	0.215 (0.161)
N	42,783	42,783	42,783	42,783	42,783

Notes: All columns include the following controls: dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status, state fixed effects and year fixed effects. The sample is all unmarried high school dropouts in the CPS MORG between the ages of 20 and 29 that report not being in school full time. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table A2: Covariates

	(1)	(2)	(3)	(4)	(5)
	Age	Female	White	Black	Immigrant
A. All unmarried individuals					
$post_{st}$	-0.018	-0.002	0.001	-0.000	-0.016***
	(0.095)	(0.002)	(0.002)	(0.002)	(0.005)
$pre*event_{st}$	0.038	0.001	0.000	0.001	0.006***
	(0.038)	(0.001)	(0.000)	(0.000)	(0.002)
N	11,533,091	11,533,091	11,533,091	11,533,091	11,533,091
B. HS dropouts age 20-29					
$post_{st} = 0$	-0.008	-0.011	0.002	-0.001	-0.030***
	(0.059)	(0.008)	(0.005)	(0.005)	(0.011)
$pre*event_{st}$	0.000	0.003	0.000	0.001	0.012***
	(0.024)	(0.003)	(0.002)	(0.002)	(0.004)
N	562,283	562,283	562,283	562,283	562,283

Notes: The sample in panel A is all unmarried CPS respondents that report not being in school full time. The sample in panel B is all unmarried high school dropouts between the ages of 20 and 29 that report not being in school full time. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table A3: Natives versus Immigrants

	Baseline			Census Division by Year FE		
	(1) All	(2) Natives	(3) Immigrants	(4) All	(5) Natives	(6) Immigrants
A. Immigrant						
post _{st}	-0.030*** (0.011)			-0.051*** (0.016)		
Mean	0.592			0.592		
B. LFP						
post _{st}	-0.016** (0.006)	-0.011 (0.009)	-0.024*** (0.008)	-0.006 (0.008)	0.002 (0.010)	-0.014 (0.009)
Mean	0.706	0.667	0.732	0.706	0.667	0.732
C. EMP						
post _{st}	-0.015* (0.008)	-0.009 (0.010)	-0.027** (0.011)	0.005 (0.008)	0.004 (0.011)	0.002 (0.011)
Mean	0.559	0.503	0.597	0.559	0.503	0.597

Notes: The sample in columns 1 and 4 is all unmarried CPS respondents with less than a high school diploma between the ages of 20 and 29 that report not being in school full time. Columns 2 and 5 restrict the sample to individuals born in the U.S. , and columns 3 and 6 restrict the sample to immigrants. All columns include the following controls: pre*event_{st}, dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age tends, and state fixed effects. The specifications in columns 4-6 include census division-by-year FE. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table A4: Effect of the Minimum Wage by 'Bites'

	(1)	(2)	(3)	(4)
	HS Drop	HS Grad	Some College	College+
A. LFP				
$post_{st}$	-0.016*** (0.006)	-0.005 (0.003)	0.003 (0.003)	-0.002 (0.002)
$pre*event_{st}$	0.005 (0.003)	0.002* (0.001)	-0.000 (0.001)	0.000 (0.001)
B. EMP				
$post_{st}$	-0.015* (0.008)	-0.003 (0.005)	0.003 (0.004)	-0.000 (0.003)
$pre*event_{st}$	0.008** (0.004)	0.006*** (0.002)	0.003 (0.002)	0.001 (0.002)
C. Real Hourly Wage				
$post_{st}$	0.219 (0.139)	0.269* (0.140)	0.279** (0.122)	0.419 (0.330)
$pre*event_{st}$	-0.046 (0.062)	-0.088 (0.066)	-0.130*** (0.049)	-0.163 (0.122)

Notes: The sample is all unmarried CPS respondents between the ages of 20 and 29 that report not being in school full time. Each column corresponds to a regression for individuals in the indicated education group. All columns include the following controls: dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status and state fixed effects. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level.

* $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table A5: Pre- versus Post-1997

	(1) 1979-2014	(2) 1979-1996	(3) 1997-2014
A. LFP			
post _{st}	-0.016*** (0.006)	-0.018 (0.011)	-0.009 (0.007)
B. EMP			
post _{st}	-0.015* (0.008)	-0.018 (0.014)	-0.012 (0.009)
C. Real Hourly Wage			
post _{st}	0.219 (0.140)	0.170 (0.324)	0.136 (0.157)
N	562,283	321,955	240,328

Notes: The sample is all unmarried CPS respondents between the ages of 20 and 29 that report not being in school full time. Each column corresponds to a regression for individuals in the indicated time period. All columns include the following controls: pre*event_{st}, dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status and state fixed effects. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table A6: Effect of the Minimum Wage on Women and Men

	(1)	(2)	(3)
	All	Women	Men
A. LFP			
post _{st}	-0.016*** (0.006)	-0.047*** (0.015)	0.005 (0.006)
Mean	0.706	0.541	0.817
B. EMP			
post _{st}	-0.015* (0.008)	-0.040*** (0.015)	0.000 (0.009)
Mean	0.559	0.410	0.659
C. Real Hourly Wage			
post _{st}	0.219 (0.140)	0.380** (0.171)	0.133 (0.2200)
Mean	10.788	9.522	11.348

Notes: The sample is all unmarried CPS respondents between the ages of 20 and 29 that report not being in school full time. Each column corresponds to a regression for individuals of the indicated gender. All columns include the following controls: pre*event_{st}, dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status and state fixed effects. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table A7: Baseline, No Sample Weights, MORG Sample only

	(1)	(2)	(3)	(4)
	Baseline	Baseline, No weights	MORG sample	MORG sample, No weights
A. LFP				
post _{st}	-0.016*** (0.006)	-0.015** (0.006)	-0.007 (0.009)	-0.009 (0.009)
Mean	0.706	0.704	0.702	0.700
B. EMP				
post _{st}	-0.015* (0.008)	-0.016** (0.007)	-0.009 (0.011)	-0.010 (0.010)
Mean	0.559	0.556	0.560	0.557

Notes: The sample in columns 1 and 2 is all unmarried CPS respondents between the ages of 20 and 29 that report not being in school full time. The sample in column 3 is restricted to CPS respondents in the MORG. All columns include the following controls: pre*event_{st}, dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quarticage trends, dummies for immigrant status and state fixed effects. Except for column 2, each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table A8: The Effect of Minimum Wages on LFP and EMP

	(1)	(2)	(3)	(4)	(5)
A. LFP					
post _{st}	-0.005 (0.003)	-0.016*** (0.006)	-0.019** (0.009)		
pre*event _{st}		0.005 (0.003)	0.007 (0.009)		0.005* (0.003)
event _{st} × post _{st}		-0.005 (0.005)	0.004 (0.012)		
pre*event _{st} ²			0.000 (0.002)		
event _{st} ² × post _{st}			-0.004** (0.002)		
post*short _{st}				-0.005 (0.004)	-0.015** (0.006)
post*medium _{st}				-0.006 (0.006)	-0.017* (0.010)
B. EMP					
post _{st}	0.001 (0.004)	-0.015* (0.008)	-0.023** (0.010)		
pre*event _{st}		0.008** (0.004)	0.014 (0.012)		0.008** (0.004)
event _{st} × post _{st}		-0.007 (0.006)	-0.002 (0.017)		
pre*event _{st} ²			0.001 (0.002)		
event _{st} ² × post _{st}			-0.005** (0.002)		
post*short _{st}				0.000 (0.005)	-0.015* (0.008)
post*medium _{st}				0.001 (0.007)	-0.015 (0.011)

Notes: All columns include the following controls: dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status, state fixed effects and year fixed effects. The sample is all unmarried high school dropouts between the ages of 20 and 29 that report not being in school full time. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table A9: Effect of the Minimum Wage on Education Choices

	(1)	(2)	(3)	(4)	(5)
	All age 20-29	Women	Men	Immigrants	Natives
$post_{st}$	0.005 (0.004)	0.002 (0.004)	0.008 (0.006)	0.008 (0.006)	-0.001 (0.003)
$pre*event_{st}$	-0.004** (0.002)	-0.003 (0.002)	-0.005*** (0.002)	-0.006** (0.002)	-0.002 (0.001)
N	4,358,749	2,033,268	2,325,481	2,365,780	1,992,969

Notes: The sample in column 1 is all unmarried CPS respondents between the ages of 20 and 29 that report not being in school full time. The sample in column 2 (3) is restricted to women (men). The sample in column 4 (5) is restricted to immigrants (natives). All columns include the following controls: dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status and state fixed effects. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

B Additional appendices (not for publication)

B.1 Alternative definition of the event time variable

In this section, I report estimates of the causal effect of minimum wage legislation on low-skilled labor force participation and employment using an alternate definition of the event time variable. Recall that the event time variable, $event_{st}$, denotes time (in years) relative to a minimum wage increase for state s . I set $event_{st} = 0$ for all state-years when the prevailing nominal minimum wage increases between $t - 1$ and t . For states that experience multiple consecutive minimum wage increases, the variable $event_{st} = 0$ only for the first minimum wage increase. State-years that are τ years before the first minimum wage increase are coded as $event_{st} = -\tau$ and state-years that are τ years after the first minimum wage increase are coded as $event_{st} = \tau$. All state-years more than four years before (resp. four years after) a minimum wage event are coded as $event_{st} = -4$ (resp. $event_{st} = 4$).

Tables B1 and B2 report estimates of α_1 using this approach. Across all specifications, the estimates of the causal effect of the minimum wage on low-skilled labor force participation are nearly identical to those in Tables 4 and 5.

Table B1: Alternative Event Variable: The Effect of Minimum Wages on LFP and EMP

	(1)	(2)	(3)	(4)	(5)
A. LFP (Mean: 0.71)					
post _{st}	-0.005 (0.003)	-0.016** (0.006)	-0.018* (0.010)		
pre*event _{st}		0.005* (0.003)	0.008 (0.009)		0.006* (0.003)
pre*event _{st} ²			0.001 (0.002)		
post*short _{st}				-0.003 (0.006)	-0.014** (0.006)
post*medium _{st}				-0.006 (0.004)	-0.018* (0.009)
B. EMP (Mean: 0.56)					
post _{st}	0.001 (0.004)	-0.015* (0.008)	-0.021* (0.011)		
pre*event _{st}		0.008** (0.003)	0.015 (0.011)		0.009** (0.004)
pre*event _{st} ²			0.001 (0.002)		
post*short _{st}				0.005 (0.008)	-0.011 (0.010)
post*medium _{st}				-0.002 (0.006)	-0.020* (0.010)

Notes: All columns include the following controls: dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status, state fixed effects and year fixed effects. The sample is all unmarried high school dropouts between the ages of 20 and 29 that report not being in school full time. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table B2: Alternative Event Variable Robustness Check: Spatial Controls

	(1)	(2)	(3)	(4)	(5)
A. LFP					
post _{st}	-0.016** (0.006)	-0.012** (0.006)	-0.009 (0.006)	-0.007 (0.008)	-0.006 (0.008)
pre*event _{st}	0.005* (0.003)	0.003 (0.002)	0.003 (0.002)	0.003 (0.003)	0.002 (0.003)
B. EMP					
post _{st}	-0.015* (0.008)	-0.013 (0.009)	-0.009 (0.008)	-0.000 (0.009)	0.005 (0.008)
pre*event _{st}	0.008** (0.003)	0.006* (0.003)	0.006** (0.003)	0.003 (0.003)	-0.000 (0.003)
C. Real Hourly Wage					
post _{st}	0.222 (0.137)	0.161 (0.121)	0.108 (0.117)	0.161 (0.148)	0.199 (0.157)
pre*event _{st}	-0.046 (0.060)	-0.030 (0.056)	0.010 (0.058)	-0.041 (0.060)	-0.084 (0.063)
State-Linear Trend	N	Y	Y	N	N
State-Quadratic Trend	N	N	Y	N	N
Region-Year FE	N	N	N	Y	N
Div.-Year FE	N	N	N	N	Y

Notes: All columns include the following controls: dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status and state fixed effects. The sample is all unmarried high school dropouts between the ages of 20 and 29 that report not being in school full time. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

B.2 Estimates for teenagers

Table B3: Teenagers: The Effect of Minimum Wages on LFP and EMP

	(1)	(2)	(3)
A. LFP (Mean: 0.42)			
post _{st}	0.002 (0.004)	0.002 (0.007)	0.002 (0.006)
pre*event _{st}		0.000 (0.002)	-0.000 (0.005)
pre*event _{st} ²			-0.000 (0.001)
B. EMP (Mean: 0.34)			
post _{st}	0.005 (0.005)	-0.001 (0.007)	-0.003 (0.007)
pre*event _{st}		0.003* (0.002)	0.005 (0.004)
pre*event _{st} ²			0.000 (0.001)
C. Real Hourly Wage (Mean: \$8.42)			
post _{st}	0.144*** (0.046)	0.188** (0.087)	0.398* (0.204)
pre*event _{st}		-0.024 (0.029)	-0.261 (0.173)
pre*event _{st} ²			-0.052 (0.035)

Notes: All columns include the following controls: dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status, state fixed effects and year fixed effects. The sample is all CPS respondents between the ages of 15 and 19. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

B.3 Comparison with Allegretto, Dube, , Reich, and Zipperer (2016)

To facilitate a comparison between the estimates in Tables B4 and B5 below and the elasticity estimates in Table 2 of Allegretto, Dube, , Reich, and Zipperer (2016), simply divide by the mean of the dependent variable. For example, using the estimate in panel A, column 1 of Table B4, the implied elasticity of low-skilled labor force participation with respect to the minimum wage is $-0.049/0.71 = -0.07$. The estimate in panel C, column 1 of Table B5 implies that the elasticity of teenage employment with respect to the minimum wage is $-0.092/0.34 = -0.27$, nearly identical to the -0.219 estimate in Panel B, column 1 of Table 2 in Allegretto, Dube, , Reich, and Zipperer (2016).

Table B4: The Effect of Minimum Wages on LFP and EMP on Low-Skilled Adults

	(1)	(2)	(3)	(4)	(5)
A. LFP (Mean: 0.71)					
$\log(\text{real min. wage})_{st}$	-0.049 (0.034)	-0.018 (0.023)	-0.010 (0.027)	-0.003 (0.035)	0.003 (0.038)
B. LFP (Mean: 0.71)					
$\log(\text{real min. wage})_{st}$	-0.053 (0.035)	-0.017 (0.029)	-0.013 (0.034)	-0.007 (0.037)	0.004 (0.039)
$\text{pre}^*\text{event}_{st}$	0.001 (0.002)	-0.000 (0.002)	0.001 (0.002)	0.001 (0.002)	-0.000 (0.002)
C. EMP (Mean: 0.56)					
$\log(\text{real min. wage})_{st}$	-0.065 (0.047)	0.005 (0.033)	0.034 (0.043)	-0.019 (0.047)	0.019 (0.049)
D. EMP (Mean: 0.56)					
$\log(\text{real min. wage})_{st}$	-0.083* (0.049)	-0.006 (0.037)	0.017 (0.049)	-0.034 (0.050)	0.013 (0.053)
$\text{pre}^*\text{event}_{st}$	0.004*** (0.002)	0.002 (0.002)	0.003 (0.002)	0.003* (0.002)	0.002 (0.002)
State-Linear Trend	N	Y	Y	N	N
State-Quadratic Trend	N	N	Y	N	N
Region-Year FE	N	N	N	Y	N
Div.-Year FE	N	N	N	N	Y

Notes: All columns include the following controls: dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status, state fixed effects and year fixed effects. The sample is all unmarried high school dropouts between the ages of 20 and 29 that report not being in school full time. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table B5: The Effect of Minimum Wages on LFP and EMP on Teenagers

	(1)	(2)	(3)	(4)	(5)
A. LFP (Mean: 0.42)					
$\log(\text{real min. wage})_{st}$	-0.098*** (0.020)	-0.000 (0.023)	0.017 (0.032)	-0.082*** (0.028)	-0.034 (0.039)
B. LFP (Mean: 0.42)					
$\log(\text{real min. wage})_{st}$	-0.111*** (0.023)	-0.002 (0.025)	0.013 (0.033)	-0.090*** (0.029)	-0.037 (0.040)
$\text{pre}^*\text{event}_{st}$	0.003** (0.002)	-0.000 (0.001)	0.001 (0.001)	0.002 (0.001)	0.001 (0.001)
C. EMP (Mean: 0.34)					
$\log(\text{real min. wage})_{st}$	-0.092*** (0.019)	0.015 (0.035)	0.036 (0.047)	-0.065** (0.028)	-0.009 (0.031)
D. EMP (Mean: 0.34)					
$\log(\text{real min. wage})_{st}$	-0.112*** (0.019)	0.004 (0.035)	0.024 (0.045)	-0.007*** (0.028)	-0.013 (0.032)
$\text{pre}^*\text{event}_{st}$	0.005*** (0.002)	0.002* (0.001)	0.002*** (0.001)	0.003*** (0.001)	0.001 (0.001)
State-Linear Trend	N	Y	Y	N	N
State-Quadratic Trend	N	N	Y	N	N
Region-Year FE	N	N	N	Y	N
Div.-Year FE	N	N	N	N	Y

Notes: All columns include the following controls: dummies for race (white and black, the excluded category is hispanic/other), linear, quadratic, cubic and quartic age trends, dummies for immigrant status, state fixed effects and year fixed effects. The sample is all unmarried high school dropouts between the ages of 20 and 29 that report not being in school full time. Each observation is weighted by its inverse sampling probability provided in the CPS. Standard errors are clustered at the state level. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.