Minimum Wages and Optimal Redistribution*

Damián Vergara
UC Berkeley

Job Market Paper

Updated frequently, click here for the latest version

Abstract

This paper analyzes whether a minimum wage should be used for redistribution on top of taxes and transfers. I characterize optimal redistribution for a government with three instruments – labor income taxes and transfers, corporate income taxes, and a minimum wage – using an empirically grounded model of the labor market with positive firm profits. I find that a minimum wage helps the government redistribute efficiently to low-skill workers when it increases average post-tax wages of low-skill labor market participants and when corporate profit incidence is large. The minimum wage can prevent firms from capturing low-wage income subsidies such as the EITC and from enjoying high profits that cannot be redistributed via corporate taxes due to capital mobility in unaffected industries. Event studies show that the average US state-level minimum wage reform over the last two decades increased average post-tax wages of low-skill labor market participants and reduced corporate profits in affected industries, namely low-skill labor intensive services. A sufficient statistics analysis implies that US minimum wages typically remain below their optimum under the current tax and transfer system.

*Email: damianvergara@berkeley.edu. First version: September, 2021. This version: October 31, 2022. I am especially grateful to Danny Yagan, Pat Kline, and Emmanuel Saez, for continuous guidance and encouragement throghout this project. I also thank Alan Auerbach, Nano Barahona, Sydnee Caldwell, David Card, Patricio Domínguez, Cecile Gaubert, Andrés González, Attila Lindner, Cristóbal Otero, Pablo Muñoz, Michael Reich, Marco Rojas, Jesse Rothstein, Harry Wheeler, Gabriel Zucman, my discussants Benjamin Glass and Thomas Winberry, and seminar participants at the Online Public Finance Workshop for Graduate Students, the NTA 114th Annual Conference on Taxation, the XIII RIDGE Forum Workshop on Public Economics, UC Berkeley, and Universidad Adolfo Ibáñez for very helpful discussions and suggestions. Early versions of this project benefited from discussions with Jakob Brounstein, Sree Kancherla, Maximiliano Lauletta, Michael Love, Billy Morrison, and Mónica Saucedo. I acknowledge financial support from the Center of Equitable Growth at UC Berkeley. This paper previously circulated under the title "When do Minimum Wages Increase Social Welfare? A Sufficient Statistics Analysis with Taxes and Transfers". Usual disclaimers apply.

1 Introduction

Governments use income taxes and transfers to redistribute to low-income individuals. However, those taxes and transfers can be distortionary, yielding an equity-efficiency tradeoff (Mirrlees, 1971; Piketty and Saez, 2013). This paper asks whether a minimum wage can relax that tradeoff and enable more efficient redistribution than income taxes and transfers alone.

Economists have long debated this question. Mill (1884) suggested that a minimum wage was the simplest way to redistribute profits to raise the incomes of low earners but Stigler (1946) articulated the natural economic argument that a minimum wage is inefficient relative to income-based taxes and transfers given its effects on employment. More recently, attempts to formally address this question have provided mixed answers by using frameworks that fail to incorporate empirically relevant channels through which minimum wages can perform redistribution (Hungerbühler and Lehmann, 2009; Lee and Saez, 2012; Cahuc and Laroque, 2014; Lavecchia, 2020). Empirically, a growing literature has found that minimum wages increase incomes of low earners with limited reductions in employment (Lee, 1999; Autor et al., 2016; Cengiz et al., 2019; Dube, 2019; Fortin et al., 2021; Manning, 2021), possibly accompanied by reduced corporate profits (Draca et al., 2011; Harasztosi and Lindner, 2019; Drucker et al., 2021). Yet, even if minimum wages redistribute from high-earning capitalists to low-earning workers, it remains unsettled whether such redistribution is preferred over analogous redistribution via corporate income taxes and income-based transfers alone, or to what extent there are interactions between minimum wages and income taxes that help governments to redistribute more efficiently when using all instruments together.

This paper proposes a novel theoretical framework to analyze the redistributive role of the minimum wage when taxes and transfers are also available for the policymaker. Formally, I characterize optimal redistribution for a social planner with three instruments: labor income taxes and transfers, corporate income taxes, and a minimum wage. The analysis illustrates the channels through which the minimum wage affects the distribution of income, explicitly describing its tradeoffs and interactions with the tax system. Results are expressed as a function of sufficient statistics, meaning that the main theoretical objects have empirical counterparts that can be estimated in data. The sufficient statistics feature, which is illustrated in an empirical exercise using publicly available US data, provides a direct connection between theory and evidence in the optimal minimum wage analysis.

The theoretical framework uses directed search and two-sided heterogeneity to allow for three potentially relevant features regarding the use of a minimum wage: the possibility of limited employment effects, spillovers to non-minimum wage jobs, and positive firm profits. On one side, there is a population of workers with heterogeneous skills and costs of participating in the labor market that decide whether to enter the labor market and on which jobs to apply to. On the other side, there is a population of capitalists with heterogeneous productivities that decide whether to create firms and the vacancies and wages they post. In the model, minimum wage changes affect workers' application strategies which, in turn, affect the posting behavior of firms. These behavioral responses can lead to limited employment effects and spillovers to non-minimum wage jobs. The model features positive profits in equilibrium and reproduces other empirically relevant characteristics of labor markets such as wage dispersion for similar workers (Card et al., 2018) and finite firm-specific labor supply elasticities (Sokolova and Sorensen, 2021).

I use this model of the labor market to characterize optimal policy when a social planner chooses the minimum wage, the labor income tax system, and the corporate tax rate to maximize social welfare taking as given social preferences for redistribution. Considering first a case with no taxes and transfers, the minimum wage affects the welfare of active low- and high-skill workers through its effects on equilibrium wages and employment probabilities, and the welfare of capitalists through its effects on profits. This implies that the minimum wage can affect both the relative welfare within labor income earners and between labor and capital income earners. The change in workers' welfare is summarized by the change in the expected utility of participating in the labor market which, under the assumptions of the model, equals the change in the average wage of labor market participants including the unemployed. This sufficient statistic aggregates all wage, employment, and participation responses that can affect workers' utility in a single elasticity. While the sign of the workers' sufficient statistic is theoretically ambiguous, welfare improvements for workers are not necessarily tied to positive employment effects. The elasticity can be used to compute, given wage effects, the disemployment effects that can be tolerated for the minimum wage to have a net positive effect on workers' welfare.

When income taxes and transfers are present, the optimal minimum wage depends additionally on fiscal externalities from both sides of the market. On the workers' side, changes in wages and employment induce a change in income tax collection and transfer spending. On the firms' side, the change in profits affects corporate tax revenue. The optimal minimum wage increases when the corporate tax rate is low, both because the revenue loss is smaller and the welfare gains from redistributing from capitalists to workers are higher. When the planner simultaneously chooses the tax system and the minimum wage, minimum wages can be desirable because they can make tax-based redistribution more efficient. To illustrate why, suppose that the optimal tax schedule considers an EITC to redistribute to low-wage workers. If firms internalize the effects of the EITC, they react by lowering pre-tax wages (Rothstein, 2010). The minimum wage prevents firms from decreasing wages, thereby increasing the efficacy of the EITC by making it work as a lump-sum transfer (Lee and Saez, 2012). The fiscal benefit of the minimum wage is accompanied by an employment distortion that is negligible at low levels but growing as the minimum wage departs from the market wage, generating a tradeoff for the planner. When corporate taxes distort pre-tax profits, they cannot fully correct this incidence distortion, suggesting that combining minimum wages and corporate taxes is possibly optimal for taxing profits.

I then provide empirical estimates of the sufficient statistics that guide the policy analysis exploiting

US state-level variation in minimum wages. I follow Cengiz et al. (2019, 2022) and estimate stacked event studies, where events are defined as state-level hourly minimum wage increases of at least $0.25 (in 2016 dollars) in states where at least 2% of the pre-event year working population earned less than the new minimum wage and where treated states did not experience other relevant minimum wage increases in the pre-event window. I identify 50 events in the period 1997-2019 for which the outcomes of interest are observed through an eight-year balanced window.

The data consists of yearly state-level aggregates of different outcomes of interest. To measure workers' outcomes, I follow Cengiz et al. (2019, 2022) and use the individual-level NBER Merged Outgoing Rotation Group of the CPS to compute average pre-tax hourly wages and the Basic CPS monthly files to compute employment and participation rates. I combine these data sources to compute pre-tax versions of the worker-level sufficient statistics at the skill-by-state-by-year level, where low- and high-skill workers are broadly defined by having or not a college degree. Novel to this paper, to estimate workers' side fiscal externalities, I use data on income maintenance benefits, medical benefits, and gross federal income taxes taken from the BEA regional accounts. Also novel to this paper, to measure capitalists' outcomes, I proxy state-level average profits using the gross operating surplus estimates from the BEA regional accounts and normalize them by the average number of private establishments reported in the QCEW data files. I combine both data sources to compute average pre-tax profits per establishment at the industry-by-state-by-year level.

The empirical results imply that minimum wages have increased low-skill workers' welfare with an estimated elasticity of around 0.1. Conversely, all specifications estimate a precise zero elasticity for the high-skill workers' analog. Results for low-skill workers are stable across demographic groups, suggesting that the welfare gains are not concentrated on particular groups of "winners". If anything, teens (aged 16-19) and black low-skill workers seem to experience larger welfare gains from minimum wage increases, but no group experiences welfare losses. Consistent with Cengiz et al. (2019, 2022), decomposing the sufficient statistic of low-skill workers across different margins shows that the entire effect is driven by an increase in the wage conditional on employment. No effect is found on hours, employment, or participation.

The estimated elasticity of income maintenance benefits to minimum wage changes ranges between -0.31 and -0.39, suggesting sizable fiscal externalities on the workers' side of the market. This result is consistent with Reich and West (2015), who find elasticities of around -0.2 for SNAP expenditures, and Dube (2019), who finds that after-tax income elasticities are one third smaller than their pre-tax analogs. These fiscal externalities attenuate the welfare gains for low-skill workers. I find no effects on medical transfers and gross federal income tax liabilities, suggesting that most of the worker-level fiscal effects are mediated by targeted transfers based on pre-tax income levels.

When looking at pre-tax profits per establishment, the estimated elasticity is zero when pooling all industries. However, I find a clear negative effect in "exposed industries" with large numbers of minimum wage workers (retail trade, low-skill health services, food, accommodation, and social services) with an estimated elasticity of -0.36. No association is found between profits and minimum wage events in manufacturing and in non-exposed service industries. The effect on the number of establishments in the exposed services is negligible, suggesting that changes in average pre-tax profits are driven by the intensive margin. The estimated effect of pre-tax profits generates a fiscal externality in terms of corporate tax revenue that is not significantly present in other sources of capital income revenue such as business and dividend income reported in the SOI state-level tables.

To interpret the estimates through the lens of the optimal policy analysis, I plug-in the empirical estimates on the theoretical results to assess the desirability of small minimum wage increases under different calibration choices. In the absence of preferences for redistribution, increasing the minimum wage is close to being welfare-neutral. However, increasing the minimum wage generates substantial welfare gains when incorporating distributional concerns between workers and capitalists. Intuitively, results show that minimum wages help low-skill workers, hurt firm owners in the exposed industries, and generate fiscal savings in income transfers but fiscal costs in terms of corporate tax revenue. Total gains for low-skill workers are comparable to total losses for capitalists, and the net fiscal effect is approximately zero. However, average post-tax profits per capitalist are between five and six times larger than average per capita post-tax wages of active low-skill workers. As a result, when including social preferences for redistribution, the gains for workers substantially outweigh the losses for capitalists. Hence, results suggest that, under existing tax and transfers systems, the average past increase in state-level minimum wages has increased social welfare, and increasing the minimum wage today would do so as well.

I close the paper with three policy applications, beginning with the interaction of the minimum wage and the corporate tax rate under international tax competition. Corporate taxes may optimally be low because capital in manufacturing can flow to lower-tax countries (Devereux et al., 2008, 2021). A minimum wage with incidence on manufacturing could similarly cause capital to flow to lower-tax countries. However, the empirical analysis finds that the minimum wage predominantly affects non-tradable services industries. As a result, I show in a multi-sector extension of the model that the minimum wage can be desirable as a kind of industry-specific corporate tax that affects immobile service firms, while leaving official corporate taxes low for unaffected mobile manufacturing firms. Formally, the welfare benefits of increasing the minimum wage are larger when the capital in non-affected industries – e.g., manufacturing – is more mobile because the optimal corporate tax rate is lower.

In the second policy application, I explore with more detail the interaction between the minimum wage and the EITC. I show that under the optimal minimum wage, the marginal tax rate on low-skill workers is negative if the social planner values redistribution toward low-skill workers, internalizing the congestion externalities generated by the policy. This extends Lee and Saez (2012) result on the complementarity between the EITC and the minimum wage using a framework where labor markets are not perfectly competitive and firms have positive profits in equilibrium.

The final policy application presents a suggestive numerical exercise that illustrates the intuitions developed throughout the paper. I calibrate a simplified version of the model to understand how does the optimal minimum wage change under different given tax systems. Social welfare is larger when all instruments are used together. Intuitively, the social planner needs to collect revenue, which is mainly done using the corporate tax. To mitigate the corporate tax distortions, the social planner increases labor supply to firms by complementing the corporate tax with negative marginal tax rates on workers. To make this plan work more efficiently, the social planner also complements this package with binding minimum wages that prevent firms to passthrough the EITC into lower pre-tax wages.

The general message is that social planners should not make the tax system and the minimum wage compete for who is the most efficient redistributive policy. By contrast, social planners can benefit from using all instruments together to make redistribution more efficient. Optimal redistribution possibly consists on a binding minimum wage, a corporate tax rate, and a targeted EITC.

Related literature and contributions The main contribution of this paper is to the normative analysis of the minimum wage in frameworks with taxes and transfers. Previous literature abstracts from firm profits, firm-level heterogeneity, and corporate taxation, and base the analysis on labor market models that do not explicitly accommodate empirically relevant general equilibrium effects of the minimum wage that can dampen employment impacts. Lee and Saez (2012) use a competitive supply-demand framework and show that the case for binding minimum wages under optimal taxes depends on labor rationing assumptions. Cahuc and Laroque (2014) contest Lee and Saez (2012)'s result by arguing that the minimum wage cannot improve welfare on top of an optimal non-linear tax schedule even if the labor demand is modeled as a standard monopsonist. Both analyses abstract from search frictions, firm-level heterogeneity, and do not give a central role to firm profits. Hungerbühler and Lehmann (2009) and Lavecchia (2020) consider random search models but also abstract from firm profits, restricting the role of the minimum wage under optimal taxes to solving search and matching inefficiencies.

A complementary literature studies the welfare consequences of the minimum wage using structural models that abstract from the tax-design question. Some papers also abstract from the distributional dimension and focus on efficiency rationales motivated by labor market imperfections (Flinn, 2006; Wu, 2021; Ahlfeldt et al., 2022; Drechsel-Grau, 2022). Two recent papers give an important role to redistribution within the analysis. Berger et al. (2022) propose a general equilibrium model of oligopsonistic labor markets and find that welfare improvements from minimum wage increases stem mainly from redistribution because reductions in labor market power can simultaneously generate misallocation as large-productive firms increase their market shares. Consistent with my analysis, they find that the main distributional benefits come from redistributing from capitalists to low-skill workers. Hurst et al.

(2022) develop a general equilibrium model to compare the short- and long-run impacts of the minimum wage, finding that the minimum wage encourages capital-labor substitution in the long-run. They argue this generates unintended distributional consequences on low-skill workers that are displaced by capital. Their results favor the tax system but also suggest that a moderate minimum wage can improve the efficacy of the EITC by setting a wage floor to firms. My analysis differs from theirs since I focus on the optimal policy design and its short- and medium-run –rather than long-run– consequences. I also explicitly discuss the interactions between the minimum wage and the corporate tax policy.[1]

In terms of optimal redistribution, this paper adds to the literature that explores whether the combination of different instruments can improve the efficiency of the tax system. Seminal examples include Diamond and Mirrlees (1971), who study the interaction between production and consumption taxes, and Atkinson and Stiglitz (1976) and Saez (2002), who study the interaction between income and commodity taxation. Recent papers include Gaubert et al. (2020), Ferey (2022), and Ferey et al. (2022). This paper also adds to the analysis of redistributive policies in labor markets with frictions (Hungerbühler et al., 2006; Stantcheva, 2014; Sleet and Yazici, 2017; Kroft et al., 2020; Bagger et al., 2021; Hummel, 2021b; Mousavi, 2021; Craig, 2022; Doligalski et al., 2022), and to the analysis of redistribution between capital and labor income (Atesagaoglu and Yazici, 2021; Eeckhout et al., 2021; Hummel, 2021a).

Finally, the empirical results add to a large literature that studies the effects of minimum wages on different outcomes. The workers' side results at the skill-level complement the vast literature that studies the effects on wages and employment (see Manning, 2021 for a survey). Results on income maintenance transfers and other fiscal outcomes complement the evidence presented in Reich and West (2015) and Dube (2019). Finally, the results on profits are in line with the findings of Draca et al. (2011), Harasztosi and Lindner (2019), and Drucker et al. (2021) and are, to my knowledge, the first using US data.

2 Model of the Labor Market

This section develops a model of the labor market with positive firm profits that can accommodate limited employment effects and spillovers to non-minimum wage jobs after minimum wage increases.

2.1 Setup: workers and capitalists

Overview The model is static and features two-sided heterogeneity. On one side, there is a population of workers that is heterogeneous in two dimensions. First, workers have different skills. For simplicity, I assume workers are either low-skill or high-skill. Second, workers have heterogeneous costs of participating in the labor market. On the other side, there is a population of capitalists with heterogeneous

[1] Dworczak et al. (2021) indirectly analyzes the redistributive consequences of the minimum wage using an alternative framework. They analyze redistribution through markets and price controls using mechanism design techniques.

productivities. Labor market interactions are modeled following a directed search approach (Moen, 1997). Capitalists decide whether to create firms based on expected profits. Conditional on creating a firm, they post wages and vacancies, with all vacancies posted at a given wage forming a *sub-market.*[2] Labor markets are segmented, meaning that wages and vacancies are skill-specific. Workers observe wages and vacancies and make their labor market participation and application decisions. In equilibrium, there is a continuum of sub-markets indexed by m, characterized by skill-specific wages, w_m^s, vacancies, V_m^s, and applicants, L_m^s, with $s \in \{l, h\}$ indexing skill.

Matching technology There are standard matching frictions within each sub-market. The number of matches within a sub-market is given by the matching function $\mathcal{M}^s(L_m^s, V_m^s)$, with $\mathcal{M}^s$ continuously differentiable, increasing and concave, and with constant returns to scale. The matching technology is allowed to be different for low- and high-skill workers (Berman, 1997; Hall and Schulhofer-Wohl, 2018).

Under these assumptions, the sub-market skill-specific job-finding rate (that is, the workers' probability of finding a job conditional on applying) can be written as

$$p_m^s = \frac{\mathcal{M}^s(L_m^s, V_m^s)}{L_m^s} = \mathcal{M}^s(1, \theta_m^s) = p^s(\theta_m^s), \tag{1}$$

with $\partial p^s(\theta_m^s)/\partial \theta_m^s \equiv p_\theta^s > 0$, where $\theta_m^s = V_m^s / L_m^s$ is the sub-market skill-specific vacancies to applicants ratio, also denoted as *sub-market tightness*. Intuitively, the higher the ratio of vacancies to applicants, the more likely that an applicant will be matched with one of those vacancies.

Likewise, the sub-market skill-specific job-filling rate (that is, the firms' probability of filling a vacancy conditional on posting it) can be written as

$$q_m^s = \frac{\mathcal{M}^s(L_m^s, V_m^s)}{V_m^s} = \mathcal{M}^s\left(\frac{1}{\theta_m^s}, 1\right) = q^s(\theta_m^s), \tag{2}$$

with $\partial q^s(\theta_m^s)/\partial \theta_m^s \equiv q_\theta^s < 0$. Intuitively, the lower the ratio of vacancies to applicants, the more likely that the firm will be able to fill the vacancy with a worker.

I assume that neither workers or firms internalize that their application and posting behavior affects equilibrium tightness, so they take p_m^s and q_m^s as given when making their individual decisions.

Workers The population of workers is normalized to 1. The exogenous shares of low- and high-skill workers are given by α_l and α_h, respectively. Conditional on skill, each worker draws a parameter $c \in \mathcal{C} = [0, C] \subset \mathbb{R}$ that represents the cost of participating in the labor market, which admits different interpretations such as search costs, disutility of (extensive margin) labor supply, or other opportunity

[2] The notion of sub-market should not be confounded with the notion of local labor market. Sub-markets only vary with wages and, in principle, all workers are equally able to apply to them. Both concepts could be closer in a more general model with multidimensional firm heterogeneity and heterogeneous application costs.

costs of working such as home production. Let f_s and F_s be the skill-specific density and cumulative distributions of c, respectively, assumed to be smooth.

Workers derive utility from the after-tax wage net of labor market participation costs.[3] The utility of not entering the labor market is $u_0 = y_0$, where y_0 is a lump-sum transfer paid by the government to non-employed individuals. u_0 is the same for all workers, regardless of their (s, c) type. When entering the labor market, workers apply to jobs. Following Moen (1997), I assume that workers can apply to jobs in only one sub-market.[4] Conditional on employment, after-tax wages in sub-market m are given by $y_m^s = w_m^s - T(w_m^s)$, where T is the (possibly non-linear) income tax-schedule, with $T(0) = -y_0$. Then, the expected utility of entering the labor market for a worker of type (s, c) is given by

$$u_1(s,c) \quad = \quad \max_m \{p_m^s y_m^s + (1-p_m^s)y_0\} - c, \tag{3}$$

since workers apply to the sub-market that gives them the highest expected after-tax wage internalizing that the application ends in employment with probability p_m^s and unemployment with probability $1-p_m^s$.

Recall that p_m^s depends on the mass of workers of skill s that apply to jobs in sub-market m: given a stock of vacancies, the more workers apply, the smaller the likelihood of being employed. Then, individuals take p_m^s as given but it is endogenously determined by the aggregate application behavior. This implies that, in equilibrium, all markets have the same expected after-tax wage, i.e., $p_i^s y_i^s + (1-p_i^s)y_0 = p_j^s y_j^s + (1-p_j^s)y_0 = \max_m\{p_m^s y_m^s + (1-p_m^s)y_0\}$, for all i, j; if not, workers have incentives to change their applications toward markets with higher expected values, pushing downward the job-filling probabilities and restoring the equilibrium. This means that workers face a trade-off between wages and employment probabilities because it is more difficult to get a job in sub-markets that pay higher wages.[5]

In what follows, I define $U^s \equiv \max_m\{p_m^s y_m^s + (1-p_m^s)y_0\}$ so $u_1(s,c) = U^s - c$. The labor market participation decision is given by $l(s,c) = 1\{u_1(s,c) \geq u_0\} = 1\{U^s - y_0 \geq c\}$. This implies that $l(s,c) = 1$ if $c \leq U^s - y_0$, $l(s,c) = 0$ otherwise. Let $L_A^s = \alpha_s \cdot \int l(s,c)dF_s(c)$ denote the mass of active workers of skill s, that is, the mass of workers of skill s that enter the labor market. Then, $L_A^s = \alpha_s \cdot F_s(U^s - y_0)$. Inactive workers are given by $L_I = L_I^l + L_I^h = 1 - L_A^l - L_A^h$. Denote by L_m^s the mass of individuals of skill s applying to jobs in sub-market m, so $L_A^s = \int L_m^s dm$. I assume away sorting patterns based on c, that is, application decisions conditional on participating in the labor market are independent from c.

Note that the expression $U^s = p_m^s y_m^s + (1-p_m^s)y_0$ implies that θ_m^s can be written as a function of w_m^s and U^s, for all m (Moen, 1997). Formally, $\theta_m^s = \theta_m^s(w_m^s, U^s)$, with $\partial\theta_m^s/\partial w_m^s < 0$ and $\partial\theta_m^s/\partial U^s > 0$.[6]

[3] Since the model abstracts from intensive margin responses, I refer to wages, incomes, or earnings indistinctly.

[4] See Kircher (2009) and Wolthoff (2018) for models where workers can simultaneously apply to several sub-markets.

[5] The model assumes that workers are risk-neutral. Incorporating concavity in the flow utility function does not affect the high-level analysis, but induces a complication when defining the empirical approximations of the relevant sufficient statistics derived in the optimal policy analysis. I come back to this discussion in the next section.

[6] Since $U^s = p^s(\theta_m^s) \cdot (w_m^s - T(w_m^s)) + (1 - p^s(\theta_m^s)) \cdot y_0$, then $dU^s = p_\theta^s \cdot d\theta_m^s \cdot y_m^s + p_m^s \cdot (1 - T'(w_m^s)) \cdot dw_m^s$. Recalling

This result simplifies the analysis below since implies that, conditional on wages, equilibrium behavior can be summarized by the scalars U^s without needing to characterize the continuous sequence of θ^s_m.

Capitalists The population of capitalists is normalized to K. Each capitalist draws a parameter $\psi \in \Psi = \left[\underline{\psi}, \overline{\psi}\right] \subset \mathbb{R}^+$ that represents firm productivity. Let o and O be the density and cumulative distributions of ψ, respectively, assumed to be smooth.

Capitalists observe ψ and choose whether to create a firm. Firms are price-takers in the output market (with the price normalized to 1). Technology is assumed to depend on ψ, low- and high-skill workers, and the (flat) corporate tax rate, t, so a firm of productivity ψ that hires (n^l, n^h) workers generates revenue equal to $\phi(\psi, n^l, n^h, t)$, with ϕ twice differentiable, $\phi_\psi > 0$, $\phi_{n^s} > 0$ and $\phi_{n^s n^s} \leq 0$, $\phi_t \leq 0$, and $\phi_{tn^s} \leq 0$, for $s \in \{l, h\}$. Allowing the revenue function to depend on t accomodates, in a reduced-form fashion, the fact that corporate tax rates can distort pre-tax profits (Kennedy et al., 2022).[7]

Firms choose skill-specific wages, w^s, and vacancies, v^s, internalizing that n^s is the result of the matching process. While firms take the job-filling probabilities as given, they internalize that paying higher wages increases the job-filling probabilities. In other words, the wage choice is equivalent to the sub-market choice. I rewrite job-filling probabilities as $\tilde{q}^s(w^s, U^s) = q(\theta^s(w^s, U^s))$, with $\tilde{q}^s_w = q^s_\theta \cdot (\partial\theta^s/\partial w^s) > 0$, so $n^s = \tilde{q}^s(w^s) \cdot v^s$. Posting v^s vacancies has a cost $\eta^s(v^s)$, with $\eta^s_v > 0$ and $\eta^s_{vv} > 0$. Then, pre-tax profits are given by revenue net of labor costs:

$$\begin{aligned}\pi\left(w^l, w^h, v^l, v^h; \psi, t\right) \quad &= \quad \phi\left(\psi, \tilde{q}^l(w^l, U^l) \cdot v^l, \tilde{q}^h(w^h, U^h) \cdot v^h, t\right) \\ &\quad - \left(w^l \cdot \tilde{q}^l(w^l, U^l) \cdot v^l + \eta^l(v^l)\right) - \left(w^h \cdot \tilde{q}^h(w^h, U^h) \cdot v^h + \eta^h(v^h)\right). \qquad (4)\end{aligned}$$

Denote the value function by $\Pi(\psi, t) = \max_{w^l, w^h, v^l, v^h} \pi\left(w^l, w^h, v^l, v^h; \psi, t\right)$. Then, after-tax profits are given by $(1 - t) \cdot \Pi(\psi, t)$.

Conditional on ψ, firms are homogeneous. Then, the solution to the profit maximizing problem can be characterized by functions $w^s(\psi)$ and $v^s(\psi)$. Appendix A derives the first-order conditions and shows that dispersion in productivities leads to dispersion in wages, with wages *marked down* relative to the marginal productivities. m indexes sub-markets as well as the productivity levels of capitalists that create firms, so $w^s_m = w^s(\psi_m)$, $v^s_m = v^s(\psi_m)$, and $V^s_m = K \cdot v^s(\psi_m) \cdot o(\psi_m)$.

Capitalists pay a fixed cost, ξ, to create firms, and receive the lump-sum transfer, y_0, when remaining inactive, so they create firms when $(1 - t) \cdot \Pi(\psi, t) \geq \xi + y_0$. Since profits are increasing in productivity, the entry rule defines a productivity threshold, ψ^*, implicitly determined by $(1 - t) \cdot \Pi(\psi^*, t) = \xi + y_0$

that $p^s_\theta > 0$ and asssuming $T'(w^s_m) < 1$ yields the result.

[7] In Appendix A, I propose two different microfoundations of ϕ that generate dependence on t: a capital allocation problem, where capitalists have a fixed stock of capital that have to allocate between the domestic firm and an international outside option, and an effort allocation problem, where revenue also depends on the owners' effort. The standard assumption of non-distortionary t arises as a particular case. The same arguments may lead ϕ to also depend on the minimum wage.

such that capitalists create firms only if $\psi \geq \psi^*$. Then, the mass of active capitalists is given by $K_A = K \cdot (1 - O(\psi^*))$. The mass of inactive capitalists, K_I, is given by $K_I = K \cdot O(\psi^*)$, with $K_A + K_I = K$.

2.2 Discussion

Before introducing a minimum wage to the model, I discuss some features and limitations of the proposed framework. This is a non-exhaustive discussion which is continued in Appendix A.

Directed search Directed search models usually generate efficient outcomes in terms of search and posting behavior (Moen, 1997; Wright et al., 2021). That is, these models don't exhibit inefficient excess or lack of applicants or vacancies, as can happen, for example, in random search models (Hosios, 1990; Mangin and Julien, 2021). In Appendix A I show that the proposed model maintain these properties.[8] This feature fosters a focus on the redistributive properties of the minimum wage rather than on its efficiency rationales (e.g., Burdett and Mortensen, 1998; Acemoglu, 2001).

Monopsony power While search and posting behavior is efficient, the model admits monopsony power through wage-dependent job-filling probabilities that have a similar flavor to the standard monopsony intuition of upward-slopping firm-specific labor supply curves (Robinson, 1933; Card et al., 2018) supported by recent empirical evidence (Staiger et al., 2010; Azar et al., 2019; Dube et al., 2020; Sokolova and Sorensen, 2021; Bassier et al., 2022). Firms internalize that paying higher wages lead to more applicants, so wages are *marked down* relative to marginal productivities.[9]

Low-wage labor markets The equilibrium of the model is consistent with other stylized facts of low-wage labor markets. The model features wage dispersion for similar workers (Card et al., 2018), wage posting rather than bargaining, which has been found to be more relevant for low-wage jobs (Hall and Krueger, 2012; Caldwell and Harmon, 2019; Lachowska et al., 2022), and can rationalize bunching in the wage distribution at the minimum wage (Cengiz et al., 2019).

Restricted heterogeneity One limitation of the model is that the dimensions of worker- and firm-level heterogeneity are limited. On the workers' side, the model assumes that all workers within skill type get the same expected utility. As suggested by Hurst et al. (2022), this could mask important distributional effects if there are winners and losers within skill-type after minimum wage changes. Extending the model in this direction would imply that U^s –which will play an important role in the optimal policy analysis–

[8] This result can be thought of as an extension of Moen (1997) result to a setting with ex-ante firm-level heterogeneity and positive profits. I show that not only posting is efficient, but also the entry thresholds at the worker- and firm-levels.

[9] Appendix A shows that the standard markdown equation can be derived from the firm's first order conditions.

can be different for different groups of low- and high-skill workers.[10] I come back to this discussion in Section 4 where I empirically test for heterogeneities in the estimated welfare changes within skill groups.

On the firms' side, one-dimensional heterogeneity is a useful simplification for notational purposes. Extending the model to multidimensional heterogeneity, $(\psi, \widetilde{\psi})$, is straightforward as long as workers do not have preferences for these attributes. This could accommodate, for example, variation in factor shares. In this setting, the problem's solution would be given by wage and vacancy functions $w^s(\psi, \widetilde{\psi})$ and $v^s(\psi, \widetilde{\psi})$, and by a set of conditional productivity thresholds, $\psi^*(\widetilde{\psi})$. Such an extension adds little intuition to the general policy analysis while introducing more complicated notation. Since the optimal policy results are based on reduced-form profit elasticities, the empirical analysis will be able to test for richer forms of firm-level heterogeneity.

The previous argument requires workers to not have preferences over $\widetilde{\psi}$ beyond its effect on wages and vacancies. Then, the simple extension to multidimensional heterogeneity does not apply to non-wage amenities.[11] Amenities can affect the policy analysis because of two reasons. First, if workers rank firms using a composite index of expected wages and amenities and the latter are not taxed, then the tax system can distort workers' preferences (Lamadon et al., 2022). Second, if amenities are endogenous, minimum wage increases may induce firms to worsen the non-wage attributes of the job (Clemens et al., 2018; Clemens, 2021). This could attenuate potential welfare gains to workers after minimum wage hikes.

2.3 Introducing a minimum wage

I introduce a minimum wage, $\overline{w}$, to explore how the predictions of the model speak to the related empirical literature. I separately explore the effects on workers and capitalists decisions.

Low-skill workers Recall that, in equilibrium, $U^l = p^l(\theta^l_m) \cdot y^l_m + (1 - p^l(\theta^l_m)) \cdot y_0$, for all sub-markets m. Let i be the sub-market constrained by the minimum wage, so $w^l_i = \overline{w}$. Differentiating yields

$$\frac{dU^l}{d\overline{w}} \quad = \quad p^l_\theta \cdot \frac{d\theta^l_i}{d\overline{w}} \cdot (\overline{w} - T(\overline{w}) - y_0) + p^l(\theta^l_i) \cdot (1 - T'(\overline{w})). \tag{5}$$

Since $p^s(\theta^l_i) > 0$, and assuming $T'(\overline{w}) < 1$, $dU^l/d\overline{w} = d\theta^l_i/d\overline{w} = 0$ is not a feasible solution to equation (5). This implies that changes in $\overline{w}$ necessarily affect the equilibrium values of U^l, θ^l_i, or both.

Intuitively, an increase in the minimum wage mechanically makes minimum wage jobs more attractive for low-skill workers. This effect is captured by $p^l(\theta^l_i)\cdot(1-T'(\overline{w}))$: the increase in the attractiveness of this sub-market is the net-of-tax gain conditional on working, $1 - T'(\overline{w})$, times the employment probability,

[10] Formally, consider an additional variable, $\widetilde{s}$ such that $U^{s,\widetilde{s}_1} \neq U^{s,\widetilde{s}_2}$. This could be the case if, for example, workers of type $(s, \widetilde{s}_1)$ can apply to a different pool of firms than workers of type $(s, \widetilde{s}_2)$.

[11] For evidence on their importance, see Bonhomme and Jolivet (2009), Mas and Pallais (2017), Maestas et al. (2018), Sorkin (2018), Taber and Vejlin (2020), Jäger et al. (2021), Le Barbanchon et al. (2021), Lindenlaub and Postel-Vinay (2021), Sockin (2021), Lamadon et al. (2022), and Roussille and Scuderi (2022).

$p^l(\theta_i^l)$. This attracts new applicants toward minimum-wage sub-markets (from other sub-markets and/or from outside the labor force), thus pushing θ_i^l downwards until the across sub-market equilibrium is restored. This decreases the employment probability in sub-market i, whose effect is captured by the change in the employment probability, $p_\theta^l \cdot (d\theta_i^l/d\overline{w})$, times after-tax income conditional on employment, $\overline{w} - T(\overline{w}) - y_0$. These two effects resemble the standard intuition around minimum wages: they affect wages and employment. How these effects balance determine the overall effect on expected utility.

This mechanic is the essence of the general equilibrium effects of the model: the initial change in applications toward minimum-wage jobs triggers a sequence of reactions that reconfigure labor market outcomes. Changes in $\overline{w}$ also affect the equilibrium of unconstrained low-skill sub-markets. To see this, let j be a sub-market that is not constrained by the minimum wage, so $w_j^l > \overline{w}$. Differentiating yields

$$\frac{dU^l}{d\overline{w}} \quad = \quad p_\theta^l \cdot \frac{d\theta_j^l}{d\overline{w}} \cdot (w_j^l - T(w_j^l) - y_0) + p^l(\theta_j^l) \cdot (1 - T'(w_j^l)) \cdot \frac{dw_j^l}{d\overline{w}}. \tag{6}$$

Equation (5) suggests that the left-hand-side of equation (6) is unlikely to be zero, implying that θ_j^l or w_j^l or both are possibly affected by changes in the minimum wage. There are two forces that mediate this spillover. First, the change in applicant flows between sub-markets and from in and out of the labor force affects the employment probabilities of all sub-markets until the equilibrium condition of equal expected utilities is restored. This is captured by the first term of equation (6). Second, as I discuss below, firms can also respond to changes in applicants. The potential wage response is captured in the second term of equation (6) and changes in vacancy posting implicitly enter the terms $d\theta_m^l/d\overline{w}$ of equations (5) and (6).

Changes in U^l also induce changes in labor market participation, since $L_A^l = \alpha_l \cdot F_l(U^l - y_0)$, so $dL_A^l/d\overline{w} = \alpha_l \cdot f_l(U^l - y_0) \cdot \left(dU^l/d\overline{w}\right)$. Then, whenever $dU^l/d\overline{w} > 0$, minimum wage hikes increase labor market participation. The behavioral response is scaled by $f_l(U^l)$, which may be negligible. This may result in positive effects of expected utilities with little participation effects at the aggregate level.

High-skill workers If $\min_m\{w_m^h\} > \overline{w}$, equilibrium effects for high-skill workers take the form of equation (6). Then, the question is whether there are equilibrium forces that rule out solutions of the form $dU^h/d\overline{w} = d\theta_i^h/d\overline{w} = dw_i^h/d\overline{w} = 0$. In this model, effects in high-skill sub-markets are mediated by the production function, since demand for high-skill workers depends on low-skill workers through ϕ. Then, this model may induce within-firm spillovers explained by a technological force. Changes in low-skill markets affect high-skill posting, thus affecting high-skill workers application decisions.

Firms Workers react to changes in the minimum wage by changing their application strategies and extensive margin decisions, thus affecting sub-markets' tightness. This affects the profit maximization problem of the firms. Appendix A provides expressions for the effects of minimum wage changes on firms' outcomes, which are analytically complex given the potential non-linearities of the matching, production, and vacancy cost functions. In what follows, I describe the main intuitions behind the analysis.

Firms for which the minimum wage binds optimize low-skill vacancies and high-skill wages and vacancies taken low-skill wages as given. The effect of the minimum wage on low-skill vacancy posting is ambiguous. On one hand, an increase in the minimum wage induces a mechanical increase in labor costs, decreasing the expected value of posting a low-skill vacancy. However, if sub-market tightness decreases given the increase in applicants, job-filling probabilities increase. This increases the expected value of posting a low-skill vacancy. Within the minimum wage sub-market, the net effect on vacancies is more likely to be negative the lower the productivity.[12] That is, the least productive firms among the constrained group reduce their size after increases in the minimum wage while the most productive firms within this group could have null or positive firm-specific employment effects.

Firms for which the minimum wage does not bind also react by adapting their posted wages and vacancies to changes in their relevant sub-market tightness. The analytical expression for the wage spillover is difficult to sign and interpret but directly depends on the change in sub-market tightness (see equation (A.XXI) of Appendix A). Since wages and vacancies are positively correlated at the firm and skill level, if wage spillovers are positive, then unconstrained firms also post more vacancies and, therefore, increase their size. Therefore, the model has potential to generate reallocation effects.

Profits are also affected by minimum wage changes. Firms for which the minimum wage binds face a reduction in profits regardless of the employment effect. This in turn leads marginal firms to exit the market after increases in the minimum wage. Firms for which the minimum wage does not bind may also have their profits affected given the change in the equilibrium job-filling probabilities.

Relation to empirical literature The purposely imposed tractability needed for the optimal policy analysis puts limits on the ability of the model to fully rationalize observed labor market reactions to minimum wage changes.[13] However, the proposed framework generates predictions consistent with the empirical literature that favor its suitability for the policy analysis.

One systematic finding of the empirical literature is that minimum wage hikes generate positive wage effects with limited –or *elusive*– disemployment effects (see Manning, 2021 for a recent review). This empirical fact is inconsistent with a perfectly competitive model of the labor market, and is difficult to rationalize with a random search framework since it requires an implausibly large labor force participation response that is at odds with the empirical literature (Cengiz et al., 2022). The proposed framework can rationalize positive wage effects with limited employment and participation effects through the equilibrium changes in applications. When the minimum wage increases, constrained firms face a mechanic increase

[12] In the model, it is possible to have productivity dispersion across firms that pay the minimum wage. Concretely, all firms whose market low-skill wage is lower than $\overline{w}$ bunch at $\overline{w}$ conditional on entry.

[13] For structural models with richer levels of heterogeneity and flexibility, see Haanwinckel (2020), Ahlfeldt et al. (2022), Berger et al. (2022), Drechsel-Grau (2022), Engbom and Moser (2022), and Hurst et al. (2022).

in their labor costs. However, job applicants reallocate applications toward these jobs, increasing the expected value of posting vacancies. This effect attenuates the negative shock in labor costs. The reorganization of applications within the mass of active workers can mediate this result when the size of the density at the margin of indifference is low enough to prevent important participation responses.

The empirical literature also finds that minimum wages generate spillovers to non-minimum wage jobs in terms of wages and employment both within and between firms (Cengiz et al., 2019; Derenoncourt et al., 2021; Dustmann et al., 2022; Forsythe, 2022; Giupponi and Machin, 2022), and have negative effects on firm profits (Draca et al., 2011; Harasztosi and Lindner, 2019; Drucker et al., 2021). The model incorporates both sets of predictions. The same responses in applications that dampen the employment effects generate spillovers to firms that pay higher wages through changes in their sub-markets' tightness, and to high-skill workers through technological restrictions embedded in the production function.[14]

3 Optimal Policy Analysis

This section uses the model of the labor market to develop an optimal policy analysis where a social planner chooses the minimum wage, the labor income tax system, and the corporate tax rate to maximize social welfare taking as given social preferences for redistribution.

3.1 Social planner's problem

The notion of *optimal policy* refers to policy parameters that maximize a social welfare function. Following related literature (Kroft et al., 2020; Lavecchia, 2020), the social planner is assumed to be utilitarian and maximize the sum of expected utilities. I assume the social planner does not observe c and ψ and, therefore, constrains the policy choice to second-best incentive-compatible policy schemes.

The social welfare function is given by

$$\begin{aligned} SW(\overline{w},T,t) \quad = \quad & \left(L_I^l + L_I^h + K_I\right) \cdot G(y_0) + \alpha_l \cdot \int_0^{U^l - y_0} G(U^l - c) dF_l(c) \\ & + \alpha_h \cdot \int_0^{U^h - y_0} G(U^h - c) dF_h(c) + K \cdot \int_{\psi^*}^{\overline{\psi}} G\left((1-t)\cdot \Pi(\psi,t) - \xi\right) dO(\psi), \end{aligned} \tag{7}$$

where $(\overline{w},T,t)$ are the policy paramters –the minimum wage, the (possibly non-linear) income tax schedule, and the flat corporate tax rate–, and G is an increasing and concave function that accounts for social preferences for redistribution. G induces curvature to the individual money-metric utilities thus

[14] The model fails to accommodate other relevant effects of the minimum wage documented in the empirical literature, namely the passthrough of minimum wages to output prices (MaCurdy, 2015; Allegretto and Reich, 2018; Harasztosi and Lindner, 2019; Leung, 2021; Ashenfelter and Jurajda, 2022; Renkin et al., 2022) and their effects on worker- and firm-level productivity (Riley and Bondibene, 2017; Mayneris et al., 2018; Coviello et al., 2021; Ruffini, 2021; Ku, 2022). Appendix A argues that these pieces are unlikely to play a central role in the optimal policy analysis.

allowing social gains from redistributing from high- to low- utility individuals. The degree of concavity of G defines the social preferences for redistribution (see the discussion below). The incentive compatibility constraints are included in the limits of integration since the planner internalizes that the policy parameters affect the participation decisions through U^l, U^h, and ψ^*. The first term of equation (7) accounts for the utility of inactive workers and inactive capitalists who get income equal to y_0. The second and third terms account for the expected utility of low- and high-skill workers that enter the labor market, also referred to as active workers. Finally, the last term accounts for the utility of active capitalists.[15]

Assuming no exogenous spending requirement, the planner's budget constraint is given by

$$\begin{aligned}\left(L_I^l + L_I^h + K_I + \rho^l \cdot L_A^l + \rho^h \cdot L_A^h\right) \cdot y_0 \quad \leq \quad & \int \left(E_m^l \cdot T(w_m^l) + E_m^h \cdot T(w_m^h)\right) dm \\ & + t \cdot K \cdot \int_{\psi^*}^{\overline{\psi}} \Pi(\psi, t) dO(\psi), \end{aligned} \tag{8}$$

where $E_m^s = p_m^s \cdot L_m^s$ is the mass of employed workers of skill s in sub-market m and ρ^s is the skill-specific unemployment rate given by $\left(L_A^s - \int E_m^s dm\right)/L_A^s$. The budget constraint establishes that the transfer paid to individuals with no market income must be funded by the tax collection on employed workers and active capitalists.

Understanding G To better understand the role of G, define the average social marginal welfare weights (SMWWs) of inactive workers, active workers of skill type s, and active capitalists of type ψ as

$$g_0 = \frac{G'(y_0)}{\gamma}, \qquad g_1^s = \frac{\alpha_s \cdot \int_0^{U^s - y_0} G'(U^s - c) dF_s(c)}{\gamma \cdot L_A^s}, \qquad g_\psi = \frac{G'((1-t)\cdot \Pi(\psi, t) - \xi)}{\gamma}, \tag{9}$$

where $\gamma > 0$ is the social planner's budget constraint multiplier. Average SMWWs represent the social value of the marginal utility of consumption normalized by the social cost of raising funds, thus measuring the social marginal value of redistributing one dollar uniformly across a group of individuals. When SMWWs are above one, the planner benefits from redistributing toward that group since the gains in the social value of utility outweight the distortions induced by the increase in revenues. That is, a given value of g_X indicates that the government is indifferent between g_X more dollars of public funds and 1 dollar of additional consumption of individuals of group X (Saez, 2001).

The utilitarian assumption used in equation (7) implies that the SMWWs are endogenous to final allocations (and, therefore, to the policy parameters) since social welfare only depends on the concave transformation of individual money-metric utilities. Alternative formulations of the problem can generate

[15] The average expected utility of active workers of skill s is $\int_0^{U^s-y_0} G(U^s - c)d\tilde{F}_s(c)$, where $\tilde{F}_s(c) = F_s(c)/F_s(U^s - y_0)$. Then, total expected utility is given by $L_A^s \cdot \int_0^{U^s-y_0} G(U^s - c)d\tilde{F}_s(c)$ which yields the expressions above noting that $L_A^s = \alpha_s \cdot F(U^s - y_0)$. The average utility of capitalists is $\int_{\psi^*}^{\overline{\psi}} G((1-t)\cdot\Pi(\psi,t)-\xi)d\tilde{O}(\psi)$, with $\tilde{O}(\psi) = O(\psi)/(1-O(\psi^*))$. Their total utility is therefore $K_A \cdot \int_{\psi^*}^{\overline{\psi}} G((1-t)\cdot\Pi(\psi,t)-\xi)d\tilde{O}(\psi)$, which yields the expression above noting that $K_A = K\cdot(1-O(\psi^*))$.

different microfoundations for the SMWWs, for example, through exogenous Pareto weights or generalized SMWWs (Saez and Stantcheva, 2016). More generally, SMWWs are sufficient statistics for preferences for redistribution since their values inform the willingness to transfer incomes between different groups of individuals. I return to this when discussing the results of the optimal policy analysis.

Rationing assumptions Since the social planner cares about expected utilities, rationing assumptions conditional on entering the labor market do not affect the welfare analysis: all workers have equal ex-ante expected utilities, so the allocation to jobs and unemployment after policy changes does not condition the planner's problem. By contrast, rationing assumptions are central in optimal policy analyses based on competitive labor markets.[16] Rationing matters if sorting to firms conditional on participation depends on c, and would play a role if adding additional layers of worker-level heterogeneity imply that some groups are more likely to work at low-wage firms or to be unemployed. This would affect the analysis since the presence of winners and losers within skill-group may distort the assessment of the distributional effects of minimum wage increases (Hurst et al., 2022). I return to this question in Section 4 when testing for heterogeneities in the empirical estimation of the worker-level sufficient statistics.

3.2 Case with no taxes

I now proceed to analyze the redistributive properties of the minimum wage using the framework described above. I start abstracting from the tax system to isolate the effects on the relative tradeoff between low-skill workers, high-skill workers, and capitalists. Taxes and transfers are introduced in the next subsection.

PROPOSITION I: *In the absence of taxes, increasing the minimum wage is welfare improving if*

$$\frac{dU^l}{d\overline{w}} \cdot L_A^l \cdot g_1^l + \frac{dU^h}{d\overline{w}} \cdot L_A^h \cdot g_1^h + K \cdot \int_{\psi^*}^{\overline{\psi}} g_\psi \frac{d\Pi(\psi)}{d\overline{w}} dO(\psi) \quad > \quad 0. \tag{10}$$

Proof: See Appendix B.

Proposition I shows that a small increase of the minimum wage can affect the welfare of active low-skill workers (first term), active high-skill workers (second term), and active capitalists (third term). Depending on the change in utility for the different groups ($dU^s/d\overline{w}$ and $d\Pi(\psi)/d\overline{w}$), the social value of those changes (g_1^s and g_ψ), and the size of the groups (L_A^s and $K \cdot o(\psi)$), increasing the minimum wage may be desirable or not for the social planner.[17] This implies that the minimum wage can affect both the relative welfare within labor income earners and between labor and capital income earners.

[16] See Lee and Saez (2012). Rationing assumptions may be controversial as suggested by Mankiw (2013).

[17] While changes in U^s and ψ^* also affect extensive margin decisions, those margins do not induce first-order welfare effects because marginal workers and capitalists are initially indifferent between states.

Welfare weights To understand why the analysis emphasizes the distributional effects of the minimum wage, consider a situation where $g_1^l = g_1^h = g_\psi = 1$, for all ψ. Then, the planner's problem is reduced to assessing changes in total output. The analysis changes when SMWWs are unrestricted. Total output could decrease after minimum wage increases, but if the gains for winners are more socially valuable than the losses for losers, then increasing the minimum wage can be welfare-improving. For example, if the social planner does not care about the utility of capitalists and high-skill workers, there could be scope to increase the minimum wage if the utility of low-skill workers increases after the policy change. The utilitarian assumption implies that SMWWs are endogenous to final allocations, so they are inversely proportional to after-tax incomes. The steepness of the relationship depends on the concavity of G.[18]

Sufficient statistics Given values for the SMWWs, if the sizes of the groups are observed, the missing piece for taking equation (10) to the data is to calibrate values for $dU^s/d\overline{w}$ and $d\Pi(\psi)/d\overline{w}$. Reduced-form estimates of these elasticities facilitate the quantitative assessment of Proposition I without needing to impose structural restrictions on the primitives of the model of the labor market. That is, empirical counterparts of $dU^s/d\overline{w}$ and $d\Pi(\psi)/d\overline{w}$ work as sufficient statistics (Chetty, 2009; Kleven, 2021) for assessing the welfare implications of minimum wage changes.

Profits are, in principle observable, so it is feasible to have reduced-form estimates of $d\Pi(\psi)/d\overline{w}$. Regarding U^s, recall that, in the absence of taxes, $U^s = p_m \cdot w_m$. Multiplying both sides by the sub-market mass of applicants, L_m^s, and integrating over m, yields

$$U^s = \frac{\int E_m^s w_m^s dm}{L_A^s} = (1-\rho^s) \cdot \mathbb{E}_m[w_m^s] + \rho^s \cdot 0, \tag{11}$$

where ρ^s is the skill-specific unemployment rate and $\mathbb{E}_m[w_m^s] = \int \nu_m^s w_m^s dm$, with $\nu_m^s = E_m^s / \int E_m^s dm$, is the average wage of employed workers. This implies that U^s is equal to the average wage of active workers including the unemployed. In the case with taxes, U^s is equal to the average pre-tax wage of active workers including the unemployed net of the their average tax liabilities.[19] In both cases, U^s can be computed using data on wages, tax liabilities, employment and participation rates. Then, $dU^s/d\overline{w}$

[18] Alternatively, equation (10) could be derived using a local perturbation approach to avoid the utilitarian microfoundation and impose exogenous social valuation criteria (Saez and Stantcheva, 2016).

[19] Recall that, in the case with taxes, $U^s = p_m^s \cdot y_m^s + (1-p_m^s) \cdot y_0$. Multiplying both sides by the sub-market mass of applicants, L_m^s, and integrating over m, gives

$$U^s = \frac{\int E_m^s (w_m^s - T(w_m^s) - y_0) dm}{L_A^s} + y_0 = \frac{\int E_m^s w_m^s dm}{L_A^s} - \frac{\int E_m^s (T(w_m^s) + y_0) dm}{L_A^s} + y_0, \tag{12}$$

where $E_m^s = p_m^s \cdot L_m^s$. If the tax schedule is constant, then

$$\frac{dU^s}{d\overline{w}} = \frac{d}{d\overline{w}} \left(\frac{\int E_m^s w_m^s dm}{L_A^s} \right) - \frac{d}{d\overline{w}} \left(\frac{\int E_m^s (T(w_m^s) + y_0) dm}{L_A^s} \right). \tag{13}$$

The first term represents the change in the average pre-tax wage among active workers (see equation (11)). The second term represents the change in average tax liabilities net of transfers among active workers.

can be estimated to quantitatively assess equation (10). Section 4 makes progress in this regard.

Two things are worth discussing about the sufficient statistic for workers, $dU^s/d\overline{w}$. First, it captures all general equilibrium effects that affect workers' utility, including effects on wages, employment, and participation. There is an unsettled discussion in the public debate about the appropriate way of weighting these different effects. The proposed framework offers an avenue for aggregating them in a single elasticity.[20] Second, equation (11) relies on the risk-neutrality assumption made in Section 2. If workers are risk-averse, then Proposition I remains valid but U^s no longer equals the average wage among active workers including the unemployed, so it cannot be estimated without further assumptions. One way to assess the concerns of using the risk-neutral sufficient statistic is to decompose the empirical estimate across the different margins. If changes in employment are negligible relative to changes in wages, then the risk-neutrality assumption should not have first-order effects on the interpretation of the estimated elasticities. I come back to this discussion in Section 4.

3.3 Case with taxes

The case without taxes informs about the direct welfare effects of the minimum wage. However, in the presence of taxes, changes in labor market outcomes and profits affect tax collection and transfer spending. These fiscal externalities matter for assessing whether increasing the minimum wage is desirable.

Fixed taxes I first consider a case where the social planner takes the tax system as given and chooses $\overline{w}$ to maximize social welfare. This characterizes the mechanic interactions between the minimum wage and the tax system. When unmodeled constraints restrict the scope for tax reforms, this may be the policy-relevant scenario for assessing the welfare effects of minimum wage reforms.

Proposition II: *If taxes are fixed, increasing the minimum wage is welfare improving if*

$$\begin{aligned}\frac{dU^l}{d\overline{w}} \cdot L_A^l \cdot g_1^l + \frac{dU^h}{d\overline{w}} \cdot L_A^h \cdot g_1^h + K \cdot (1-t) \cdot \int_{\psi^*}^{\overline{\psi}} g_\psi \frac{d\Pi(\psi, t)}{d\overline{w}} dO(\psi) \\ + \int \left(\frac{dE_m^l}{d\overline{w}} \left(T(w_m^l) + y_0 \right) + E_m^l T'(w_m^l) \frac{dw_m^l}{d\overline{w}} \right) dm \\ + \int \left(\frac{dE_m^h}{d\overline{w}} \left(T(w_m^h) + y_0 \right) + E_m^h T'(w_m^h) \frac{dw_m^h}{d\overline{w}} \right) dm \\ + t \cdot K \cdot \int_{\psi^*}^{\overline{\psi}} \frac{d\Pi(\psi, t)}{d\overline{w}} dO(\psi) - \frac{dK_I}{d\overline{w}} \cdot (t \cdot \Pi(\psi^*, t) + y_0) \quad > \quad 0.\end{aligned} \tag{14}$$

Proof: See Appendix B.

[20] While the sign of $dU^s/d\overline{w}$ is in principle ambiguous, it is not determined by the sign of the employment effects. Appendix A shows the disemployment effects that can be tolerated for the minimum wage to increase average workers' welfare given positive wage effects. If employment and wage effects are positive, welfare effects on workers are unambiguously positive.

The first line of Proposition II reproduces the same welfare tradeoff described in Proposition I. The second to fourth lines summarize the fiscal externalities on both sides of the market. These fiscal externalities matter for the analysis since they either relax or restrict the planner's budget constraint, consequently relaxing or restricting the redistribution already done by the existing tax system.

The second line describes the fiscal externalities on low-skill labor markets. The first term shows that, if low-skill employment increases, there is an increase in tax collection (or expenditure if there are transfers to workers), $T(w_m^l)$, and a decrease in transfers paid to unemployed individuals, y_0. The opposite happens when employment decreases. The second term shows that if the wages of employed workers change, income tax collection changes according to the shape of the income tax schedule, $T'(w_m^l)$. The third line represents the same effects but for high-skill labor markets.

The fourth line describes the fiscal externalities on the capitalists' side. The first term shows that changes in profits affect the corporate tax revenue. If profits decrease, the social planner collects less revenue. The second term shows that firms that exit the market generate a negative fiscal externality since they switch from paying taxes to receiving a transfer. Both effects are increasing in the corporate tax rate: the larger t, the larger the revenue loss produced by smaller profits and extensive margin responses.

Firm-level fiscal externalities seem particularly relevant in the current state of international tax competition (Devereux et al., 2008, 2021). Under international capital mobility, it may be difficult to enforce large corporate tax rates because capital can fly to low-tax countries. If corporate taxes are low, then the rationale for using the minimum wage becomes stronger. One concern with this argument is that the same reasons that limit corporate tax rates could apply to the minimum wage: international capital could also react to minimum wage changes. In Section 4 I document that the profit effects are concentrated in labor intensive industries whose capital is presumably less mobile relative to other industries. By contrast, the effects of corporate tax changes on pre-tax profits seem to be concentrated in capital intensive industries (Kennedy et al., 2022). This suggests that the economic reasons that push corporate tax rates down are not extendable to minimum wages. I formalize this intuition in Section 5.

Optimal taxes The previous analysis illustrates the mechanical interaction between the minimum wage and the tax schedule but does not answer if both policies are desirable at the joint optimum. The following proposition explores the desirability of the minimum wage when the social planner jointly optimizes the tax system and the minimum wage. For analytical simplicity, I assume that either $\max_i w_i^l < \min_j w_j^h$, or that the social planner can implement skill-specific income tax schedules.[21]

[21] This allows me to solve the planner's problem doing pointwise maximization. These assumptions increase the attractiveness of the tax system, making more restrictive the case for a binding minimum wage.

Proposition III: *If taxes are optimal, increasing the minimum wage above the market level is always welfare improving. Additional increases are welfare improving if*

$$\begin{aligned}\frac{\partial U^l}{\partial \overline{w}} \cdot L_A^l \cdot g_1^l + \frac{\partial U^h}{\partial \overline{w}} \cdot L_A^h \cdot g_1^h + K \cdot (1-t) \cdot \int_{\psi^*}^{\overline{\psi}} g_\psi \frac{\partial \Pi(\psi,t)}{\partial \overline{w}} dO(\psi) \\ + \int \left(\frac{\partial E_m^l}{\partial \overline{w}} \left(T(w_m^l) + y_0 \right) + E_m^l \frac{\partial w_m^l}{\partial \overline{w}} \right) dm \\ + \int \left(\frac{\partial E_m^h}{\partial \overline{w}} \left(T(w_m^h) + y_0 \right) + E_m^h \frac{\partial w_m^h}{\partial \overline{w}} \right) dm \\ + t \cdot K \cdot \int_{\psi^*}^{\overline{\psi}} \frac{\partial \Pi(\psi,t)}{\partial \overline{w}} dO(\psi) - \frac{\partial K_I}{\partial \overline{w}} \cdot (t \cdot \Pi(\psi^*, t) + y_0) \quad > 0 \quad . \end{aligned} \tag{15}$$

Furthermore, at the joint optimum: (i) the SMMWs of inactive individuals, active low-skill workers, and active high-skill workers average to 1, and (ii) the average SMMW among active capitalists is below 1. Proof: See Appendix B.

At a high-level, Proposition III reproduces the same intuition than Proposition II: the desirability of the minimum wage depends on both the effects on the relative welfare of active low-skill workers, active high-skill workers, and active capitalists, and on the fiscal externalities generated on labor markets and profits. However, when taxes are optimized together with the minimum wage, the way in which the minimum wage affects welfare and generate fiscal effects changes. This is reflected in two important differences between equations (14) and (15) that illustrate the forces that play a role in the joint optimum.

First, all relevant elasticities are *micro* rather than *macro* elasticities (Landais et al., 2018b,a; Kroft et al., 2020; Lavecchia, 2020), which I denote by partial rather than total derivatives. Macro elasticities (Propositions I and II) internalize all general equilibrium effects of the minimum wage, while micro elasticities (Proposition III) mute some of these effects by keeping after-tax allocations constant. Formally, recall that $U^s = p_m^s \cdot y_m^s + (1-p_m^s) \cdot y_0 \equiv p_m^s \cdot \Delta y_m^s + y_0$, with $\Delta y_m^s = y_m^s - y_0$. When taxes are fixed, both Δy_m^s and p_m^s can react to minimum wage changes. However, taxes being optimal imply that Δy_m^s is fixed so the minimum wage directly affects workers' welfare only through changes in the employment probabilities. Then, the labor market equilibrium effects of the minimum wage are only driven by changes in vacancy posting as a response to the increase in labor costs so the direct welfare effects on workers are proportional to the (presumably negative) employment effects.[22]

Second, the fiscal externalities are also affected by optimal taxes. Since after-tax allocations are fixed by the optimal tax schedule, changes in wages due to the minimum wage affect within-firm redistribution.

[22] If $U^s = p_m^s \cdot \Delta y_m^s + y_0$, multiplying by L_m^s and integrating over m yields $(U^s - y_0) \cdot L_A^s = \int E_m^s \cdot \Delta y_m^s dm$. Then

$$\frac{\partial U^s}{\partial \overline{w}} \cdot (L_A^s + (U^s - y_0) \cdot f^s(U^s - y_0)) \quad = \quad \int \frac{\partial E_m^s}{\partial \overline{w}} \Delta y_m^s dm. \tag{16}$$

This is captured by the term $E_m^s \cdot (\partial w_m^s / \partial \overline{w})$: there are fiscal gains from wage increases because they switch the burden of redistribution from the government to firms and, therefore, relax the social planner's budget constraint. This cannot be done completely by the corporate tax since it distorts pre-tax profits.[23] For the minimum wage sub-market we have that $\partial w_m^l / \partial \overline{w} = 1$, so the fiscal gain is equivalent to the mass of low-skill workers earning the minimum wage. To develop intuition, consider a marginal increase to the optimal after-tax allocation to low-skill workers. This increases labor supply, so firms optimally react by lowering pre-tax wages (Rothstein, 2010). The minimum wage mutes this behavioral response, therefore, increasing the efficiency of the transfer to low-skill workers by making it work as a lump-sum transfer.

The desirability of the minimum wage at the joint optimum depends on how these two forces balance; the assessment of equation (15) is ultimately a quantitative question. Distortions in vacancy posting are negligible when the minimum wage is exactly set at the market level, so it is always optimal to have a binding minimum wage just above the market equilibrium because of the fiscal gain.[24] As the minimum wage departs from the market level, the employment costs increase and become more likely to outweight the fiscal benefits, hinting the existence of an interior solution on the optimal minimum wage policy. I revisit these interactions at the joint optimum in Section 5 to get more concrete policy recommendations about the desirability of the minimum wage under different tax schedules.[25]

Caveats I briefly discuss three elements that may affect the optimal policy analysis whose formal treatment is beyond the scope of this paper.

First, the theoretical attractiveness of the income tax system to implement the desired allocations relies on its flexibility and its perfect enforcement. In the real world, income tax schedules are usually not fully non-linear (e.g., they follow a bracket structure), are not perfectly enforced, and are costly to administrate. Taxes are avoided or evaded (Andreoni et al., 1998; Slemrod and Yitzhaki, 2002; Kleven et al., 2011; Guyton et al., 2021) and the take up of benefits at the bottom of the distribution is not perfect (Currie, 2006; Kopczuk and Pop-Eleches, 2007; Chetty et al., 2013; Bhargava and Manoli, 2015; Guyton et al., 2017; Goldin, 2018; Cranor et al., 2019; Finkelstein and Notowidigdo, 2019; Linos et al., 2021). These frictions generate additional efficiency costs to the tax system. Also, abstracting from tax evasion rules out additional complementarities between the minimum wage and the tax system. For example, if workers under report their incomes, then the minimum wage can increase tax collection by

[23] Firm-level heterogeneity, revenue distortions, and entry distortions impede t to fully redistribute from capitalists to workers. That is why the average SMWW of active capitalists is less than 1 at the joint optimum. Part of this result relies on the corporate tax rate being flat. Extending to non-linear frameworks may induce additional welfare gains from using the tax system relative the minimum wage. For example, Saez and Zucman (2021) propose non-linear payroll taxes to mimic minimum wages. Non-linear corporate taxes, however, are not common practice and may induce additional behavioral responses that affect the efficiency of the tax system (e.g., Bachas and Soto, 2021).

[24] This result is a direct consequence of the envelope theorem. If $\overline{w} = \min_m w_m^l$, then the minimum wage respect the first-order condition of constrained firms, impeding a first-order distortion in posted vacancies.

[25] The proposition also states that, in the joint optimum, the SMMWs of inactive individuals, active low-skill workers, and active high-skill workers average to 1, which is a standard result of optimal tax analyses with quasi-linear utility functions.

setting a floor on reported labor income (Bíró et al., 2022; Feinmann et al., 2022).

Second, an additional benefit from using the tax system that is not in the analysis is the ability of the tax and transfers systems to tag on additional variables such as family size. That type of flexibility is unlikely to apply to the minimum wage policy (Stigler, 1946).

Third, the minimum wage affects the distribution before taxes and transfers while taxes and transfers alter pre-tax values to generate the after-tax distribution.[26] The optimal policy analysis assumes that the social value of after-tax allocations does not depend on the composition between pre-tax incomes and taxes and transfers. However, recent evidence suggests that affecting the pre- and the post-tax and transfer distribution has different implications for long-run trends in inequality (Bozio et al., 2020; Blanchet et al., 2022). Moreover, social preferences may put different weights on the two types of interventions. For example, McCall (2013) provides survey evidence that suggests that the US public care about inequality and redistribution, but prefer policies that address inequality within the firm rather than with taxes and transfers. This is consistent with the results of state-level ballot initiatives that have favored minimum wage changes relative to reforms to top marginal income tax rates (Saez, 2021). Such social preferences could be incorporated by generalizing the SMWWs (Saez and Stantcheva, 2016).

4 Sufficient Statistics Estimation

The previous section derives conditions for the desirability of the minimum wage. These conditions depend on the welfare impacts of minimum wages on low-skill workers, high-skill workers, and capitalists, which are summarized by sufficient statistics. This section estimates these sufficient statistics using publicly available US data. The focus is on the macro-version of the sufficient statistics, that is, the version that considers all general equilibrium effects of minimum wages, which relates to Propositions I and II.[27]

4.1 Empirical strategy

The empirical strategy exploits state-level variation in minimum wages to estimate stacked event studies.

Events I follow Cengiz et al. (2019, 2022) strategy to define state-level events. A state-by-year minimum wage is defined as the maximum between the statutory values of the federal and state minimum wages throughout the calendar year. I use data from Vaghul and Zipperer (2016) for the 1997-2019 period for which I can observe all the outcomes of interest within eight-year balanced windows. Nominal values are transformed to 2016 dollars using the R-CPI-U-RS index including all items. An event is defined as

[26]This claim is true only to a first-approximation since changes in taxes can also affect the pre-tax income distribution (Roine et al., 2009; Alvaredo et al., 2013; Piketty et al., 2014; Vergara, 2022).

[27]Estimating micro elasticities –Proposition III– requires a stricter empirical design that exploits variation in minimum wages while keeping after-tax allocations fixed.

a state-level hourly minimum wage increase above the federal minimum wage of at least \$0.25 (in 2016 dollars) in a state with at least 2% of the employed population affected, where the affected population is computed using the NBER Merged Outgoing Rotation Group of the CPS (henceforth, CPS-MORG).[28] These restrictions are imposed to focus on minimum wage increases that are likely to have effects on the labor market. Small state-level or binding federal minimum wage increases are not recorded as events, however, regressions control for small state-level and federal minimum wage increases. I also restrict the attention to events where treated states do not experience other events in the three years previous to the event and whose timing allow to observe the outcomes from three years before to four years after. This results in 50 valid state-level events, whose time distribution is plotted in Figure C.I of Appendix C. Table C.I of Appendix C display the list of the considered events with their corresponding treated states.

Estimating equation Estimating event studies in this setting is challenging because of two reasons. First, events do not induce an absorbing status: states may increase their minimum wages several times throughout the period considered. Second, treatment effect heterogeneity may induce bias when treatment adoption is staggered (de Chaisemartin and D'Haultfœuille, 2022; Roth et al., 2022).

To deal with these issues, I implement stacked event studies (Cengiz et al., 2019, 2022; Gardner, 2021; Baker et al., 2022) as follows. For each event, I define a time window that goes from 3 years before the event to 4 years after. All states that do not experience events in the event-specific time-window define an event-specific control group. This, in turn, defines an event-specific dataset. Finally, all event-specific datasets are appended and used to estimate a standard event study with event-specific fixed effects. This leads to the following estimating equation:

$$\log Y_{ite} = \sum_{\tau=-3}^{4} \beta_\tau D_{ite}^\tau + \alpha_{ie} + \gamma_{te} + \rho_{ite} + \epsilon_{ite}, \tag{17}$$

where i, t, and e index state, year, and event, respectively, Y_{ite} is an outcome of interest (see next subsection), D_{ite}^τ are event indicators with τ the distance from the event (in years), α_{ie} are state-by-event fixed effects, γ_{te} are year-by-event fixed effects, and ρ_{iqe} are state-by-year-by-event varying controls that include small state-level minimum wage increases and binding federal minimum wage increases.[29] I also consider specifications where the year-by-event fixed effects are allowed to vary across census regions and census divisions. β_{-1} is normalized to 0. To allow for correlation within states across events, standard errors are clustered at the state level. Regressions are weighted by state-by-year average total population.

[28] This is done by computing employment counts by wage bins and checking whether, on average, the previous year share of workers with wages below the new minimum wage is above 2% (Cengiz et al., 2019).

[29] Following Cengiz et al. (2019, 2022), controls for small state-level and binding federal minimum wage increases are included as follows. Let $\hat{t}$ be the year in which the small state-level or binding federal minimum wage increase takes place. Then, define $Early_t = 1\{t \in \{\hat{t}-3, \hat{t}-2\}\}$, $Pre_t = 1\{t = \hat{t}-1\}$ and $Post_t = 1\{t \in \{\hat{t}, \hat{t}+1, \hat{t}+2, \hat{t}+3, \hat{t}+4\}\}$, and let $Small_i$ and Fed_i be indicators of states that face small state-level and binding federal minimum wage increases, respectively. Then ρ_{ite} includes all the interactions between $\{Early_t, Pre_t, Post_t\} \times \{Small_i, Fed_i\}$ for each event separately.

Regressions for capitalists' outcomes that vary at the industry-level allow for state-by-industry-by-event fixed-effects, cluster standard errors at the state-by-industry level, and weight observations using the average state-by-industry employment in the pre-period reported in the QCEW files.

I also consider standard differences-in-differences regressions:

$$\log Y_{ite} = \beta T_{ie} \text{Post}_{te} + \alpha_{ie} + \gamma_{te} + \rho_{ite} + \epsilon_{ite}, \tag{18}$$

where T_{ie} is an indicator variable that takes value 1 if state i is treated in event e, Post_{te} is an indicator variable that takes value 1 if year t is larger or equal than the treatment year in event e, and all other variables are defined as in equation (17). The coefficient of interest is β, which captures the average treatment effect in the post-event years (from $\tau = 0$ to $\tau = 4$).

4.2 Data

Outcomes consist on state-level aggregates for 1997-2019 computed using publicly available data.

Workers The sufficient statistic for changes in active workers' welfare is $dU^s/d\overline{w}$, for $s \in \{l, h\}$. Equations (11) and (13) show that $dU^s/d\overline{w}$ equals the change in the average post-tax wage of active workers including the unemployed, which can be decomposed between changes in their average pre-tax wage (including the unemployed) and changes in their average net tax liabilities.

I use the CPS-MORG data to compute average pre-tax hourly wages and the Basic CPS monthly files to compute employment and participation rates at the state-by-year-by-skill level. The pre-tax component of the sufficient statistic, then, can be computed as the average wage times the employment rate. Low-skill (high-skill) workers are defined as not having (having) a college degree. Hourly wages are either directly reported or indirectly computed by dividing reported weekly earnings by weekly hours worked. I drop individuals aged 15 or less, self-employed individuals, and veterans. Nominal wages are transformed to 2016 dollars using the R-CPI-U-RS index including all items. Observations whose hourly wage is computed using imputed data (on wages, earnings, and/or hours) are excluded to minimize the scope for measurement error. To avoid distorting low-skill workers' statistics with non-affected individuals at the top of the wage distribution, I restrict the low-skill workers sample to workers that are either out of the labor force, unemployed, or in the bottom half of the wage distribution when employed. I test how results change when considering different wage percentile thresholds. To compute changes in net tax liabilities at the state-by-year level, I use data from the BEA regional accounts. I consider income maintenance benefits, medical benefits, and gross federal income tax liabilities.[30]

[30] The BEA definition of income maintenance benefits is as follows: "Income maintenance benefits consists largely of Supplemental Security Income (SSI) benefits, Earned Income Tax Credit (EITC), Additional Child Tax Credit, Supplemental Nutrition Assistance Program (SNAP) benefits, family assistance, and other income maintenance benefits, including general

Capitalists The sufficient statistic that summarizes changes in active capitalists' welfare driven by minimum wage changes is the change in firm profits, $d\Pi(\psi, t)/d\overline{w}$, for $\psi \in [\psi^*, \overline{\psi}]$.

Absent firm-level microdata, I compute a measure of average profits per firm at the industry-by-state-by-year level. I use the Gross Operating Surplus (GOS) estimates from the BEA regional accounts as a proxy of state-level aggregate profits and divide them by the average number of private establishments reported in the QCEW data files.[31] Nominal profits are transformed to 2016 dollars using the R-CPI-U-RS index including all items. I consider 25 industries that have a relatively large coverage across states and years. Noting that minimum wage workers are not evenly distributed across industries (e.g., BLS, 2020), I group industries into three large groups: manufacturing, exposed services, and non-exposed services.[32] Fiscal effects are proportional to the effect on profits. I also use data on taxes on production and imports net of subsidies reported on the BEA regional accounts at the industry-level, and data on business and dividend income reported in the state-level SOI tables to test for additional fiscal externalities.

Descriptive statistics Table 1 shows descriptive statistics for the non-stacked panel. The total number of observations is 1,173 (51 states times 23 years). All monetary values are annual and in 2016 dollars.[33] Average pre-tax incomes (including the unemployed) are more than 3 times larger for high-skill workers relative to low-skill workers. This is explained by higher hourly wages and weekly hours conditional on employment, and also by higher employment rates. Average income maintenance benefits per working-age individual are 1,051 dollars, which represents around 5% of low-skill workers pre-tax income. Average medical benefits and gross federal income taxes per working-age individual are 4,541 and 7,179 dollars, respectively. Average pre-tax profits per establishment are substantially larger than disposable incomes for workers. In exposed services, the average pre-tax profit per establishment is almost 9 times the average pre-tax income of low-skill workers including the unemployed. The ratio increases to almost 50 times and to 100 times when looking at non-exposed services and manufacturing.

assistance." Medical benefits consider both Medicaid and Medicare programs.

[31] The BEA definition of gross operating surplus is as follows: "Value derived as a residual for most industries after subtracting total intermediate inputs, compensation of employees, and taxes on production and imports less subsidies from total industry output. Gross operating surplus includes consumption of fixed capital (CFC), proprietors' income, corporate profits, and business current transfer payments (net)."

[32] I exclude agriculture and mining. I also exclude construction and finance since they experience particularly abnormal profit dynamics around the 2009 financial crisis. Manufacturing industries include SIC codes 41, 43, 44, 46, 50, 54, 56, and 57, that is, nonmetallic mineral products, fabricated metal products, machinery, electrical equipment, food and beverages and tobacco, printing and related support activities, chemical manufacturing, plastics and rubber products. Exposed services include SIC codes 9, 19, 21, 27, 28, and 34, that is, retail trade, ambulatory health services, nursing and residential care facilities, food, accommodation, and social services and other services. Non-exposed services include SIC codes 8, 10, 11, 13, 14, 15, 16, 17, 20, 24 and 25, that is, wholesale trade, transport, information, real estate, professional services, management of businesses, administrative support, educational services, hospitals, arts, and recreation industries.

[33] While the theoretical and empirical analysis on workers' outcomes is based on average hourly wages, I annualize these values by multiplying them by 52 weeks and the average number of hours worked by skill-group. I below show that weekly hours worked conditional on employment are not affected by minimum wage changes.

4.3 Results

Worker-level pre-tax outcomes Figure 1 plots the estimated coefficients $\{\beta_\tau\}_{\tau=-3}^{4}$ of equation (17) with their corresponding 95% confidence intervals using the average pre-tax hourly wage of active low- and high-skill workers including the unemployed as dependent variables to proxy for U^l and U^h. Each figure plots regressions with two different types of time fixed-effects: year-by-event fixed effects and census-region-by-year-by-event fixed effects. Table 2 presents the estimated coefficients β of equation (18) that summarize the average treatment effect in the post-event period, and also includes specifications that control for census-division-by-year-by-event fixed effects. Panel (a) of Figure 1 shows that state-level minimum wage increases have increased active low-skill workers' welfare. Table 2 shows that the implied elasticity, $d\log U^l/d\log\overline{w}$, ranges between 0.10 and 0.13. Panel (b) of Figure 1 shows that these minimum wage increases have had null effects on high-skill workers' welfare. Table 2 shows that the estimate of $d\log U^h/d\log\overline{w}$ is a precise zero, regardless of the fixed-effects considered. These results suggest that state-level minimum wages have reduced welfare gaps between low- and high-skill active workers.[34]

To better understand how minimum wages have affected low-skill workers, Figure 2 presents separate results for each of the margins that can play a role in the evolution of active low-skill workers' welfare: hourly wages, weekly hours (both conditional on employment), employment rates, and participation rates. Results indicate that all the effect of minimum wage increases on U^l is driven by an increase in the wage conditional on employment, with no effect on hours, employment, or participation.[35,36]

To test for patterns of heterogeneity, Figure 3 plots the estimated β coefficient of equation (18) with its corresponding 95% confidence interval using different groups of low-skill workers. Panel (a) uses the average pre-tax wage of active low-skill workers including the unemployed (U^l) as dependent variable. Panel (b) uses the average pre-tax hourly wage of low-skill workers conditional on employment as dependent variable. Panel (c) uses the average employment rate of low-skill workers as dependent variable. The effects are very stable across groups: all experience an increase in welfare driven by changes in wages conditional on employment with no effects on employment. If anything, teen (aged 16-19) and black low-skill workers seem to experience larger welfare gains. This suggests there are no clear groups of winners and losers within the broad population of low-skill workers.[37]

[34]Figure C.II of Appendix C tests the sensitivity of the result on low-skill workers to the choice of the wage percentile used to trim the sample of employed low-skill workers. Results are robust to using more restrictive samples and to the incorporation of low-skill workers until percentile 80. Including low-skill workers belonging to the top 20% attenuates the estimated result, which is expected given the unlikely response of top wages to changes in the minimum wage.

[35]The lack of employment responses suggests that the result is robust to including curvature in the flow utility of workers.

[36]My results differ from Gandhi and Ruffini (2022) and Jardim et al. (2022) who find effects on hours worked.

[37]Results are consistent with Cengiz et al. (2019, 2022) who find positive wage effects, limited employment effects, and limited participation effects on low-wage workers, i.e., the part of the distribution close to the minimum wage, using similar data and empirical strategy. My results differ from theirs along two dimensions. First, I focus on broad skill groups that are not exclusively composed by minimum wage workers. Second, the main focus of my analysis is the estimation of the composite sufficient statistic rather than the effect on the different margins.

Worker-level fiscal effects The previous analysis focuses on pre-tax workers' outcomes. Figure 4 plots the estimated coefficients $\{\beta_\tau\}_{\tau=-3}^{4}$ of equation (17) with their corresponding 95% confidence intervals using fiscal variables as dependent variables to estimate worker-level fiscal externalities. Table 3 presents the estimated coefficients β of equation (18) that summarize the average treatment effect in the post-event period, and also includes specifications that control for census-division-by-year-by-event fixed effects. Panel (a) of Figure 4 uses total income maintenance transfers per working-age individual as a dependent variable. Consistent with Reich and West (2015) and Dube (2019), results suggest that income maintenance benefits have decreased after state-level minimum wage increases, with the implied elasticity ranging between -0.31 and -0.39. Neither medical benefits (Panel (b)) or gross federal income taxes (Panel (c)) show a response to changes in minimum wages, suggesting that the worker-level fiscal effects are mediated by targeted transfers based on pre-tax income levels.

Capitalist-level pre-tax outcomes Figure 5 plots the estimated coefficients $\{\beta_\tau\}_{\tau=-3}^{4}$ of equation (17) with their corresponding 95% confidence intervals using capitalist-level pre-tax outcomes as dependent variable. Table 4 presents the estimated coefficients β of equation (18) that summarize the average treatment effect in the post-event period. Panels (a) and (b) plot regressions that pool all industries and that control by three different types of time fixed-effects: year-by-event fixed effects, census-region-by-year-by-event fixed effects, and census-division-by-year-by-event fixed effects. Panels (c) and (d) plot regressions splitting by industry group that control by census-division-by-year-by-event fixed effects since they may better capture time-varying shocks at the industry-level.

Panel (a) shows that, when pooling all industries, trends in average profits per establishment seem to be unaffected by minimum wage shocks. Panel (b) shows a similar pattern on the average number of establishments. However, Panel (c) shows a substantial decrease in the average profit per establishment in exposed services, with an implied elasticity of -0.35. Panel (d) shows that this is mainly driven by an intensive margin response, since trends in establishments for these industries also seem to be unaffected by minimum wage changes.[38] While these results should be interpreted with caution since they are based on non-ideal aggregate data of profits and establishments, they suggest that there is substantial profit incidence in industries where this effect is expected.

Capitalist-level fiscal effects The fall in profits in exposed services implies a direct fiscal loss proportional to the corporate tax rate. However, the effect on capitalists' outcomes could generate additional fiscal externalities in other parts of the tax system. Figure 6 shows little support to that hypothesis. In Panels (a) and (b), data varies at the state-by-year level. In Panel (c), data varies at the state-by-industry-by year level. Panel (a) shows no effect on business income per income tax return, Panel (b)

[38]Table 4 suggest a significant elasticity of -0.1 of number of establishments to changes in minimum wages. However, Panel (d) of Figure 5 suggests that the estimated effect is confounded by a differential pre-trend.

shows no effect on dividend income per income tax return, and Panel (c) shows no effect on taxes on production and imports net of subsidies. This suggest that the capitalist-level fiscal effects are mediated by the direct effect on profits and the corresponding loss in corporate tax revenue.

4.4 Back to the optimal policy analysis

Results suggest that minimum wages benefit low-skill workers, hurt capitalists in exposed industries, and generate fiscal savings in income transfers but fiscal costs in corporate tax revenue. I plug-in the estimates on the theoretical results to interpret the estimations through the lens of the optimal policy analysis. A modified version of equation (14) suggests that increasing the minimum wage increases social welfare if[39]

$$\frac{d\log U^l}{d\overline{w}} \cdot U^l \cdot L_A^l \cdot g_1^l + \frac{d\log \Pi^S}{d\overline{w}} \cdot \Pi^S \cdot K_A^S \cdot g_K^S + \text{Fiscal effects} \quad > \quad 0, \tag{19}$$

where I omit the high-skill workers component –because $dU^h/d\overline{w}$ is estimated to be zero– and denote as Π^S the average profit per establishment in exposed services and g_K^S its corresponding marginal welfare weight. The fiscal effects component considers both worker- and capitalist-level fiscal externalities.

$U^l \cdot L_A^l$ equals the sum of total pre-tax income of low-skill workers plus total income maintenance transfers,[40] so the first term of equation (19) can be written as $(\epsilon_{U_{PT}^l} \cdot \text{PTW} + \epsilon_{IT} \cdot \text{IT}) \cdot g_1^l$, where $\epsilon_{U_{PT}^l}$ is the pre-tax version of $d\log U^l/d\overline{w}$, ϵ_{IT} is the fiscal effect on income maintenance transfers, PTW accounts for total annual pre-tax wages, and IT accounts for total income maintenance benefits. Likewise, the second component of equation (19) can be written as $\epsilon_{\Pi^S} \cdot \text{PTP} \cdot (1-t) \cdot g_K^S$, where ϵ_{Π^S} is the profit elasticity on exposed services, and PTP accounts for total annual pre-tax profits of exposed services. Finally, fiscal effects can be written as $-\epsilon_{IT} \cdot \text{IT} + \epsilon_{\Pi^S} \cdot t \cdot \text{PTP}$. Putting altogether, I can write equation (19) as

$$(\epsilon_{U_{PT}^l} \cdot \text{PTW} + \epsilon_{IT} \cdot \text{IT}) \cdot g_1^l + \epsilon_{\Pi} \cdot \text{PTP} \cdot (1-t) \cdot g_K^S - \epsilon_{IT} \cdot \text{IT} + \epsilon_{\Pi^S} \cdot t \cdot \text{PTP} \quad > \quad 0. \tag{21}$$

Values for $\{\epsilon_{U_{PT}^l}, \epsilon_{IT}, \epsilon_{\Pi^S}\}$ can be taken from Tables 2, 3, and 4. I focus on the estimates using the stricter set of time fixed effects for making elasticities comparable and consider two values for ϵ_{Π^S} depending on the interpretation of the extensive margin response estimate.[41] Likewise, values for $\{\text{PTW}, \text{IT}, \text{PTP}\}$ are

[39]The estimated elasticities correspond to the macro version of the sufficient statistics and, therefore, are relevant to the calibration of Proposition II.

[40]Note from equation (13) that

$$U^l \cdot L_A^l \quad = \quad \int E_m^l w_m^l dm - \int E_m^l T(w_m^l) dm + y_0 \cdot L_A^l \cdot \rho^l, \tag{20}$$

that is, $U^l \cdot L_A^l$ equals total pre-tax income plus the net tax liabilities which are composed by the taxes paid by employed workers and the transfers received by the unemployed workers. I use the total income maintenance benefits as a proxy for total net tax liabilities of low-skill workers.

[41]I impute $\epsilon_{U_{PT}^l} = 0.017$ and $\epsilon_{IT} = -0.05$. Regarding ϵ_{Π^S}, I first assume $\epsilon_{\Pi^S} = -0.047$. I also assume that the estimated decrease in number of establishments reported in Table 4 is real, which yields $\epsilon_{\Pi^S} = -0.047 - 0.015 = -0.062$.

directly observed in the data. I follow two approaches for their computation: the population-weighted average of treated states in the pre-event year –to assess the welfare desirability of past minimum wage increases– and the population-weighted average of all states in 2019 –to speculate about the effects of small minimum wages today.[42] Given this, I impute values for t using two assumptions. First, I consider statutory corporate tax rates, thus imputing $t = 35\%$ for assessing past minimum wage increases and $t = 21\%$ for assessing minimum wage increases today. Second, I consider the effective corporate tax rates estimated by Zucman (2014).[43] For the first case I set $t = 20\%$, which is the average value for the period 1997-2017. For the second case I consider $t = 13\%$, which is the most recent available value of the series.

There are two unknowns left to quantitatively assess equation (21): the welfare weights, $\{g_1^l, g_K^S\}$. I calibrate g_K^S and then back-up the welfare weight on low-skill workers that makes equation (21) hold with equality, g_1^{l*}, which can be interpreted as the minimum social value on redistribution toward low-skill workers such that increasing the minimum wage is welfare improving. g_1^{l*} is a measure of the restrictions on social preferences that make the policy change desirable. The smaller g_1^{l*}, the weaker the required preferences for redistribution toward low-skill workers. For this purpose, I follow two approaches to calibrate g_K^S. First, I set $g_K^S = 1$, which emulates an scenario in which the social planner does not have a particular preference to redistribute from or to capitalists. Second, I assume that the social welfare function, G, is given by $G(V) = V^{1-\zeta}/(1-\zeta)$, with $\zeta > 0$. I consider $\zeta \in \{1, 1.5, 2\}$. Under this functional form, higher ζ represents stronger preferences for redistribution, and $\{g_1^l, g_K^S\}$ are endogenous to final allocations. Therefore, relative welfare weights are proportional to average after-tax allocations. Formally, $g_1^l/g_K^S = (U^l/(1-t)\cdot\Pi^S)^{-\zeta}$, so g_K^S can be jointly determined with g_1^{l*}.[44]

Results Table 5 summarizes the results. Each cell reports g_1^{l*} for a different permutation of the 32 calibration choices discussed above. Table 5 suggests that past minimum wage increases have been welfare improving, and that small minimum wage increases today are likely to be as well. When $g_K^S = 1$, the policy change is close to being welfare-neutral. Using a low value for ϵ_{Π^S} yields g_1^{l*} close to 1, meaning that the policy breaks even. Using a higher value for g_1^{l*} yields a minimum marginal welfare weight on

[42] PTW is computed by multiplying the annualized average pre-tax sufficient statistic by state and year by the working-age population and the share of low-skill workers. IT and PTP are observed directly from the raw data.

[43] Effective corporate tax rates are computed by dividing all the corporate taxes paid by US firms (to US and foreign governments) by total US corporate profits using national accounts data taken from the BEA NIPA tables.

[44] For simplicity, I do not consider the participation and entry costs that ultimately matter for the computation of the welfare weights. To get an empirical estimate for the average post-tax sufficient statistic, I compute IT/PTW = 14%, to amplify the annualized average pre-tax sufficient statistic by 14%. When $g_K^S = 1$, g_1^{l*} is given by

$$g_1^{l*} = \frac{-\left(\epsilon_\Pi \cdot \mathrm{PTP}\cdot(1-t) - \epsilon_{IT}\cdot \mathrm{IT} + \epsilon_{\Pi^S}\cdot t \cdot \mathrm{PTP}\right)}{\epsilon_{U^l_{PT}}\cdot \mathrm{PTW} + \epsilon_{IT}\cdot \mathrm{IT}}. \tag{22}$$

When $g_K^S = g_1^l/\omega(\zeta)$, with $\omega(\zeta) = (U^l/(1-t)\cdot\Pi^S)^{-\zeta}$, g_1^{l*} is given by

$$g_1^{l*} = \frac{-\left(-\epsilon_{IT}\cdot \mathrm{IT} + \epsilon_{\Pi^S}\cdot t\cdot \mathrm{PTP}\right)}{\epsilon_{U^l_{PT}}\cdot \mathrm{PTW} + \epsilon_{IT}\cdot \mathrm{IT} + \epsilon_\Pi \cdot \mathrm{PTP}\cdot(1-t)\cdot\omega(\zeta)^{-1}}. \tag{23}$$

low-skill workers that makes equation (21) true equal to 1.52 and 1.54. That is, if the planner does not care about inequality between low-skill workers and exposed services capitalists, a moderate preference for redistribution toward low-wage workers justifies the minimum wage reform. However, when preferences for redistribution are incorporated in the form of a concave social welfare function, the minimum wage becomes unambiguously welfare improving. In all 24 cases, (21) is true even if $g_1^l = 0$.

This highlights the importance of including redistributive preferences to the analysis of the minimum wage policy. Even if total output falls, the incorporation of distributional concerns makes the case for the minimum wage unambiguously favorable. The degree of concavity of the social welfare function (ζ) does not affect the analysis. This is explained by the fact that average post-tax profits are several times larger than average post-tax incomes of active low-skill workers (between five and six times larger), so the redistributive forces in (21) manifest even when the concavity of the social welfare function is moderate.

Intuitively, the empirical analysis shows that minimum wages benefit low-skill workers, hurt firm owners in the exposed industries, and generate fiscal savings in income transfers but fiscal costs in terms of corporate tax revenue. Total after-tax gains for low-skill workers are comparable to total after-tax losses for capitalists. In addition, the net fiscal effect is positive in all cases except when the corporate tax rate is calibrated at 35%, but in all cases the net effect is small relative to baseline incomes: it never represents more than 0.5% of total pre-tax incomes of low-skill workers. This implies that, in the absence of preferences for redistribution, the policy is close to breaking even. When preferences for redistribution enter the analysis, the change in profits only affects the fiscal effect but plays a negligible role in the welfare assessment of the change in after-tax incomes. This makes a positive case for the minimum wage because the distinction between winners and losers is aligned with the social planner's preferences.

5 Policy Applications

The optimal policy analysis uses an equilibrium framework to analyze at a high-level the desirability of the minimum wage. The generality of the model and the reduced-form focus of the sufficient statistics analysis put limits to the insights that can be obtained in terms of concrete policy recommendations.

In this last section, I consider a simplified version of the model with additional structure to explore some additional economic intuitions that are not captured by the previous analysis. I consider three policy applications. First, I focus on the interaction between the minimum wage and the corporate tax rate. Second, I focus on the interaction between the minimum wage and the EITC. Finally, I develop a simple numerical exercise to better understand the interactions at the joint optimum.

5.1 Modified model: simplifications and additional structure

I base the simplification of the model on the following empirical findings discussed in the previous section. First, all groups of low-skill workers seem to experience a similar benefit from minimum wage increases. Second, high-skill workers are not affected by minimum wage changes. Third, the number of private establishments does not respond to minimum wage changes. Fourth, profit responses for capitalists show substantial industry-level heterogeneity. Concretely, the decrease in profits is driven by exposed services while non-exposed services and manufacturing are not affected by minimum wage changes.

Based on these insights, I restrict firm-level heterogeneity to represent a two-industry economy with no within-industry heterogeneity. I consider fixed populations of capitalists indexed by $I = \{M, S\}$ with sizes K_I, where capitalists of type $I = S$ –representing exposed services– only employ low-skill workers, and capitalists of type $I = M$ –representing manufacturing– only employ high-skill workers. The respective production functions are given by $\phi^S(n^l, t)$ and $\phi^M(n^h, t)$. This simplification puts restrictions on the wage distribution but accommodates the fact that firms that are affected by the minimum wage may be different from the non-affected ones.[45] I also assume that the representative firms in both industries are inframarginal, so small changes in the minimum wage do not affect the number of firms.

Services and manufacturing not only differ in their exposure to minimum wage workers, but also in their capital intensity and mobility. Manufacturing is more capital intensive, and its capital is presumably more internationally mobile than the one employed in services industries. This suggests that the behavioral response of profits to changes in corporate taxes is also likely to differ between industries. I explicitly allow for this possibility by incorporating a capital-allocation microfoundation of the dependence of ϕ on t discussed in Appendix A (see discussion in Section 2). In this microfoundation, capitalists are endowed with a fixed stock of capital that have to allocate between domestic and foreign investment. The domestic corporate tax rate distorts the amount of capital invested domestically. Formally, the domestic production function is given by $\widetilde{\phi}^I(n, k, t)$, where k is capital and $\widetilde{\phi}^I(n, k(t), t) = \phi(n, t)$ (see Appendix A). Under this version of the model, the response of pre-tax profits to changes in the corporate tax rate is proportional to the degree of capital mobility. With this additional structure, I can allow for differential capital mobility between services and manufacturing, which I denote by $\varepsilon^I_{k,t}$, for $I \in \{S, M\}$.

The rest of the section revisits the optimal policy analysis using this simplified framework. For simplicity, I assume that the social planner cannot obtain revenue from the returns of the capital invested abroad and that the planner does not consider the returns abroad within the social welfare function.

[45] Under this assumption, the wage distribution consists on two mass points, a low-skill wage (possibly the minimum wage) and a high-skill wage. The lack of heterogeneous effects at the worker-level suggests that abstracting from wage dispersion within skill types is unlikely to play an important role in these applications.

5.2 Minimum wages and corporate tax rates under international capital mobility

The first policy application explores the interaction between the minimum wage and the corporate tax rate, abstracting from the rest of the income tax system. Section 3 suggests a negative interaction between the minimum wage and the corporate tax rate: when corporate tax rates are low, welfare benefits from increasing the minimum wage are larger. This application takes advantage of the two-sector model with capital mobility to better understand the policy implications of this result.

The following proposition shows that, if capital and labor are complements ($\widetilde{\phi}_{kn} > 0$), the desirability of the minimum wage is increasing in the capital mobility of the non-affected sector –M–, while the capital mobility of the affected sector –S– has an ambiguous effect on the optimal minimum wage.[46]

Proposition IV: *Assume that capital and labor are complements. The marginal social welfare of increasing $\overline{w}$ when t is optimal is increasing in $\varepsilon^M_{k,t}$. The effect of $\varepsilon^S_{k,t}$ on the optimal $\overline{w}$ is ambiguous.*
Proof: See Appendix B.

The intuition of this result is the following. Capital mobility has a negative impact on the optimal corporate tax rate. The higher the mobility, the more distortionary is t, and therefore the lower its optimal level. This increases the optimal minimum wage, $\overline{w}$, because both policies redistribute profits so, when t is smaller, then the benefits and costs from increasing $\overline{w}$ are smaller. However, capital mobility in the affected sector ($I = S$) decreases the optimal minimum wage because of similar distortions on domestic capital. This implies that increasing $\varepsilon^M_{k,t}$ unambiguously increases the optimal minimum wage, because it increases the distortion of t without affecting the distortion of $\overline{w}$, but increasing $\varepsilon^S_{k,t}$ has an ambiguous effect on its level given the two forces that work in opposite directions.

This result gives more relevance to the empirical results documented in the previous section. Governments have trouble enforcing large corporate tax rates because of international capital mobility. This is especially driven by capital-intensive industries, such as manufacturing, where corporate tax rates are more likely to have real productive distortions (Kennedy et al., 2022). However, minimum wages do not affect profits in capital intensive industries. By contrast, low-skill labor-intensive industries are unlikely to be affected by capital mobility distortions, but are affected by the minimum wage through the directs effects on wages and profits. That is, in practice, $\varepsilon^M_{k,t}$ is possibly large and $\varepsilon^S_{k,t}$ is possibly low.

This suggests that the minimum wage can be interpreted as a industry-specific corporate tax rate that minimize distortions related to capital mobility. Said differently, the minimum wage can effectively tax

[46]For analytical simplicity, the following proofs assume that technological second-order effects are negligible. Intuitively, when there is a negative shock to firms that pushes optimal vacancies downwards, there is an unintended benefit in marginal productivity given the smaller scale, given decreasing returns to scale. This attenuates the shock but it is unlike to outweight the first-order effect. I assume those second-order effects never dominate the direct effects on vacancies and wages.

profits in affected sectors without distorting capital allocation in other productive non-affected sectors. This insight arises as particularly policy-relevant given the documented decline in effective capital taxation in developed countries due to globalization (Bachas et al., 2022).

5.3 Minimum wages and subsidies to employed low-skill workers

The previous application focus on the interaction between the minimum wage and the corporate tax rate, abstracting from the income tax system. In what follows, I study the interaction between the minimum wage and the income tax rate for employed low-skill workers. The main objective is to better understand whether binding minimum wages complement negative marginal tax rates at the bottom such as the EITC, as previous literature has suggested (OECD, 2009; Lee and Saez, 2012; Hurst et al., 2022).

The following proposition shows that, when the social planner jointly optimizes the minimum wage and the income tax system, binding minimum wages are always optimal. That is, increasing the minimum wage above the market level is always welfare improving. Then, I derive sufficient conditions on the welfare weights on low-skill workers under which optimal minimum wages are complemented by negative marginal tax rates on employed low-skill workers at the joint optimum. The result extends existing results to a more general labor market framework with matching frictions and firm profits.

PROPOSITION V: *It is always optimal to complement the optimal income tax system with a binding minimum wage. Under the optimal binding minimum wage, the optimal marginal tax rate on employed low-skill workers is negative if*

$$g_1^l \quad > \quad \frac{1 - c \cdot \varepsilon_{\theta,\Delta}^l \cdot \left[(1-t) \cdot g_K^S + t\right]}{1 - \varepsilon_{\theta,\Delta}^l} \tag{24}$$

where $c \in (0,1)$ and $\varepsilon_{\theta,\Delta}^l$ is the (absolute value of) the elasticity of low-skill labor market tightness with respect to changes in low-skill net-of-tax wage.

Proof: See Appendix B.

When the income tax system is optimal, increasing the minimum wage just above the market level generates a fiscal gain with no efficiency cost, so a binding minimum wage is always optimal when taxes are optimal. Further increases need to balance the additional fiscal gain with congestion externalities (see Section 3). Intuitively, distortions on vacancy posting are second-order when the firm is optimizing but become relevant as the firm departs from the unrestricted first order condition.

The proposition also states that, under the optimal minimum wage, tax-based transfers to employed low-skill workers in the form of negative marginal tax rates such as the EITC are optimal when the social planner values low-skill workers' welfare beyond a given threshold. The threshold depends on the effects of these transfers on labor market tightness. In the absence of matching frictions, $\varepsilon^l_{\theta,\Delta} = 0$, so equation (24) is reduced to $g^l_1 > 1$, which is the standard result on the desirability of the EITC in frictionless labor markets with extensive margin responses (Lee and Saez, 2012; Piketty and Saez, 2013). Matching frictions and corporate profits add two ingredients to the desirability of negative marginal tax rates on low-skill workers under an optimal minimum wage. Since the minimum wage prevents pre-tax wage adjustments, firms react to the increase in labor supply by decreasing vacancies. This implies that an EITC generates an externality in the congestion of the labor market. That efficiency cost is captured by $\varepsilon^l_{\theta,\Delta}$. This generates two forces that work in opposite directions. First, as captured by the denominator, it generates an inefficiency that makes the critical g^l_1 higher. Second, it generates an increase in profits that may slack the restriction because of two reasons. First, if the planner values redistribution toward firms, this transfer is socially valuable. Second, even if g^S_K is negligible, the transfer to firms allows the social planner to enforce larger optimal corporate tax rates by reducing its distortions on pre-tax profits.

5.4 Numerical analysis

To end the paper, I perform a numerical exercise to illustrates and corroborate the intuitions developed throughout the paper. I calibrate the simplified version of the model and compute the optimal minimum wage change under different tax systems. This informs about the interactions between the policy instruments, and allow a better characterization of the joint optimum away from the local analysis. This subsection discusses the main conclusions of the analysis. All the details about the simulations, including calibration and results, are presented in Appendix D. It is important to emphasize that, given the theoretical simplifications, this numerical exercise should be only taken as suggestive. It is not expected to be interpreted as a robust policy counterfactual whose exact quantitative output is precise enough to inform policy. For robust policy counterfactuals, see Haanwinckel (2020), Ahlfeldt et al. (2022), Berger et al. (2022), Drechsel-Grau (2022), Engbom and Moser (2022), and Hurst et al. (2022).

There are three big conclusions from the numerical analysis. First, given a tax system, social welfare is a globally concave function of the minimum wage, so the model generally generates an interior solution for the optimal minimum wage. This is explained by the fact that wage effects tend to dominate employment effects at low-levels, but employment effects become larger and eventually dominate wage effects as the minimum wage departs from the market level. This result, while expected and present in most labor market models, works as a sanity check for the proposed framework. Second, the optimal minimum wage under fixed taxes varies with the tax parameters. The optimal minimum wage is larger when the EITC is larger –that is, the minimum wage seems to complement the EITC– and is larger when the corporate tax rate is smaller –that is, the minimum wage seems to substitute corporate taxation. Third, the joint optimum seems to use all policies in tandem. That is, optimal redistribution consists on a large EITC, a binding minimum wage far above the market wage, and a non-trivial corporate tax rate.[47] Interestingly, while the minimum wage and the corporate tax rate partially work as substitutes, the planner prefers to use them both at intermediate levels given that each policy's distortion is increasing in its level.

These results reinforce the main message of this paper: the minimum wage can increase the efficiency of tax-based redistribution. Policy-makers should not make the minimum wage compete against the tax system on who is more efficient. Optimally combining all instruments can lead to larger social welfare.

6 Conclusion

The desirability of the minimum wage has been a controversial policy debate for decades. The wide recent evidence on its effects on wages, employment, and other relevant labor market outcomes has encouraged economists to conceptually revisit its role as part of the available instruments for governments. Concerning inequality, a central question is whether there are rationales for governments to use the minimum wage to make tax-based redistribution more efficient. This paper aims to contribute to this discussion.

I propose a general theoretical framework based on an empirically grounded model of the labor market with positive profits that characterizes optimal redistribution for a social planner that can use the income tax system, the corporate tax rate, and a minimum wage to maximize social welfare. I derive conditions that justify the use of the minimum wage as a function of sufficient statistics for welfare, social preferences for redistribution, and fiscal externalities. I find that a minimum wage helps the government redistribute efficiently from capitalists to low-skill workers when it increases average post-tax wages of low-skill labor market participants and when corporate profit incidence is large. Intuitively, the minimum wage can prevent firms from capturing low-wage income subsidies such as the EITC and from enjoying high profits that cannot be redistributed via corporate taxes due to capital mobility in unaffected industries.

Event studies show that the average US state-level minimum wage reform over the last two decades increased average post-tax wages of low-skill labor market participants and reduced corporate profits in affected industries, namely low-skill labor intensive services. It also show a substantial decrease on income maintenance transfers. A sufficient statistics analysis implies that US minimum wages typically remain below their optimum under the current tax and transfer system. The empirical results motivate policy applications that suggest that the minimum wage can be particularly appealing when international tax competition prevents governments from enforcing large corporate tax rates, and that the optimal redistributive package possibly consists on binding minimum wages far above the market wage combined with large EITC subsidies and non-trivial corporate tax rates.

The general message of the paper is that there are rationales to complement tax-based redistribution

[47] Among the cases considered, the optimal policy consists on an EITC of 100%, an hourly minimum wage of $12, and a corporate tax rate of 35%. The optimal minimum wage is substantially larger than the market wage which is simulated to be below $7, and should be considered a lower bound given the efficiency properties of the model (see Section 2).

with binding minimum wages. Governments should not make the tax system and the minimum wage compete for who is the most efficient redistributive policy. By contrast, social planners can benefit from using all instruments together to make redistribution more efficient. Optimal redistribution possibly consists on a binding minimum wage, a corporate tax rate, and a targeted EITC.

Bibliography

Acemoglu, D. (2001). Good jobs versus bad jobs. *Journal of Labor Economics 19*(1), 1–21.

Ahlfeldt, G. M., D. Roth, and T. Seidel (2022). Optimal minimum wages. *Working Paper*.

Allegretto, S. and M. Reich (2018). Are local minimum wages absorbed by price increases? Estimates from internet-based restaurant menus. *ILR Review 71*(1), 35–63.

Alvaredo, F., A. B. Atkinson, T. Piketty, and E. Saez (2013). The top 1 percent in international and historical perspective. *Journal of Economic Perspectives 27*(3), 3–20.

Andreoni, J., B. Erard, and J. Feinstein (1998). Tax compliance. *Journal of Economic Literature 36*(2), 818–860.

Ashenfelter, O. and Š. Jurajda (2022). Minimum wages, wages, and price pass-through: The case of McDonald's Restaurants. *Journal of Labor Economics 40*(S1), S179–S201.

Atesagaoglu, O. E. and H. Yazici (2021). Optimal taxation of capital in the presence of declining labor share. *Working Paper*.

Atkinson, A. B. and J. E. Stiglitz (1976). The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics 6*(1-2), 55–75.

Autor, D., A. Manning, and C. L. Smith (2016). The contribution of the minimum wage to US wage inequality over three decades: A reassessment. *American Economic Journal: Applied Economics 8*(1), 58–99.

Azar, J., S. Berry, and I. E. Marinescu (2019). Estimating labor market power. *Working Paper*.

Bachas, P., M. H. Fisher-Post, A. Jensen, and G. Zucman (2022). Globalization and factor income taxation. *Working Paper*.

Bachas, P. and M. Soto (2021). Corporate taxation under weak enforcement. *American Economic Journal: Economic Policy 13*(4), 36–71.

Bagger, J., E. R. Moen, and R. Vejlin (2021). Equilibrium worker-firm allocations and the deadweight losses of taxation. *Working Paper*.

Baker, A., D. F. Larcker, and C. C. Wang (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics 144*(2), 370–395.

Bassier, I., A. Dube, and S. Naidu (2022). Monopsony in movers: The elasticity of labor supply to firm wage policies. *Journal of Human Resources 57*(S), S50–S86.

Berger, D. W., K. F. Herkenhoff, and S. Mongey (2022). Minimum wages and welfare. *Working Paper*.

Berman, E. (1997). Help wanted, job needed: Estimates of a matching function from employment service data. *Journal of Labor Economics 15*(1), S251–S292.

Bhargava, S. and D. Manoli (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an IRS field experiment. *American Economic Review 105*(11), 3489–3529.

Bíró, A., D. Prinz, L. Sándor, et al. (2022). The minimum wage, informal pay and tax enforcement. *Journal of Public Economics 215*(1), 104728.

Blanchet, T., L. Chancel, and A. Gethin (2022). Why is Europe more equal than the United States? *American Economic Journal: Applied Economics 14*(4), 480–518.

BLS (2020). Characteristics of minimum wage workers, 2020. *Bureau of Labor Statistics*.

Bonhomme, S. and G. Jolivet (2009). The pervasive absence of compensating differentials. *Journal of Applied Econometrics 24*(5), 763–795.

Bozio, A., J. Goupille-Lebret, and M. Guillot (2020). Predistribution vs. redistribution: Evidence from France and the US. *Working Paper*.

Burdett, K. and D. T. Mortensen (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, 257–273.

Cahuc, P. and G. Laroque (2014). Optimal taxation and monopsonistic labor market: Does monopsony justify the minimum wage? *Journal of Public Economic Theory 16*(2), 259–273.

Caldwell, S. and N. Harmon (2019). Outside options, bargaining, and wages: Evidence from coworker networks. *Working Paper*.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics 36*(S1), S13–S70.

Cengiz, D., A. Dube, A. Lindner, and D. Zentler-Munro (2022). Seeing beyond the trees: Using machine learning to estimate the impact of minimum wages on labor market outcomes. *Journal of Labor Economics 40*(S1), S203–S247.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics 134*(3), 1405–1454.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annual Review of Economics 1*(1), 451–488.

Chetty, R., J. N. Friedman, and E. Saez (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings. *American Economic Review 103*(7), 2683–2721.

Clemens, J. (2021). How do firms respond to minimum wage increases? Understanding the relevance of non-employment margins. *Journal of Economic Perspectives 35*(1), 51–72.

Clemens, J., L. B. Kahn, and J. Meer (2018). The minimum wage, fringe benefits, and worker welfare. *Working Paper*.

Coviello, D., E. Deserranno, and N. Persico (2021). Minimum wage and individual worker productivity: Evidence from a large US retailer. *Journal of Political Economy*.

Craig, A. C. (2022). Optimal income taxation with spillovers from employer learning. *American Economic Journal: Economic Policy*.

Cranor, T., J. Goldin, and S. Kotb (2019). Does informing employees about tax benefits increase take-up? Evidence from EITC notification laws. *National Tax Journal 72*(2), 397–434.

Currie, J. (2006). The take up of social benefits. In A. Auerbach and J. Quigley (Eds.), *Distribution of Income, and Public Policy*, pp. 80–148.

de Chaisemartin, C. and X. D'Haultfœuille (2022). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. *Econometrics Journal*.

Derenoncourt, E., C. Noelke, and D. Weil (2021). Spillovers from voluntary employer minimum wages. *Working Paper*.

Devereux, M. P., M. Keen, A. J. Auerbach, P. Oosterhuis, J. Vella, and W. Schön (2021). *Taxing profit in a global economy.* Oxford University Press.

Devereux, M. P., B. Lockwood, and M. Redoano (2008). Do countries compete over corporate tax rates? *Journal of Public Economics 92*(5-6), 1210–1235.

Diamond, P. A. and J. A. Mirrlees (1971). Optimal taxation and public production I: Production efficiency. *The American Economic Review 61*(1), 8–27.

Doligalski, P., A. Ndiaye, and N. Werquin (2022). Redistribution with performance pay. *Working Paper*.

Draca, M., S. Machin, and J. Van Reenen (2011). Minimum wages and firm profitability. *American Economic Journal: Applied Economics 3*(1), 129–51.

Drechsel-Grau, M. (2022). Macroeconomic and distributional effects of higher minimum wages. *Working Paper*.

Drucker, L., K. Mazirov, and D. Neumark (2021). Who pays for and who benefits from minimum wage increases? Evidence from Israeli tax data on business owners and workers. *Journal of Public Economics 199*, 104423.

Dube, A. (2019). Minimum wages and the distribution of family incomes. *American Economic Journal: Applied Economics 11*(4), 268–304.

Dube, A., J. Jacobs, S. Naidu, and S. Suri (2020). Monopsony in online labor markets. *American Economic Review: Insights 2*(1), 33–46.

Dustmann, C., A. Lindner, U. Schönberg, M. Umkehrer, and P. Vom Berge (2022). Reallocation effects of the minimum wage: Evidence from Germany. *Quarterly Journal of Economics 137*(1), 267–328.

Dworczak, P., S. D. Kominers, and M. Akbarpour (2021). Redistribution through markets. *Econometrica 89*(4), 1665–1698.

Eeckhout, J., C. Fu, W. Li, and X. Weng (2021). Optimal taxation and market power. *Working Paper*.

Engbom, N. and C. Moser (2022). Earnings inequality and the minimum wage: Evidence from Brazil. *American Economic Review*.

Feinmann, J., M. Lauletta, and R. Rocha (2022). Employer-employee collusion and payments under the table: Evidence from Brazil. *Working Paper*.

Ferey, A. (2022). Redistribution and unemployment insurance. *Working Paper*.

Ferey, A., B. B. Lockwood, and D. Taubinsky (2022). Sufficient statistics for nonlinear tax systems with general across-income heterogeneity. *Working Paper*.

Finkelstein, A. and M. J. Notowidigdo (2019). Take-up and targeting: Experimental evidence from SNAP. *The Quarterly Journal of Economics 134*(3), 1505–1556.

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

Forsythe, E. (2022). The effect of minimum wage policies on the wage and occupational structure of establishments. *Working Paper*.

Fortin, N. M., T. Lemieux, and N. Lloyd (2021). Labor market institutions and the distribution of wages: The role of spillover effects. *Journal of Labor Economics 39*(S2), S369–S412.

Gandhi, A. and K. Ruffini (2022). Minimum wages and employment composition. *Working Paper*.

Gardner, J. (2021). Two-stage differences in differences. *Working Paper*.

Gaubert, C., P. Kline, D. Vergara, and D. Yagan (2020). Place-based redistribution. *Working Paper*.

Giupponi, G. and S. J. Machin (2022). Company wage policy in a low-wage labor market. *Working Paper*.

Goldin, J. (2018). Tax benefit complexity and take-up: Lessons from the Earned Income Tax Credit. *Tax Law Review 72*, 59.

Guyton, J., P. Langetieg, D. Manoli, M. Payne, B. Schafer, and M. Sebastiani (2017). Reminders and recidivism: Using administrative data to characterize nonfilers and conduct EITC outreach. *American Economic Association: Papers and Proceedings 107*(5), 471–75.

Guyton, J., P. Langetieg, D. Reck, M. Risch, and G. Zucman (2021). Tax evasion at the top of the income distribution: Theory and evidence. *Working Paper*.

Haanwinckel, D. (2020). Supply, demand, institutions, and firms: A theory of labor market sorting and the wage distribution. *Working Paper*.

Hall, R. E. and A. B. Krueger (2012). Evidence on the incidence of wage posting, wage bargaining, and on-the-job search. *American Economic Journal: Macroeconomics 4*(4), 56–67.

Hall, R. E. and S. Schulhofer-Wohl (2018). Measuring job-finding rates and matching efficiency with heterogeneous job-seekers. *American Economic Journal: Macroeconomics 10*(1), 1–32.

Harasztosi, P. and A. Lindner (2019). Who pays for the minimum wage? *The American Economic Review 109*(8), 2693–2727.

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

Hummel, A. J. (2021a). Monopsony power, income taxation and welfare. *Working Paper*.

Hummel, A. J. (2021b). Unemployment and tax design. *Working Paper*.

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. *The Review of Economic Studies 73*(3), 743–767.

Hurst, E., P. Kehoe, E. Pastorino, and T. Winberry (2022). The Distributional Impact of the Minimum Wage in the Short and Long Run. *Working Paper*.

Jäger, S., C. Roth, N. Roussille, and B. Schoefer (2021). Worker beliefs about outside options. *Working Paper*.

Jardim, E., M. C. Long, R. Plotnick, E. Van Inwegen, J. Vigdor, and H. Wething (2022). Minimum-wage increases and low-wage employment: Evidence from seattle. *American Economic Journal: Economic Policy 14*(2), 263–314.

Kennedy, P., C. Dobridge, P. Landefeld, and J. Mortenson (2022). The Efficiency-Equity Tradeoff of the Corporate Income Tax: Evidence from the Tax Cuts and Jobs Act. *Working Paper*.

Kircher, P. (2009). Efficiency of simultaneous search. *Journal of Political Economy 117*(5), 861–913.

Kleven, H. (2021). Sufficient statistics revisited. *Annual Review of Economics 13*.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica 79*(3), 651–692.

Kopczuk, W. and C. Pop-Eleches (2007). Electronic filing, tax preparers and participation in the Earned Income Tax Credit. *Journal of Public Economics 91*(7-8), 1351–1367.

Kroft, K., K. Kucko, E. Lehmann, and J. Schmieder (2020). Optimal income taxation with unemployment and wage responses: A sufficient statistics approach. *American Economic Journal: Economic Policy 12*(1), 254–92.

Ku, H. (2022). Does minimum wage increase labor productivity? Evidence from piece rate workers. *Journal of Labor Economics 40*(2), 325–359.

Lachowska, M., A. Mas, R. Saggio, and S. A. Woodbury (2022). Wage Posting or Wage Bargaining? A Test Using Dual Jobholders. *Journal of Labor Economics 40*(S1), S469–S493.

Lamadon, T., M. Mogstad, and B. Setzler (2022). Imperfect competition, compensating differentials and rent sharing in the US labor market. *American Economic Review 112*(1), 169–212.

Landais, C., P. Michaillat, and E. Saez (2018a). A macroeconomic approach to optimal unemployment insurance: Applications. *American Economic Journal: Economic Policy 10*(2), 182–216.

Landais, C., P. Michaillat, and E. Saez (2018b). A macroeconomic approach to optimal unemployment insurance: Theory. *American Economic Journal: Economic Policy 10*(2), 152–81.

Lavecchia, A. M. (2020). Minimum wage policy with optimal taxes and unemployment. *Journal of Public Economics 190*, 104228.

Le Barbanchon, T., R. Rathelot, and A. Roulet (2021). Gender differences in job search: Trading off commute against wage. *The Quarterly Journal of Economics 136*(1), 381–426.

Lee, D. (1999). Wage inequality in the United States during the 1980s: Rising dispersion or falling minimum wage? *The Quarterly Journal of Economics 114*(3), 977–1023.

Lee, D. and E. Saez (2012). Optimal minimum wage policy in competitive labor markets. *Journal of Public Economics 96*(9-10), 739–749.

Leung, J. H. (2021). Minimum wage and real wage inequality: Evidence from pass-through to retail prices. *Review of Economics and Statistics 103*(4), 754–769.

Lindenlaub, I. and F. Postel-Vinay (2021). The worker-job surplus. *Working Paper*.

Linos, E., A. Prohofsky, A. Ramesh, J. Rothstein, and M. Unrath (2021). Can nudges increase take-up of the EITC?: Evidence from multiple field experiments. *American Economic Journal: Economic Policy*.

MaCurdy, T. (2015). How effective is the minimum wage at supporting the poor? *Journal of Political Economy 123*(2), 497–545.

Maestas, N., K. J. Mullen, D. Powell, T. Von Wachter, and J. B. Wenger (2018). The value of working conditions in the United States and implications for the structure of wages. *Working Paper*.

Mangin, S. and B. Julien (2021). Efficiency in search and matching models: A generalized Hosios condition. *Journal of Economic Theory 193*, 105208.

Mankiw, G. (2013). Some Observations on Minimum Wages, available at http://gregmankiw.blogspot.com/2013/09/some-observations-on-minimum-wages.html.

Manning, A. (2021). The elusive employment effect of the minimum wage. *Journal of Economic Perspectives 35*(1), 3–26.

Mas, A. and A. Pallais (2017). Valuing alternative work arrangements. *American Economic Review 107*(12), 3722–59.

Mayneris, F., S. Poncet, and T. Zhang (2018). Improving or disappearing: Firm-level adjustments to minimum wages in China. *Journal of Development Economics 135*, 20–42.

McCall, L. (2013). *The undeserving rich: American beliefs about inequality, opportunity, and redistribution.* Cambridge University Press.

Mill, J. S. (1884). *Principles of political economy.* D. Appleton.

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

Moen, E. R. (1997). Competitive search equilibrium. *Journal of Political Economy 105*(2), 385–411.

Mousavi, N. (2021). Optimal labor income tax, incomplete markets, and labor market power. *Working Paper*.

OECD (2009). *OECD Employment Outlook 2009: Tackling the Jobs Crisis*. Organisation for Economic Co-operation and Development.

Piketty, T. and E. Saez (2013). Optimal labor income taxation. In *Handbook of Public Economics*, Volume 5, pp. 391–474. Elsevier.

Piketty, T., E. Saez, and S. Stantcheva (2014). Optimal taxation of top labor incomes: A tale of three elasticities. *American Economic Journal: Economic Policy 6*(1), 230–71.

Reich, M. and R. West (2015). The effects of minimum wages on food stamp enrollment and expenditures. *Industrial Relations 54*(4), 668–694.

Renkin, T., C. Montialoux, and M. Siegenthaler (2022). The pass-through of minimum wages into US retail prices: Evidence from supermarket scanner data. *Review of Economics and Statistics 104*(5), 890–908.

Riley, R. and C. R. Bondibene (2017). Raising the standard: Minimum wages and firm productivity. *Labour Economics 44*, 27–50.

Robinson, J. (1933). *The economics of imperfect competition*. Springer.

Roine, J., J. Vlachos, and D. Waldenström (2009). The long-run determinants of inequality: What can we learn from top income data? *Journal of Public Economics 93*(7-8), 974–988.

Roth, J., P. H. Sant'Anna, A. Bilinski, and J. Poe (2022). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Working Paper*.

Rothstein, J. (2010). Is the EITC as good as an NIT? Conditional cash transfers and tax incidence. *American Economic Journal: Economic Policy 2*(1), 177–208.

Roussille, N. and B. Scuderi (2022). Bidding for talent: Equilibrium wage dispersion on a high-wage online job board. *Working Paper*.

Ruffini, K. (2021). Higher wages, service quality, and firm profitability: Evidence from nursing homes and minimum wage reforms. *The Review of Economics and Statistics*.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *The Review of Economic Studies 68*(1), 205–229.

Saez, E. (2002). The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics 83*(2), 217–230.

Saez, E. (2021). Public economics and inequality: Uncovering our social nature. *AEA Papers and Proceedings 111*, 1–26.

Saez, E. and S. Stantcheva (2016). Generalized social marginal welfare weights for optimal tax theory. *The American Economic Review 106*(1), 24–45.

Saez, E. and G. Zucman (2021). Increasing the minimum wage through tax policy. *Working Paper*.

Sleet, C. and H. Yazici (2017). Taxation, redistribution and frictional labor supply. *Working Paper*.

Slemrod, J. and S. Yitzhaki (2002). Tax avoidance, evasion, and administration. In *Handbook of Public Economics*, Volume 3, pp. 1423–1470. Elsevier.

Sockin, J. (2021). Show me the amenity: Are higher-paying firms better all around? *Working Paper*.

Sokolova, A. and T. Sorensen (2021). Monopsony in labor markets: A meta-analysis. *ILR Review 74*(1), 27–55.

Sorkin, I. (2018). Ranking firms using revealed preference. *The Quarterly Journal of Economics 133*(3), 1331–1393.

Staiger, D. O., J. Spetz, and C. S. Phibbs (2010). Is there monopsony in the labor market? Evidence from a natural experiment. *Journal of Labor Economics 28*(2), 211–236.

Stantcheva, S. (2014). Optimal income taxation with adverse selection in the labour market. *The Review of Economic Studies 81*(3), 1296–1329.

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

Taber, C. and R. Vejlin (2020). Estimation of a roy/search/compensating differential model of the labor market. *Econometrica 88*(3), 1031–1069.

Vaghul, K. and B. Zipperer (2016). Historical state and sub-state minimum wage data. *Washington Center for Equitable Growth*.

Vergara, D. (2022). Do policies and institutions matter for pre-tax income inequality? Cross-country evidence. *International Tax and Public Finance 29*(1), 30–52.

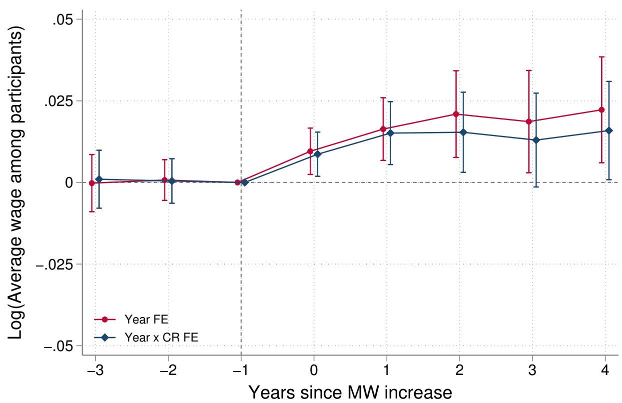
Wolthoff, R. (2018). Applications and interviews: Firms recruiting decisions in a frictional labour market. *The Review of Economic Studies 85*(2), 1314–1351.

Wright, R., P. Kircher, B. Julien, and V. Guerrieri (2021). Directed search and competitive search: A guided tour. *Journal of Economic Literature 59*(1), 90–148.

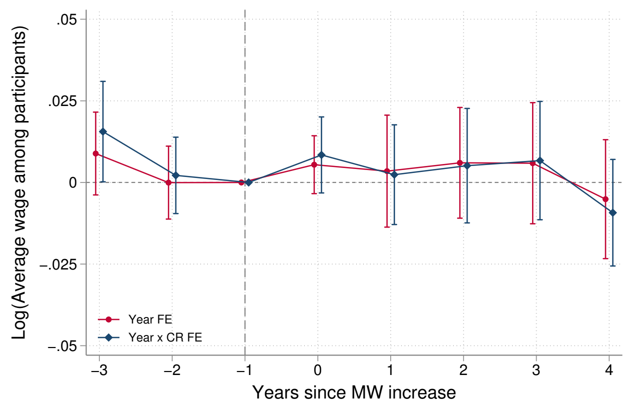
Wu, L. (2021). Partially directed search in the labor market. *Working Paper*.

Zucman, G. (2014). Taxing across borders: Tracking personal wealth and corporate profits. *Journal of Economic Perspectives 28*(4), 121–48.

Figure 1: Changes in workers' welfare after minimum wage increases



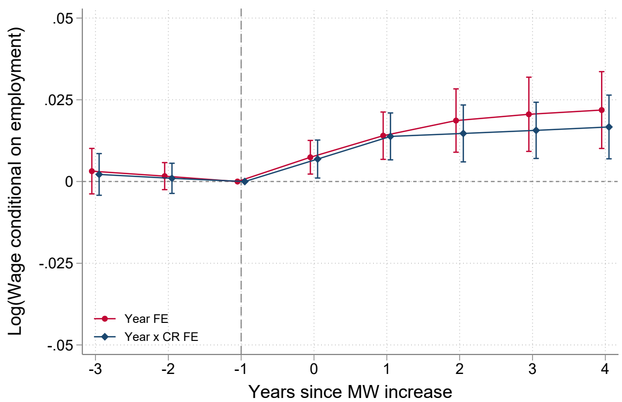
(a) Low-skill workers



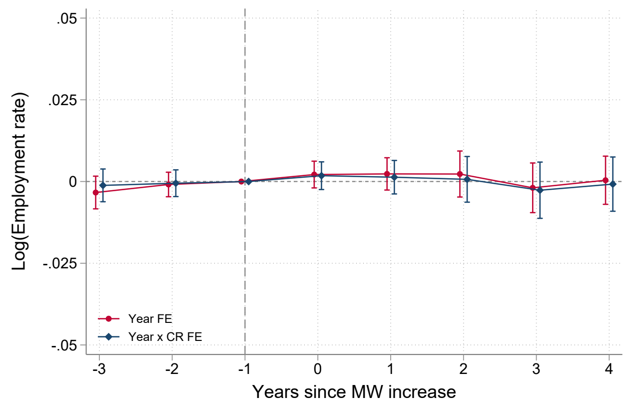
(b) High-skill workers

Notes: These figures plot the estimated β_τ coefficients with their corresponding 95% confidence intervals from equation (17). Panel (a) uses the average pre-tax wage of active low-skill workers including the unemployed as dependent variable. Panel (b) uses the average pre-tax wage of active high-skill workers including the unemployed as dependent variable. Low- and high- skill workers are defined as not having (having) a college degree. Red lines represent specifications that control by year-by-event fixed effects. Blue lines represent specifications that control by census-region-by-year-by-event fixed effects. Standard errors are clustered at the state level, and regressions are weighted by state-by-year average population.

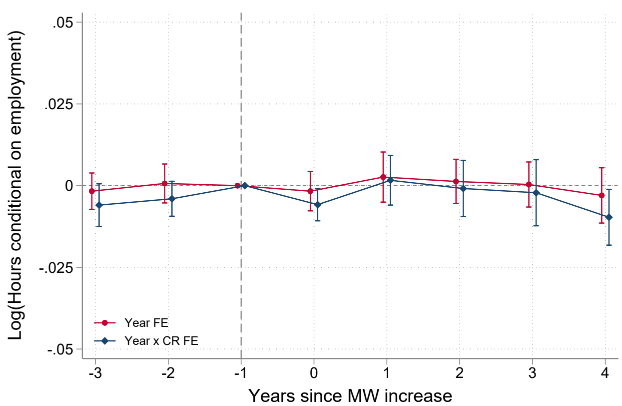
Figure 2: Decomposing the effect of minimum wages on low-skill workers



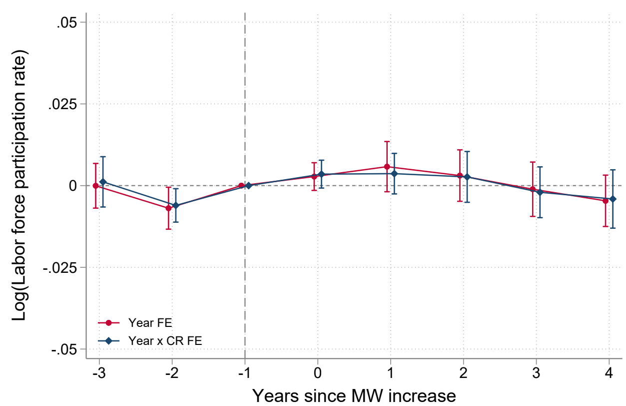
(a) Wage (conditional on employment)



(b) Employment rate



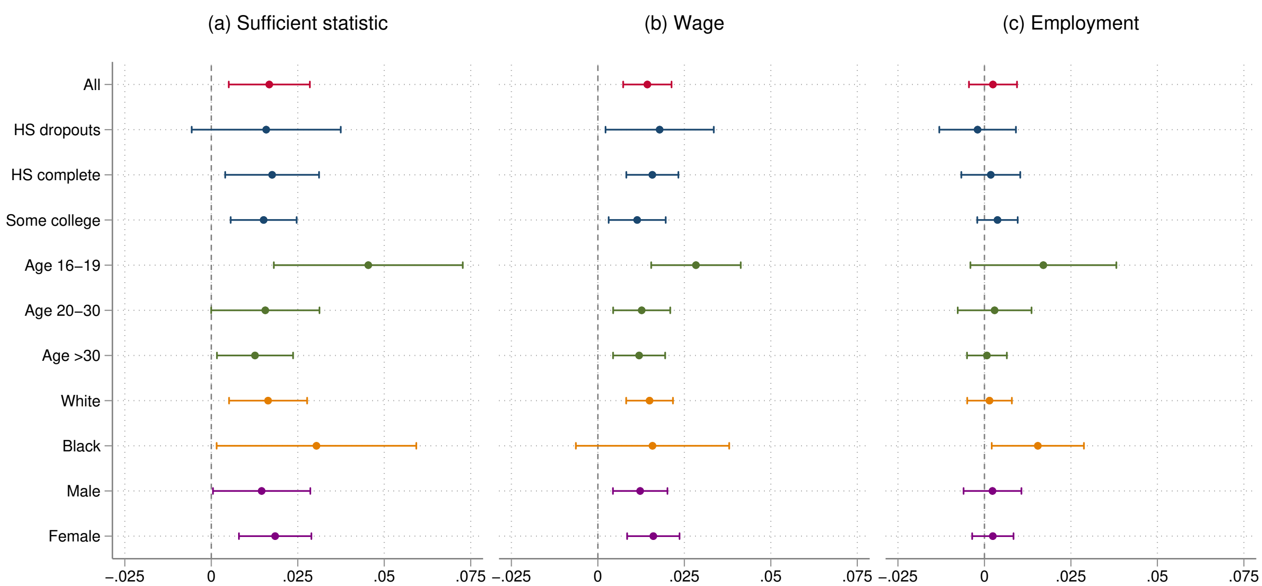
(c) Weekly hours (conditional on employment)



(d) Participation rate

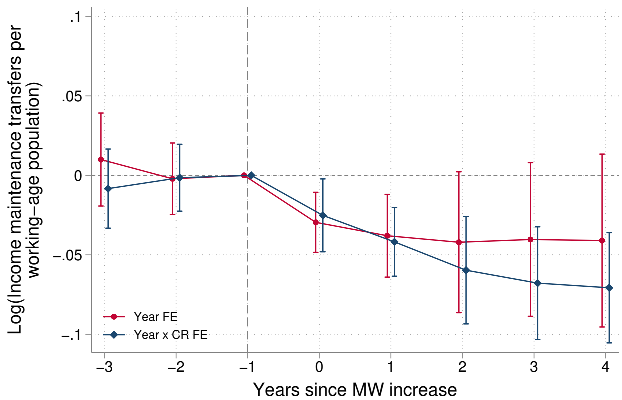
Notes: These figures plot the estimated β_τ coefficients with their corresponding 95% confidence intervals from equation (17). Panel (a) uses the average pre-tax hourly wage of low-skill workers conditional on employment as dependent variable. Panel (b) uses the average employment rate of low-skill workers as dependent variable. Panel (c) uses the average weekly hours worked of low-skill workers conditional on employment as dependent variable. Panel (d) uses the average participation rate of low-skill workers as dependent variable. Low-skill workers are defined as not having a college degree. Red lines represent specifications that control by year-by-event fixed effects. Blue lines represent specifications that control by census-region-by-year-by-event fixed effects. Standard errors are clustered at the state level, and regressions are weighted by state-by-year average population.

Figure 3: Minimum wage effects on low-skill workers: heterogeneity

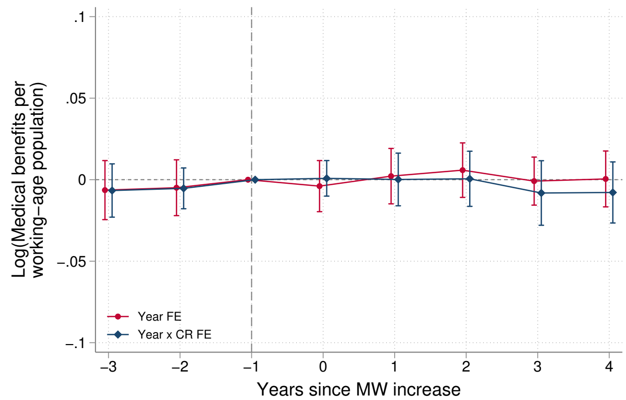


Notes: This figure plots the estimated β coefficient with its corresponding 95% confidence intervals from equation (18) for different groups of low-skill workers and different dependent variables. Panel (a) uses the average pre-tax wage of active low-skill workers including the unemployed as dependent variable. Panel (b) uses the average pre-tax hourly wage of low-skill workers conditional on employment as dependent variable. Panel (c) uses the average employment rate of low-skill workers as dependent variable. Low-skill workers are defined as not having a college degree. Red coefficients reproduce the analysis with the complete sample. Blue coefficients split low-skill workers by education (high-school dropouts, high-school complete, and college incomplete). Green coefficients split low-skill workers by age (16-19, 20-30, and more than 30). Orange coefficients split low-skill workers by race (white and black). Purple coefficients split low-skill workers by sex (male and female). All regressions include year-by-event fixed effects. Standard errors are clustered at the state level, and regressions are weighted by state-by-year average population.

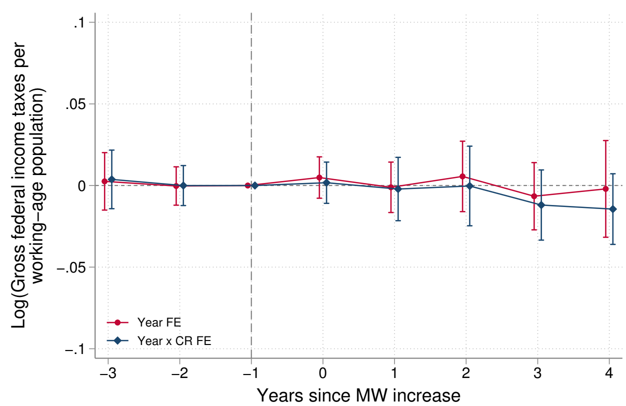
Figure 4: Worker-level fiscal effects after minimum wage increases



(a) Income maintenance transfers



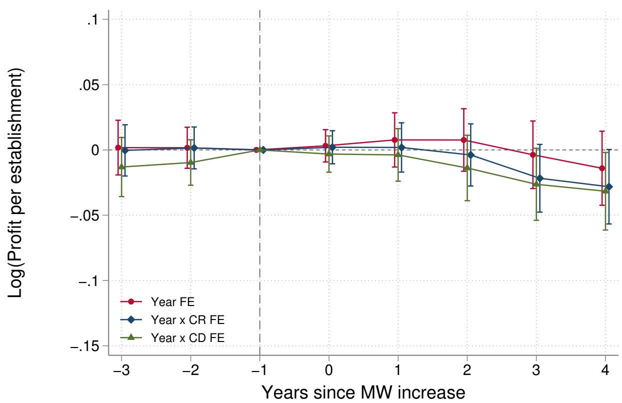
(b) Medical benefits



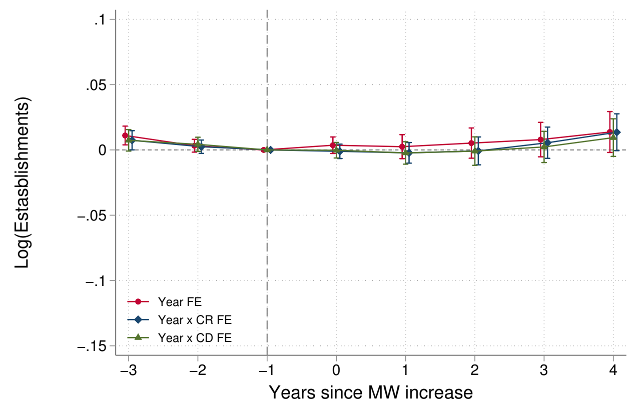
(c) Gross federal income taxes

Notes: These figures plot the estimated β_τ coefficients with their corresponding 95% confidence intervals from equation (17). Panel (a) uses total income maintenance transfers per working-age individual as a dependent variable. Panel (b) uses total medical benefits per working-age individual as a dependent variable. Panel (c) uses total gross federal income taxes per working-age individual as a dependent variable. Red lines represent specifications that control by year-by-event fixed effects. Blue lines represent specifications that control by census-region-by-year-by-event fixed effects. Standard errors are clustered at the state level, and regressions are weighted by state-by-year average population.

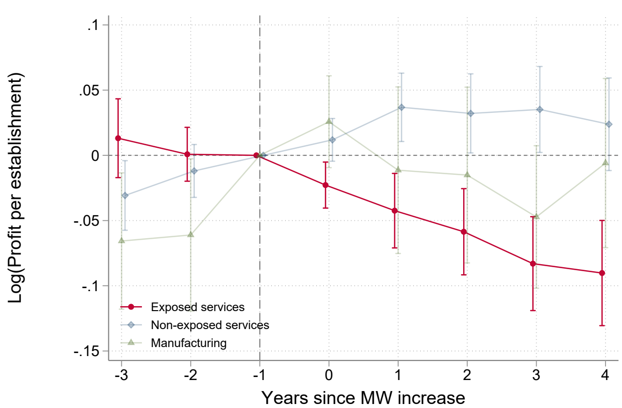
Figure 5: Changes in capitalists' welfare after minimum wage increases



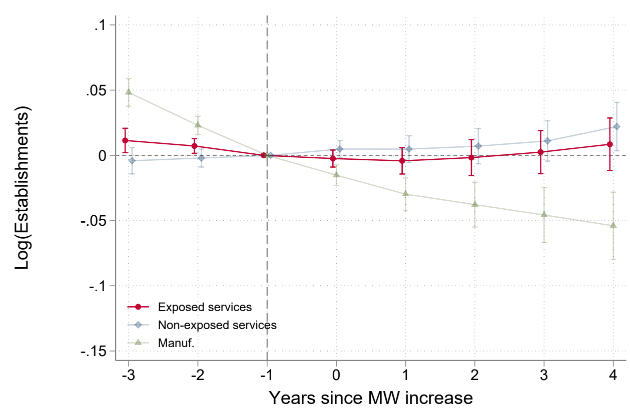
(a) Profit per establishment (all industries)



(b) Establishments (all industries)



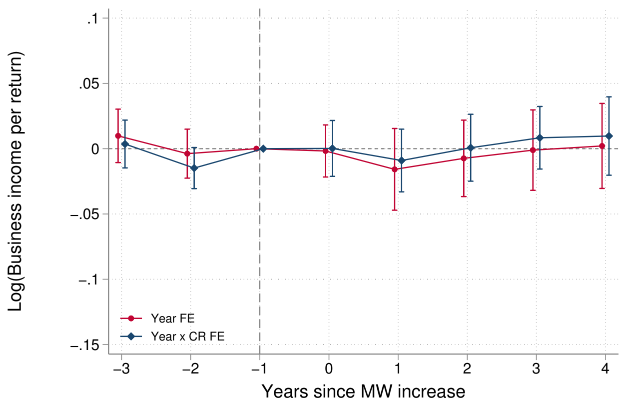
(c) Profit per establishment (by industry)



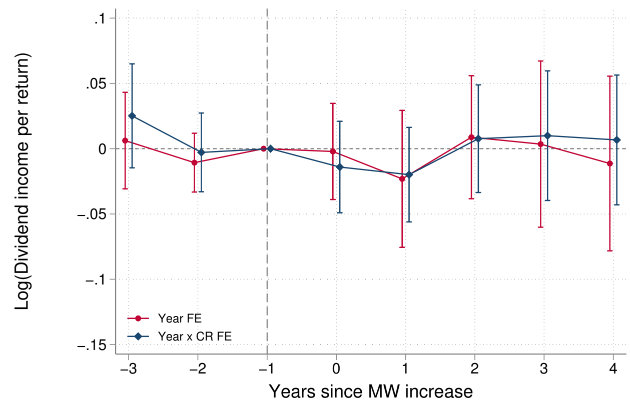
(d) Establishments (by industry)

Notes: These figures plot the estimated β_τ coefficients with their corresponding 95% confidence intervals from equation (17). Panels (a) and (c) use the average profit per establishment as dependent variable. Panels (b) and (d) use the number of establishments as dependent variable. In panels (a) and (b), red lines represent specifications that control by year-by-event fixed effects, blue lines represent specifications that control by census-region-by-year-by-event fixed effects, and green lines represent specifications that control by census-division-by-year-by-event fixed effects. In panels (c) and (d) regressions control by census-division-by-year-by-event fixed effects. Regressions consider a total of 25 industries that are grouped into three categories as follows. Manufacturing industries include SIC codes 41, 43, 44, 46, 50, 54, 56, and 57, that is, nonmetallic mineral products, fabricated metal products, machinery, electrical equipment, food and beverages and tobacco, printing and related support activities, chemical manufacturing, plastics and rubber products. Exposed services include SIC codes 9, 19, 21, 27, 28, and 34, that is, retail trade, ambulatory health services, nursing and residential care facilities, food, accommodation, and social services and other services. Non-exposed services include SIC codes 8, 10, 11, 13, 14, 15, 16, 17, 20, 24 and 25, that is, wholesale trade, transport, information, real estate, professional services, management of businesses, administrative support, educational services, hospitals, arts, and recreation industries. Standard errors are clustered at the state-by-industry level, and regressions are weighted by the average state-by-industry employment in the pre-period.

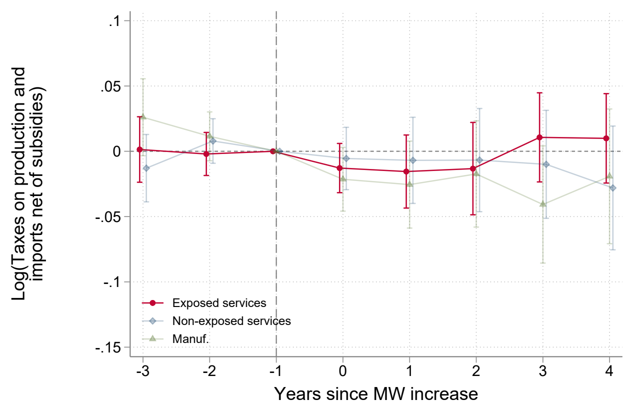
Figure 6: Capitalist-level fiscal effects after minimum wage increases



(a) Business income per tax return



(b) Dividend income per tax return



(c) Taxes on production and imports net of subsidies

Notes: These figures plot the estimated β_τ coefficients with their corresponding 95% confidence intervals from equation (17). In Panels (a) and (b), the unit of observation is at the state-by-year level. In Panel (c), the unit of observation is at the state-by-industry-by-year level. Panel (a) uses the average business income per return as dependent variable. Panel (b) uses the average dividend income per return as dependent variable. Panel (c) uses total taxes on production and imports net of subsidies as dependent variable. In panels (a) and (b), red lines represent specifications that control by year-by-event fixed effects, and blue lines represent specifications that control by census-region-by-year-by-event fixed effects. In panels (c), the regression controls by census-division-by-year-by-event fixed effects. Panel (c) considers a total of 25 industries that are grouped into three categories as follows. Manufacturing industries include SIC codes 41, 43, 44, 46, 50, 54, 56, and 57, that is, nonmetallic mineral products, fabricated metal products, machinery, electrical equipment, food and beverages and tobacco, printing and related support activities, chemical manufacturing, plastics and rubber products. Exposed services include SIC codes 9, 19, 21, 27, 28, and 34, that is, retail trade, ambulatory health services, nursing and residential care facilities, food, accommodation, and social services and other services. Non-exposed services include SIC codes 8, 10, 11, 13, 14, 15, 16, 17, 20, 24 and 25, that is, wholesale trade, transport, information, real estate, professional services, management of businesses, administrative support, educational services, hospitals, arts, and recreation industries. In Panels (a) and (b), standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. In Panel (c), standard errors are clustered at the state-by-industry level, and the regression is weighted by the average state-by-industry employment in the pre-period.

Table 1: Descriptive statistics

	Obs.	Mean	Std. Dev.	Min	Max
Low-skill workers:					
U^l (annualized)	1,173	19,396.69	1,225.82	16,176.45	24,002.46
Hourly wage	1,173	11.55	0.62	9.74	13.99
Weekly hours worked	1,173	34.83	1.57	29.84	38.50
Employment rate	1,173	0.93	0.03	0.79	0.97
Participation rate	1,173	0.61	0.05	0.47	0.72
High-skill workers:					
U^h (annualized)	1,173	61,401.02	7,771.19	42,370.24	89,741.55
Hourly wage	1,173	29.73	3.87	20.70	43.24
Weekly hours worked	1,173	40.85	0.89	37.56	44.02
Employment rate	1,173	0.97	0.01	0.92	1.00
Participation rate	1,173	0.78	0.04	0.65	0.88
Fiscal variables (per working-age individual):					
Income maintenance benefits	1,173	1,056.56	328.81	402.09	2,194.19
Medical benefits	1,173	4,540.73	1,388.37	1,691.10	9,536.34
Gross federal income taxes	1,173	7,179.38	2,091.98	3,780.21	16,346.43
Capitalists:					
Profit per establishment (exposed services)	1,173	170,217.33	50,459.38	95,477.16	539,061.13
Establishments (exposed services)	1,173	70,313.94	103,291.48	5,397	914,454
Profit per establishment (non-exposed services)	1,173	943,530.83	261,578.15	441,352.63	1,765,250.88
Establishments (non-exposed services)	1,173	59,394.48	64,998.93	5,436	438,230
Profit per establishment (manufacturing)	1,173	1,957,057.43	1,359,081.14	-237,418.27	7,436,421.00
Establishments (manufacturing)	1,173	4,314.69	4,723.98	92	30,725

Notes: This table shows descriptive statistics for the non-stacked panel. The unit of observation is a state-year pair. Nominal values are transformed to 2016 dollars using the R-CPI-U-RS index including all items. U^l and U^h are the average pre-tax wage including the unemployed annualized by computing Hourly Wage $\times$ Weekly Hours $\times$ Employment Rate $\times$ 52. Worker-level aggregates are computed using the CPS-MORG data and the Basic Monthly CPS files. Income maintenance benefits, medical benefits, and gross federal income taxes are taken from the BEA regional accounts. Profit per establishment corresponds to the gross operating surplus taken from the BEA regional accounts normalized by the number of private establishments reported in the QCEW data.

Table 2: Worker-level results: sufficient statistic

	Low-skill Workers			**High-skill Workers**		
$\hat{\beta}$	0.017	0.013	0.015	0.000	-0.003	0.002
	(0.006)	(0.006)	(0.005)	(0.007)	(0.006)	(0.008)
Year FE	Y	N	N	Y	N	N
Year x CR FE	N	Y	N	N	Y	N
Year x CD FE	N	N	Y	N	N	Y
Obs.	10,300	10,300	9,653	10,300	10,300	9,653
Events	50	50	50	50	50	50
$\Delta \log \text{MW}$	0.131	0.131	0.127	0.131	0.131	0.127
Elasticity	0.128	0.099	0.120	0.002	-0.023	0.013

Notes: This table shows the estimated β coefficient from equation (18). All columns represent different regressions using different dependent variables (all in logarithms) and fixed effects. Columns 1 to 3 use the average pre-tax wage of low-skill workers including the unemployed (U^l) as dependent variable. Columns 4 to 6 use the average pre-tax wage of high-skill workers including the unemployed (U^h) as dependent variable. Year FE means that the regression includes year-by-event fixed effects. Year x CR FE means that the regression includes year-by-census region-by-event fixed effects. Year x CD FE means that the regression includes year-by-census division-by-event fixed effects. $\Delta \log \text{MW}$ is the average change in the log of the real state-level minimum wage across events in the year of the event. The implied elasticity is computed dividing the point estimate by $\Delta \log \text{MW}$. Standard errors (in parentheses) are clustered at the state level and regressions are weighted by state-by-year population.

Table 3: Worker-level results: fiscal effects

	Income maintenance transfers			**Medical benefits**			**Gross federal income taxes**		
$\hat{\beta}$	-0.040	-0.049	-0.050	0.004	0.001	0.006	-0.000	-0.006	0.005
	(0.015)	(0.012)	(0.015)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
Year FE	Y	N	N	Y	N	N	Y	N	N
Year x CR FE	N	Y	N	N	Y	N	N	Y	N
Year x CD FE	N	N	Y	N	N	Y	N	N	Y
Obs.	10,300	10,300	9,653	10,300	10,300	9,653	10,300	10,300	9,653
Events	50	50	50	50	50	50	50	50	50
$\Delta \log \text{MW}$	0.131	0.131	0.127	0.131	0.131	0.127	0.131	0.131	0.127
Elast.	-0.306	-0.371	-0.389	0.029	0.008	0.044	-0.002	-0.049	0.042

Notes: This table shows the estimated β coefficient from equation (18). All columns represent different regressions using different dependent variables (all in logarithms) and fixed effects. Columns 1 to 3 use total income maintenance transfers per working-age individual as dependent variable. Columns 4 to 6 use total medical benefits per working-age individual as dependent variable. Columns 7 to 9 use total gross federal income taxes per working-age individual as dependent variable. Year x CR FE means that the regression includes year-by-census region-by-event fixed effects. Year x CD FE means that the regression includes year-by-census division-by-event fixed effects. $\Delta \log \text{MW}$ is the average change in the log of the real state-level minimum wage across events in the year of the event. The implied elasticity is computed dividing the point estimate by $\Delta \log \text{MW}$. Standard errors (in parentheses) are clustered at the state level and regressions are weighted by state-by-year population.

Table 4: Capitalist-level results: sufficient statistic

	Profits per establishment				**Establishments**			
	All industries			**Exp. Serv.**	**All industries**			**Exp. Serv.**
$\hat{\beta}$	-0.005	-0.007	-0.007	-0.047	0.012	0.005	-0.000	-0.015
	(0.011)	(0.009)	(0.011)	(0.014)	(0.013)	(0.007)	(0.006)	(0.010)
Year FE	Y	N	N	N	Y	N	N	N
Year x CR FE	N	Y	N	N	N	Y	N	N
Year x CD FE	N	N	Y	Y	N	N	Y	Y
Obs.	519,311	519,311	519,311	519,311	552,792	552,792	552,792	552,792
Events	50	50	50	50	50	50	50	50
$\Delta \log \mathrm{MW}$	0.132	0.132	0.132	0.132	0.131	0.131	0.131	0.131
Elast.	-0.034	-0.053	-0.051	-0.357	0.094	0.041	-0.003	-0.116

Notes: This table shows the estimated β coefficient from equation (18). All columns represent different regressions using different dependent variables (all in logarithms) and fixed effects. Columns 1 to 4 use the average profit per establishment as dependent variable. Columns 5 to 8 use the total number of establishments as dependent variable. Year x CR FE means that the regression includes year-by-census region-by-event fixed effects. Year x CD FE means that the regression includes year-by-census division-by-event fixed effects. All regressions include state-by-industry-by-event fixed effects. $\Delta \log \mathrm{MW}$ is the average change in the log of the real state-level minimum wage across events in the year of the event. The implied elasticity is computed dividing the point estimate by $\Delta \log \mathrm{MW}$. Regressions consider a total of 25 industries that are grouped into three categories as follows. Manufacturing industries include SIC codes 41, 43, 44, 46, 50, 54, 56, and 57, that is, nonmetallic mineral products, fabricated metal products, machinery, electrical equipment, food and beverages and tobacco, printing and related support activities, chemical manufacturing, plastics and rubber products. Exposed services include SIC codes 9, 19, 21, 27, 28, and 34, that is, retail trade, ambulatory health services, nursing and residential care facilities, food, accommodation, and social services and other services. Non-exposed services include SIC codes 8, 10, 11, 13, 14, 15, 16, 17, 20, 24 and 25, that is, wholesale trade, transport, information, real estate, professional services, management of businesses, administrative support, educational services, hospitals, arts, and recreation industries. Standard errors (in parentheses) are clustered at the state-by-industry level and regressions are weighted by the average state-by-industry employment in the pre-period.

Table 5: Welfare effects of minimum wage reforms under fixed taxes

Panel (a): Low ϵ_{Π^S}

	Past minimum wage increases				**Minimum wage increases today**			
	$g_K^S = 1$	$\zeta = 1$	$\zeta = 1.5$	$\zeta = 2$	$g_K^S = 1$	$\zeta = 1$	$\zeta = 1.5$	$\zeta = 2$
Statutory t	0.98	0.00	0.00	0.00	0.99	0.00	0.00	0.00
Effective t	0.98	0.00	0.00	0.00	0.99	0.00	0.00	0.00

Panel (b): High ϵ_{Π^S}

	Past minimum wage increases				**Minimum wage increases today**			
	$g_K^S = 1$	$\zeta = 1$	$\zeta = 1.5$	$\zeta = 2$	$g_K^S = 1$	$\zeta = 1$	$\zeta = 1.5$	$\zeta = 2$
Statutory t	1.52	0.12	0.09	0.08	1.54	0.00	0.00	0.00
Effective t	1.52	0.00	0.00	0.00	1.54	0.00	0.00	0.00

Notes: This table shows estimates for g_1^{l*} for different calibration choices. All cells consider $\epsilon_{U^l_{PT}} = 0.017$ and $\epsilon_{IT} = -0.05$. Panel (a) considers $\epsilon_{\Pi^S} = -0.047$ and Panel (b) considers $\epsilon_{\Pi^S} = -0.062$. Left-panels compute $\{\mathrm{PTW}, \mathrm{IT}, \mathrm{PTP}\}$ using the population-weighted average of treated states in the pre-event year, while right-panels compute $\{\mathrm{PTW}, \mathrm{IT}, \mathrm{PTP}\}$ using the population-weighted average of all states in 2019. Within each sub-panel, columns consider the different approaches for computing g_1^{l*}. The first column assumes $g_K^S = 1$, so g_1^{l*} is computed using equation (22). Columns two to four assume $g_K^S = g_1^l/\omega(\zeta)$, with $\omega(\zeta) = (U^l/(1-t) \cdot \Pi^S)^{-\zeta}$, so g_1^{l*} is computed using equation (23) using $\zeta \in \{1, 1.5, 2\}$. Within each sub-panel, the rows consider either the statutory corporate tax rate or the effective corporate tax rate. The statutory and effective corporate tax rates are (35%, 20%) in the left sub-panel and (21%, 13%) in the right sub-panel, respectively.

Minimum Wages and Optimal Redistribution

Online Appendix

Damián Vergara - UC Berkeley[1]

[1]Email: damianvergara@berkeley.edu. This version: October 31, 2022.

A Theory appendix

Firm's problem The first-order conditions of firms are given by

$$w^s: \qquad (\phi_s - w^s) \cdot \tilde{q}^s_w = \tilde{q}^s, \tag{A.I}$$

$$v^s: \qquad (\phi_s - w^s) \cdot \tilde{q}^s = \eta^s_v, \tag{A.II}$$

for $s \in \{l, h\}$, where $\phi_s = \partial\phi/\partial n^s$ and arguments are omitted from functions to simplify notation. Is direct from the FOCs that wages are below the marginal productivities, that is, that $\phi_s > w^s$. Moreover, defining the firm-specific labor supply elasticity as $\varepsilon^s = (\partial n^s/\partial w^s) \cdot (w^s/n^s) = \widetilde{q}^s_w \cdot w^s/\widetilde{q}^s$, we can rearrange (A.I) and write $\phi_s/w_s = 1/\varepsilon^s + 1$, which is the standard markdown equation (Robinson, 1933). In this model, ε^s is endogenous and finite because of the matching frictions.

Also, combining both FOCs yields $\tilde{q}^{s2} = \eta^s_v \cdot \tilde{q}^s_w$. Differentiating and rearranging terms yields

$$\frac{dw^s}{dv^s} = \frac{\eta^s_{vv} \cdot \tilde{q}^s_w}{2\tilde{q}^s \cdot \tilde{q}^s_w - \eta^s_v \cdot \tilde{q}^s_{ww}} > 0, \tag{A.III}$$

provided $\tilde{q}^s_{ww} < 0$.[2] Moreover, differentiating (A.II) yields

$$(d\phi_s - dw^s) \cdot \tilde{q}^s + (\phi_s - w^s) \cdot \tilde{q}^s_w \cdot dw^s \quad = \quad \eta^s_{vv} \cdot dv^s. \tag{A.IV}$$

Note that

$$d\phi_s \quad = \quad \phi_{ss} \cdot (\tilde{q}^s_w \cdot dw^s \cdot v^s + \tilde{q}^s \cdot dv^s) + \phi_{sj} \cdot \left(\tilde{q}^j_w \cdot dw^j \cdot v^j + \tilde{q}^j \cdot dv^j\right), \tag{A.V}$$

where j is the other skill-type. Replacing (A.I) and (A.V) in (A.IV), yields

$$\left(\phi_{ss} \cdot [\tilde{q}^s_w \cdot dw^s \cdot v^s + \tilde{q}^s \cdot dv^s] + \phi_{sj} \cdot \left[\tilde{q}^j_w \cdot dw^j \cdot v^j + \tilde{q}^j \cdot dv^j\right]\right) \cdot \tilde{q}^s = \eta^s_{vv} \cdot dv^s. \tag{A.VI}$$

[2] Ignoring the superscripts, note that $\tilde{q}_w = q_\theta \cdot (\partial\theta/\partial w)$, which is positive in equilibrium since U is fixed. Then

$$\tilde{q}_{ww} = q_{\theta\theta} \cdot \left(\frac{\partial\theta}{\partial w}\right)^2 + q_\theta \cdot \frac{\partial^2\theta}{\partial w^2}.$$

In principle the sign of $\tilde{q}_{ww}$ is ambiguous, since $q_{\theta\theta} > 0$ and $\partial^2\theta/\partial w^2 > 0$. I assume that the second term dominates so $\tilde{q}_{ww} < 0$. If $\mathcal{M}(L, V) = L^\delta V^{1-\delta}$, $\text{sgn}\,[\tilde{q}_{ww}] = \text{sgn}\left[\frac{-(1-T'(w))^2}{1-\delta} - T''(w)\right]$, so the condition holds as long as the tax system is not too concave. For the result above, $\tilde{q}_{ww} < 0$ is a sufficient but not necessary condition, that is, $\tilde{q}_{ww}$ is allowed to be *moderately* positive, which is plausible since the opposite forces in $\tilde{q}_{ww}$ are interrelated. $q_{\theta\theta} > 0$ follows the concavity and constant returns to scale of the matching function. To see why $\partial^2\theta/\partial w^2 > 0$, recall that $dU = p_\theta \cdot d\theta_m \cdot (w_m - T(w_m) - y_0) + p_m \cdot (1 - T'(w)) \cdot dw_m$. Setting $dU = 0$ and differentiating again yields

$$0 \quad = \quad \left(y_m \cdot p_{\theta\theta} \cdot \frac{\partial\theta_m}{\partial w_m} + 2 \cdot p_\theta \cdot (1 - T'(w_m))\right) \cdot \frac{\partial\theta_m}{\partial w_m} + p_\theta \cdot y_m \cdot \frac{\partial^2\theta_m}{\partial w_m^2} - p_m \cdot T''(w_m),$$

which implies that $\partial^2\theta/\partial w^2 > 0$ as long as the tax system is not "too concave".

Rearranging terms gives

$$\frac{dv^s}{dv^j} = \left[\phi_{sj} \cdot \left(\tilde{q}_w^j \cdot \frac{dw^j}{dv^j} \cdot v^j + \tilde{q}^j\right)\right]^{-1} \cdot \left[\frac{\eta_{vv}^s}{\psi \cdot \tilde{q}^s} - \phi_{ss} \cdot \left(\tilde{q}_w^s \cdot \frac{dw^s}{dv^s} \cdot v^s + \tilde{q}^s\right)\right], \tag{A.VII}$$

which, given (A.III), implies that $\operatorname{sgn}(dv^s/dv^j) = \operatorname{sgn}\phi_{sj}$.

Efficiency properties of the decentralized equilibrium Without loss of generality, consider a case where there is a unique skill type. A social planner that only cares about efficiency decides on sequences of vacancies and applicants to maximize total output net of costs for firms and workers, internalizing the existence of matching frictions. The objective function is given by

$$\mathcal{V} = K \cdot \int_{\psi^*}^{\overline{\psi}} [\phi(\psi, n) - \eta(v_\psi) - \xi]\, dO(\psi) - \alpha \cdot \int_0^{c^*} c \cdot dF(c), \tag{A.VIII}$$

subject to

$$n = q\left(\frac{K \cdot v_\psi \cdot o(\psi)}{L_\psi}\right) \cdot v_\psi, \tag{A.IX}$$

$$\int_{\psi^*}^{\overline{\psi}} L_\psi d\psi = \alpha \cdot F(c^*), \tag{A.X}$$

where $\{c^*, \psi^*\}$ are the thresholds for workers and firms to enter the labor market, and $\{v_\psi, L_\psi\}$ are the sequences of vacancies and applicants, with $\theta_\psi = (K \cdot v_\psi \cdot o(\psi))/L_\psi$. The planner chooses $\{c^*, \psi^*\}$ and $\{v_\psi, L_\psi\}$ to maximize (A.VIII) subject to (A.IX) (matches are endogenous to the number of applicants and vacancies) and (A.X) (the distribution of applicants across firms has to be consistent with the number of active workers). The Lagrangian is given by

$$\begin{aligned}\mathcal{L} = {} & K \cdot \int_{\psi^*}^{\overline{\psi}} \left[\phi\left(\psi, q\left(\frac{K \cdot v_\psi \cdot o(\psi)}{L_\psi}\right) \cdot v_\psi\right) - \eta(v_\psi) - \xi\right] dO(\psi) \\ & -\alpha \cdot \int_0^{c^*} c \cdot dF(c) + \mu \cdot \left[\alpha \cdot F(c^*) - \int_{\psi^*}^{\overline{\psi}} L_\psi d\psi\right],\end{aligned} \tag{A.XI}$$

where μ is the multiplier. The first order conditions with respect to v_ψ, L_ψ, and c^* are given by

$$v_\psi : \quad \phi_n \cdot (q_\theta \cdot \theta_\psi + q) = \eta_v, \tag{A.XII}$$

$$L_\psi : \quad -\theta_\psi^2 \cdot q_\theta \cdot \phi_n = \mu, \tag{A.XIII}$$

$$c^* : \quad -\alpha \cdot c^* \cdot f(c^*) + \mu \cdot \alpha \cdot f(c^*) = 0. \tag{A.XIV}$$

Equation (A.XIV) implies that $\mu = c^{*}$. Using that and combining (A.XII) and (A.XIII) yields

$$q \cdot \phi_n - \frac{c^{*}}{\theta_\psi} \quad = \quad \eta_v. \tag{A.XV}$$

To assess the efficiency of vacancy posting decisions and applications decisions, I check whether (A.XV) is consistent with the decentralized equilibrium. In the absence of taxes, the threshold for workers' entry is given by $U = p(\theta_\psi) \cdot w_\psi$, that holds for any ψ. We also know, from the properties of the matching function, that $p(\theta_\psi) = \theta_\psi \cdot q(\theta_\psi)$. Replacing in (A.XV) yields $q \cdot (\phi_n - w_\psi) = \eta_v$, which coincides with the decentralized first order condition of the firms for vacancies (see equation (A.II)). Then, the decentralized equilibrium is efficient in terms of vacancy posting and applications.

The first order condition with respect to ψ^{*} is given by

$$\psi^{*}: \qquad -K \cdot \left(\phi\left(\psi^{*}, q\left(\theta_{\psi^{*}}\right) \cdot v_{\psi^{*}}\right) - \eta\left(v_{\psi^{*}}\right) - \xi\right) \cdot o\left(\psi^{*}\right) - \mu \cdot L_{\psi^{*}} = 0. \tag{A.XVI}$$

Equation (A.XVI) can be written as

$$\phi\left(\psi^{*}, q\left(\theta_{\psi^{*}}\right) \cdot v_{\psi^{*}}\right) - \eta\left(v_{\psi^{*}}\right) - \frac{\mu \cdot L_{\psi^{*}}}{K \cdot o\left(\psi^{*}\right)} \quad = \quad \xi \tag{A.XVII}$$

Note that $\left(\mu \cdot L_{\psi^{*}}\right) / \left(K \cdot o\left(\psi^{*}\right)\right) = \left(c^{*} \cdot v_{\psi^{*}}\right) / \left(\theta_{\psi^{*}}\right)$. Then, $c^{*} = p\left(\theta_\psi\right) \cdot w_{\psi^{*}}$ and $p(\theta_{\psi^{*}}) = \theta_{\psi^{*}} \cdot q(\theta_{\psi^{*}})$ imply that $\left(c^{*} \cdot v_{\psi^{*}}\right) / \theta_{\psi^{*}} = w_{\psi^{*}} \cdot q(\theta_{\psi^{*}}) \cdot v_{\psi^{*}}$, which implies that equation (A.XVI) is equivalent to $\Pi\left(\psi^{*}\right) = \xi$, which coincides with the decentralized equilibrium. Therefore, the decentralized equilibrium is efficient. $\square$

Microfounding $\phi(\psi, n^{l}, n^{h}, t)$ I provide two examples to microfound the production function as a function of t. I consider a capital-allocation problem, and an effort-allocation problem.

Regarding the capital-allocation problem, assume that the structural production function, $\widetilde{\phi}$ depends on capital, k, as well as the other inputs described in Section 2, with $\widetilde{\phi}_k > 0$ and $\widetilde{\phi}_{kk} < 0$. Capitalists have a fixed endowment of capital, $\overline{k}$, that can be invested domestically, k_D, or abroad, k_A, with $k_D + k_A = \overline{k}$. If capitalists invest k_D domestically, they get after-tax profits $(1-t) \cdot \widetilde{\Pi}(k_D, \psi)$, where $\widetilde{\Pi}(k_D, \psi)$ is the value function that optimizes wages and vacancies given capital. If capitalists invest k_A abroad, they get $k_A \cdot r^{*} \cdot (1-t^{*})$, where $r^{*} \cdot (1-t^{*})$ is the after-tax return of capital abroad. Capitalists choose k_D to maximize $(1-t) \cdot \widetilde{\Pi}(k_D, \psi) + \left(\overline{k} - k_D\right) \cdot r^{*} \cdot (1-t^{*})$. The first-order condition is given by $(1-t) \cdot \widetilde{\Pi}_k = r^{*} \cdot (1-t^{*})$, which characterizes the optimal capital invested domestically, k_D^{*}, as a function of ψ and t.[3] Then, $k_D^{*} = k_D(\psi, t)$, so $\widetilde{\phi}(k_D(\psi, t), n^{l}, n^{h}, t) = \phi(\psi, n^{l}, n^{h}, t)$ and $\widetilde{\Pi}(k_D(\psi, t), \psi) = \Pi(\psi, t)$. If capitalists have no investment opportunities abroad or transportation costs of capital are large (which would be analogous

[3] This is well-defined given decreasing returns to capital in $\widetilde{\phi}$. Since capitalists own the capital, we have that $\widetilde{\Pi}_k = \partial \widetilde{\Pi}(k_D, \psi) / \partial k_D > 0$ and $\widetilde{\Pi}_{kk} = \partial^2 \widetilde{\Pi}(k_D, \psi) / \partial k_D^2 < 0$.

to $r^* \cdot (1-t^*) \to 0$), then $k_D = \overline{k}$ and the production function no longer depends on t.

A similar argument can be developed with respect to the minimum wage, $\overline{w}$. If $\overline{w}$ binds, then it also affects the allocation of capital to domestic investment. Then, ϕ can also be written as a function of $\overline{w}$.

Under this formulation, the behavioral response of profits to corporate taxes and minimum wages can be written as a formula of capital mobility. Define the elasticity of domestic capital to changes in corporate taxes by $\varepsilon_{k,t} = -(\partial k_D/\partial t) \cdot (t/k_D)$, and the elasticity of domestic capital to changes in the minimum wage (when it binds) by $\varepsilon_{k,\overline{w}} = -(\partial k_D/\partial \overline{w}) \cdot (\overline{w}/k_D)$. $\varepsilon_{k,t}$ and $\varepsilon_{k,\overline{w}}$ are interpreted as the magnitude (absolute value) of the behavioral response. Both elasticities are related through the technological role of capital in the production function. Formally, $\varepsilon_{k,\overline{w}} = a \cdot \varepsilon_{k,t}$, with $a > 0$.[4]

Then, pre-tax domestic profit effects to changes in the policy parameters are given by:

$$\frac{\partial \Pi\left(\psi, t, \overline{w}\right)}{\partial t} = \frac{\widetilde{\Pi}\left(k_D(\psi, t, \overline{w}), \psi, \overline{w}\right)}{\partial t} = \widetilde{\Pi}_k \cdot \frac{\partial k_D}{\partial t} = -\gamma_t \cdot \varepsilon_{k,t} < 0, \tag{A.XVIII}$$

where $\gamma_t = \widetilde{\Pi}_k \cdot k_D/t > 0$, and

$$\frac{\partial \Pi\left(\psi, t, \overline{w}\right)}{\partial \overline{w}} = \frac{\partial \widetilde{\Pi}\left(k_D(\psi, t, \overline{w}), \psi, \overline{w}\right)}{\partial \overline{w}} = \widetilde{\Pi}_k \cdot \frac{\partial k_D}{\partial \overline{w}} + \widetilde{\Pi}_{\overline{w}} = -\gamma_{\overline{w}} \cdot \varepsilon_{k,t} + \widetilde{\Pi}_{\overline{w}} < 0, \tag{A.XIX}$$

where $\gamma_{\overline{w}} = \widetilde{\Pi}_k \cdot k_D \cdot a/\overline{w} > 0$. Note that the envelope theorem does not hold for this object since pre-tax profits of domestic firms represent only a fraction of the value function of the capital allocation problem. The effect of corporate tax rates on pre-tax profits is driven by the reduction in capital. The effect of minimum wages on pre-tax profits is driven by both the reduction in capital and the direct effect on labor costs (see below).

As an alternative microfoundation, assume that the structural production function, $\widetilde{\phi}$, depends on the managerial effort of the capitalist, e, as well as the other inputs described in Section 2, with $\widetilde{\phi}_e > 0$ and $\widetilde{\phi}_{ee} < 0$. If capitalists exert effort e, they get after-tax profits $(1-t) \cdot \widetilde{\Pi}(e, \psi)$, where $\widetilde{\Pi}(e, \psi)$ is the value function that optimizes wages and vacancies given effort. Exerting effort e has a cost $c(e)$, with $c_e > 0$ and $c_{ee} > 0$. Optimal effort solves the first order condition $(1-t) \cdot \widetilde{\Pi}_e = c_e$, which characterizes the optimal effort, e^*, as a function of ψ and t. Then, $e^* = e(\psi, t)$, so $\widetilde{\phi}(e(\psi, t), n^l, n^h, t) = \phi(\psi, n^l, n^h, t)$ and $\widetilde{\Pi}(e(\psi, t), \psi) = \Pi(\psi, t)$. If effort plays little role in revenue or the costs are negligible, then the production function no longer depends on t.

[4] Differentiating the first-order condition and setting $dr^* = dt^* = 0$ yields:

$$-dt \cdot \widetilde{\Pi}_k + (1-t) \cdot \left(\widetilde{\Pi}_{kk} \cdot dk_D + \widetilde{\Pi}_{k\overline{w}} \cdot d\overline{w}\right) = 0,$$

where $\widetilde{\Pi}_{k\overline{w}} < 0$. Then, $\varepsilon_{k,t} = -(\partial k_D/\partial t) \cdot (t/k_D) = -(\widetilde{\Pi}_k \cdot t)/((1-t) \cdot \widetilde{\Pi}_{kk} \cdot k_D)$, and $\varepsilon_{k,\overline{w}} = -(\partial k_D/\partial \overline{w}) \cdot (\overline{w}/k_D) = (\widetilde{\Pi}_{k\overline{w}} \cdot \overline{w})/((1-t) \cdot \widetilde{\Pi}_{kk} \cdot k_D) = a \cdot \varepsilon_{k,t}$, where $a = -\widetilde{\Pi}_{k\overline{w}} \cdot \overline{w}/\widetilde{\Pi}_k \cdot t > 0$.

Firms' responses to changes in the minimum wage To see the effect of the minimum wage on firms' decisions, note that the four first order conditions (equations (A.I) and (A.II) for $s = \{l, h\}$) hold for firms that are not constrained by the minimum wage, while (A.I) no longer holds for firms that are constrained by the minimum wage. Then, for firms that operate in sub-markets with $w_m^l > \overline{w}$, it is sufficient to verify the reaction of one of the four endogenous variables to changes in the minimum wage and use the within-firm correlations to predict reactions in the other variables. For firms that operate in sub-markets where $w_m^l = \overline{w}$, it is necessary to first compute the change in low-skill vacancies and then infer the changes in high-skill vacancies and wages using the within-firm between-skill correlations that still hold for the firm.

In both cases, it is easier to work with equation (A.II) for $s = l$. For an unconstrained firm, totally differentiating the first order condition yields

$$\begin{aligned}\left(\left[\phi_{ll} \cdot \left(q_\theta^l \cdot d\theta^l \cdot v^l + q^l \cdot dv^l\right) + \phi_{lh} \cdot \left(q_\theta^h \cdot d\theta^h \cdot v^h + q^h \cdot dv^h\right)\right] - dw^l\right) \cdot q^l & \\ +(\phi_l - w^l) \cdot q_\theta^l \cdot d\theta^l &= \eta_{vv}^l \cdot dv^l, \end{aligned} \tag{A.XX}$$

where I omitted sub-market sub-indices to simplify notation. Rearranging terms gives

$$\begin{aligned} & dw^l \cdot \left[\frac{dv^l}{dw^l} \cdot \left(\eta_{vv}^l - \phi_{ll} \cdot q^{l2} - \phi_{lh} \cdot q^h \cdot q^l \cdot \frac{dv^h}{dv^l}\right) + q^l\right] \\ & = d\theta^l \cdot q_\theta^l \cdot \left[(\phi_l - w^l) + \phi_{ll} \cdot v^l \cdot q^l\right] + d\theta^h \cdot q_\theta^h \cdot \phi_{lh} \cdot q^l. \end{aligned} \tag{A.XXI}$$

Note that the sign and magnitude of $dw^l/d\overline{w}$ depends on $d\theta^l/d\overline{w}$. With the variation in wages it is possible to predict variation in vacancies (and, therefore, firm-size) and spillovers to high-skill workers.

On the other hand, for a constrained firm, totally differentiating the first order condition yields

$$\begin{aligned}\left(\left[\phi_{ll} \cdot \left(q_\theta^l \cdot d\theta^l \cdot v^l + q^l \cdot dv^l\right) + \phi_{lh} \cdot \left(q_\theta^h \cdot d\theta^h \cdot v^h + q^h \cdot dv^h\right)\right] - d\overline{w}\right) \cdot q^l & \\ +(\phi_l - \overline{w}) \cdot q_\theta^l \cdot d\theta^l &= \eta_{vv}^l \cdot dv^l, \end{aligned} \tag{A.XXII}$$

where I omitted sub-market sub-indices to simplify notation. Rearranging terms gives

$$\begin{aligned}\frac{dv^l}{d\overline{w}} \cdot \left(\eta_{vv}^l - \phi_{ll} \cdot q^{l2} - \phi_{lh} \cdot q^h \cdot q^l \cdot \frac{dv^h}{dv^l}\right) &= \frac{d\theta^l}{d\overline{w}} \cdot q_\theta^l \cdot \left[(\phi_l - \overline{w}) + \phi_{ll} \cdot v^l \cdot q^l\right] \\ & + \frac{d\theta^h}{d\overline{w}} \cdot q_\theta^h \cdot \phi_{lh} \cdot q^l - q^l. \end{aligned} \tag{A.XXIII}$$

The sign and magnitude depends on the reaction on equilibrium sub-market tightness. However, note that the first-order effect is decreasing in productivity, since ϕ_l is decreasing in ψ and $(\phi_l - \overline{w}) \to 0$ as $\overline{w}$ increases. That is, among firms that pay the minimum wage, the least productive ones are more likely to decrease their vacancies, and therefore shrink and eventually exit the market.

Finally, to see the effect of the minimum wage on profits, we can use the envelope theorem and conclude that the total effect is equal to the partial effect ignoring general equilibrium changes on endogenous variables. This implies that for constrained firms

$$\frac{d\Pi(\psi,t)}{d\overline{w}} = \frac{\partial\Pi(\psi,t)}{\partial\overline{w}} = q^l_\theta \cdot \frac{\partial\theta^l}{\partial\overline{w}} \cdot v^l \cdot (\phi_l - \overline{w}) - v^l \cdot q^l. \tag{A.XXIV}$$

This effect is negative given that the first-order condition with respect to low-skill wages holds with inequality and is stronger for less productive firms. For unconstrained firms

$$\frac{d\Pi(\psi,t)}{d\overline{w}} = \frac{\partial\Pi(\psi,t)}{\partial\overline{w}} = q^l_\theta \cdot \frac{\partial\theta^l}{\partial\overline{w}} \cdot v^l \cdot (\phi_l - w^l), \tag{A.XXV}$$

so the effect on profits is uniquely mediated by the effect on job-filling probabilities.

Employment effects and workers' welfare For simplicity, assume away taxes. In equilibrium, $U^s = p^s_m \cdot w^s_m$. Multiplying by L^s_m at both sides and integrating over m yields $L^s_A \cdot U^s = \int E^s_m w^s_m dm$, where $E^s_m = L^s_m \cdot p^s_m$ is the mass of employed workers of skill s in sub-market m. Differentiating gives

$$\frac{dU^s}{d\overline{w}} \cdot (L^s_A + U^s \cdot \alpha_s \cdot f_s(U^s)) \quad = \quad \int \left(\frac{dE^s_m}{d\overline{w}} \cdot w^s_m + E_m \cdot \frac{dw^s_m}{d\overline{w}} \right) dm, \tag{A.XXVI}$$

where I used $L^s_A = \alpha_s \cdot F_s(U^s)$. The left-hand side is the welfare effect on workers times a positive constant. Then, the right-hand side can be used to calculate the wage-weighted disemployment effects, $\int (dE^s_m/d\overline{w}) w^s_m dm$, that can be tolerated for the minimum wage to increase aggregate welfare for workers given employment-weighted wage effects. If both employment and wage effects are positive, the welfare effect on workers is unambiguously positive.

Additional discussion on the limitations of the model I briefly discuss the implications of abstracting from dynamics, intensive margin responses, and informal labor markets.

Dynamics: The model is static. The implications of this assumption for the optimal policy analysis are, in principle, ambiguous. Dube et al. (2016) and Gittings and Schmutte (2016) show that minimum wage shocks decrease employment flows –separation, hires, and turnover rates– while keeping the employment stock constant, thus increasing job stability. In the presence of labor market frictions, this induces a dynamic efficiency gain from minimum wage increases that is not captured by the model. On the other hand, Sorkin (2015), Aaronson et al. (2018), and Hurst et al. (2022) argue that the long-run employment distortions of minimum wage shocks are larger than the short-run responses, because of long-run capital substitution through technological change.

Intensive margin responses: The model assumes segmented labor markets. This assumption implies that the model abstracts from intensive margin responses (Saez, 2002). For example, increasing the minimum wage could induce high-skill workers to apply to low-skill vacancies. To the extent that these responses are empirically relevant, this is a caveat of the policy analysis. Note that this is different from changes in demand for skills, as suggested by Butschek (2022) and Clemens et al. (2021). The model can rationalize this by changes in the skill composition of posted vacancies mediated by ϕ. Intensive margin responses could also affect incentives conditional on labor market segmentation. For example, workers may want to work more hours if the after-tax wage increases, or firms could offer jobs with shorter schedules (Jardim et al., 2022). This mechanism is muted in the model, mainly motivated by the fact that at the bottom of the wage distribution extensive margin responses tend to play a more important role to understand workers behavior. My empirical analysis finds no effect on hours worked conditional on employment, providing empirical support to the assumption.

Informality: In some contexts, the interaction between the minimum wage and the degree of formality of the labor market may be a first order consideration. In the model, the costs of participating in the labor market, which are not taxed, may rationalize heterogeneity in outside options, including informal labor market opportunities. However, changing the characteristics of the formal sector may affect both the supply and demand for formal jobs. For detailed analyses, see Bosch and Manacorda (2010), Meghir et al. (2015), Pérez (2020), and Haanwinckel and Soares (2021).

Additional discussion on the empirical effects of minimum wages I briefly discuss the price and productivity effects documented in the empirical literature and its implications for the policy analysis.

Price effects: The model assumes that output prices are fixed, ruling out price increases after minimum wage shocks. However, the empirical literature finds substantial passthrough to prices (Allegretto and Reich, 2018; Harasztosi and Lindner, 2019; Renkin et al., 2022; Leung, 2021; Ashenfelter and Jurajda, 2022). Modeling price increases after minimum wage shocks in the presence of limited employment effects is challenging: if employment does not fall and demand curves are downward sloping, prices should decrease rather than increase. Bhaskar and To (1999) and Sorkin (2015) reconcile limited employment effects with price increases in dynamic frameworks. Price effects matter for welfare since they can erode nominal minimum wage increases. Also, the unemployed and non-employed households can be made worse off given the absence of nominal improvements (MaCurdy, 2015). The distributional effect depends on which consumers buy the goods produced by firms that pay the minimum wage, and the relative importance of these goods in aggregate consumption. It also depends on the share of minimum wage workers since it affects the mapping from product-level prices to economy-level price indexes.

While more research is needed to assess the distributional impacts of the price effects, the available evidence suggests that they are unlikely to play a big role in the aggregate distributional analysis.

Minimum wage workers represent a small share of the aggregate labor market, so it is unlikely that a small share of price increases can have first-order effects on aggregate price indexes. Also, Harasztosi and Lindner (2019) show that the goods produced by firms that pay the minimum wage are evenly consumed across the income distribution, which neutralizes the potential unintended consequences through redistribution from high-income consumers to low-skill workers. Ashenfelter and Jurajda (2022) analyze McDonald's restaurants responses to local minimum wage shocks and show that the elasticity of the number of Big Mac's that can be purchased by minimum wage workers is around 80% of the own-wage elasticity, meaning that even if workers spend all their money in Big Mac's, their real wage increases are still sizable. Renkin et al. (2022) also suggest that the price effects do not neutralize the redistributive potential of the minimum wage, arguing that: "the rise in grocery store prices following a $1 minimum wage increase reduces real income by about $19 a year for households earning less than $10,000 a year. (...). The price increases in grocery stores offset only a relatively small part of the gains of minimum wage hikes. Minimum wage policies thus remain a redistributive tool even after accounting for price effects in grocery stores." Based on these pieces of evidence, I conjecture that ignoring price effects is unlikely to dramatically affect the conclusions of the policy analysis and leave this extension to future research.

Productivity effects: The model assumes that labor productivity is independent from the minimum wage. This abstracts from recent literature that finds that minimum wages can increase both workers' (Coviello et al., 2021; Ruffini, 2021; Ku, 2022) and firms' (Riley and Bondibene, 2017; Mayneris et al., 2018) productivities. Potential mechanisms include efficiency wages (Shapiro and Stiglitz, 1984) and effects on investment in training (Acemoglu and Pischke, 1999). Harasztosi and Lindner (2019) argue that it is unlikely that productivity increases play a major role at the firm level as it would contradict the heterogeneous employment effects found between tradable and non-tradable sectors. If these effects are substantial, abstracting from these worker- and firm-specific increases in productivity after minimum wage hikes is likely to make the case for a positive minimum wage conservative. Note, however, that the model can accommodate aggregate increases in productivity through reallocation effects, as in Dustmann et al. (2022). Importantly, the main policy results depend on reduced-form profit elasticities that are robust to productivity increases.

B Proofs

Proof of Proposition I In the absence of taxes, there is no budget constraint and the social welfare function is given by

$$\begin{aligned} SW(\overline{w}) &= \left(L_I^l + L_I^h + K_I\right) \cdot G(0) + \alpha_l \cdot \int_0^{U^l} G(U^l - c) dF_l(c) \\ &+ \alpha_h \cdot \int_0^{U^h} G(U^h - c) dF_h(c) + K \cdot \int_{\psi^*}^{\overline{\psi}} G\left(\Pi(\psi) - \xi\right) dO(\psi). \end{aligned} \tag{B.I}$$

Replacing $L_I^l + L_I^h = 1 - L_A^l - L_A^h$, the total derivative with respect to the minimum wage is given by

$$\begin{aligned} \frac{dSW}{d\overline{w}} &= \left(\frac{dK_I}{d\overline{w}} - \frac{dL_A^l}{d\overline{w}} - \frac{dL_A^h}{d\overline{w}}\right) \cdot G(0) \\ &+ \alpha_l \cdot G(0) \cdot f_l(U^l) \cdot \frac{dU^l}{d\overline{w}} + \alpha_l \cdot \frac{dU^l}{d\overline{w}} \cdot \int_0^{U^l} G'(U^l - c) dF_l(c) \\ &+ \alpha_h \cdot G(0) \cdot f_h(U^h) \cdot \frac{dU^h}{d\overline{w}} + \alpha_h \cdot \frac{dU^h}{d\overline{w}} \cdot \int_0^{U^h} G'(U^h - c) dF_h(c) \\ &+ K \cdot \left(\int_{\psi^*}^{\overline{\psi}} G'(\Pi(\psi) - \xi) \frac{d\Pi(\psi)}{d\overline{w}} dO(\psi) - \frac{d\psi^*}{d\overline{w}} \cdot G(0) \cdot o(\psi^*)\right). \end{aligned} \tag{B.II}$$

Note that $dL_A^s/d\overline{w} = d(\alpha_s \cdot F_s(U^s))/d\overline{w} = \alpha_s \cdot f_s(U^s) \cdot (dU^s/d\overline{w})$, for $s \in \{l, h\}$, and that $dK_I/d\overline{w} = d(K \cdot O(\psi^*))/d\overline{w} = K \cdot o(\psi^*) \cdot (d\psi^*/d\overline{w})$. Then, (B.II) is reduced to

$$\begin{aligned} \frac{dSW}{d\overline{w}} &= \alpha_s \cdot \frac{dU^l}{d\overline{w}} \cdot \int_0^{U^l} G'(U^l - c) dF_l(c) + \alpha_h \cdot \frac{dU^h}{d\overline{w}} \cdot \int_0^{U^h} G'(U^h - c) dF_h(c) \\ &+ K \cdot \int_{\psi^*}^{\overline{\psi}} G'(\Pi(\psi) - \xi) \frac{d\Pi(\psi)}{d\overline{w}} dO(\psi). \end{aligned} \tag{B.III}$$

Using the marginal welfare weights definitions, (B.III) is reduced to

$$\frac{dSW}{d\overline{w}} = \gamma \cdot \left(\frac{dU^l}{d\overline{w}} \cdot L_A^l \cdot g_1^l + \frac{dU^h}{d\overline{w}} \cdot L_A^h \cdot g_1^h + K \cdot \int_{\psi^*}^{\overline{\psi}} g_\psi \frac{d\Pi(\psi)}{d\overline{w}} dO(\psi)\right). \quad \square \tag{B.IV}$$

Proof of Proposition II The Lagrangian is given by

$$\begin{aligned}\mathcal{L}(\overline{w}, y_0) &= \left(L_I^l + L_I^h + K_I\right) \cdot G(y_0) \\ &+\alpha_l \cdot \int_0^{U^l - y_0} G(U^l - c) dF_l(c) + \alpha_h \cdot \int_0^{U^h - y_0} G(U^h - c) dF_h(c) \\ &+K \cdot \int_{\psi^*}^{\overline{\psi}} G((1-t) \cdot \Pi(\psi, t) - \xi) dO(\psi) + \gamma \cdot \left[\int \left(E_m^l T(w_m^l) + E_m^h T(w_m^h)\right) dm \right.\\ &\left.+t \cdot K \cdot \int_{\psi^*}^{\overline{\psi}} \Pi(\psi, t) dO(\psi) - y_0 \left(L_I^l + L_I^h + K_I + \rho^l \cdot L_A^l + \rho^h \cdot L_A^h\right)\right], \end{aligned} \tag{B.V}$$

where γ is the budget constraint multiplier. Since $\rho^s \cdot L_A^s = L_A^s - \int E_m^s dm$, and using the fact that $L_I^l + L_I^h + L_A^l + L_A^h = 1$, the expression is simplified to

$$\begin{aligned}\mathcal{L}(\overline{w}, y_0) &= \left(L_I^l + L_I^h + K_I\right) \cdot G(y_0) \\ &+\alpha_l \cdot \int_0^{U^l - y_0} G(U^l - c) dF_l(c) + \alpha_h \cdot \int_0^{U^h - y_0} G(U^h - c) dF_h(c) \\ &+K \cdot \int_{\psi^*}^{\overline{\psi}} G((1-t) \cdot \Pi(\psi, t) - \xi) dO(\psi) + \gamma \cdot \left[\int \left(E_m^l (T(w_m^l) + y_0) \right.\right.\\ &\left.\left.+E_m^h (T(w_m^h) + y_0)\right) dm + t \cdot K \cdot \int_{\psi^*}^{\overline{\psi}} \Pi(\psi, t) dO(\psi) - y_0 \left(1 + K_I\right)\right]. \end{aligned} \tag{B.VI}$$

The derivative with respect to $\overline{w}$, taking y_0, t, and $T(\cdot)$ as given, is given by

$$\begin{aligned}\frac{d\mathcal{L}}{d\overline{w}} &= \left(\frac{dK_I}{d\overline{w}} - \frac{dL_A^s}{d\overline{w}} - \frac{dL_A^h}{\overline{w}}\right) \cdot G(y_0) \\ &+G(y_0) \cdot \alpha_l \cdot f_l(U^l - y_0) \cdot \frac{dU^l}{d\overline{w}} + \alpha_l \cdot \frac{dU^l}{d\overline{w}} \cdot \int_0^{U^l - y_0} G'(U^l - c) dF_l(c) \\ &+G(y_0) \cdot \alpha_h \cdot f_h(U^h - y_0) \cdot \frac{dU^h}{d\overline{w}} + \alpha_h \cdot \frac{dU^h}{d\overline{w}} \cdot \int_0^{U^h - y_0} G'(U^h - c) dF_h(c) \\ &+K \cdot \left[\int_{\psi^*}^{\overline{\psi}} G'((1-t) \cdot \Pi(\psi, t) - \xi)(1-t) \frac{d\Pi(\psi)}{d\overline{w}} dO(\psi) - G(y_0) \cdot o(\psi^*) \cdot \frac{d\psi^*}{d\overline{w}}\right] \\ &\gamma \cdot \left[\int \left(\frac{dE_m^l}{d\overline{w}} \left(T(w_m^l) + y_0\right) + E_m^l T'(w_m^l) \frac{dw_m^l}{d\overline{w}} \right.\right.\\ &\left.+\frac{dE_m^h}{d\overline{w}} \left(T(w_m^h) + y_0\right) + E_m^h T'(w_m^h) \frac{dw_m^h}{d\overline{w}}\right) dm \\ &\left.+t \cdot K \cdot \left(\int_{\psi^*}^{\overline{\psi}} \frac{d\Pi(\psi, t)}{d\overline{w}} dO(\psi) - \Pi(\psi^*, t) \cdot o(\psi^*) \cdot \frac{d\psi^*}{d\overline{w}}\right) - y_0 \cdot \frac{dK_I}{d\overline{w}}\right]. \end{aligned} \tag{B.VII}$$

Recall that $dK_I/d\overline{w} = K \cdot o(\psi^*) \cdot (d\psi^*/d\overline{w})$ and $dL_A^s/d\overline{w} = \alpha_s \cdot f_s(U^s - y_0) \cdot (dU^s/d\overline{w})$ for $s \in \{l, h\}$.

Using the social marginal weights definitions, and grouping common terms, (B.VII) can be written as

$$\begin{aligned}\frac{d\mathcal{L}}{d\overline{w}}\cdot\frac{1}{\gamma} \quad = \quad & \frac{dU^l}{d\overline{w}}\cdot L_A^l\cdot g_1^l+\frac{dU^h}{d\overline{w}}\cdot L_A^h\cdot g_1^h+K\cdot(1-t)\cdot\int_{\psi^*}^{\overline{\psi}}g_\psi\frac{d\Pi(\psi,t)}{d\overline{w}}dO(\psi)\\ & +\int\left(\frac{dE_m^l}{d\overline{w}}\left(T(w_m^l)+y_0\right)+E_m^lT'(w_m^l)\frac{dw_m^l}{d\overline{w}}\right)dm\\ & +\int\left(\frac{dE_m^h}{d\overline{w}}\left(T(w_m^h)+y_0\right)+E_m^hT'(w_m^h)\frac{dw_m^h}{d\overline{w}}\right)dm\\ & +t\cdot K\cdot\int_{\psi^*}^{\overline{\psi}}\frac{d\Pi(\psi,t)}{d\overline{w}}dO(\psi)-\frac{dK_I}{d\overline{w}}\cdot\left(t\cdot\Pi(\psi^*,t)+y_0\right). \quad \square \end{aligned} \tag{B.VIII}$$

Proof of Proposition III I assume that the planner optimizes allocations instead of taxes. Assuming either that $\max_i w_i^l < \min_j w_j^h$ or that the social planner can implement skill-specific income tax schedules, allow to set up the problem to be solved by pointwise maximization. That is, the planner chooses $\Delta y_m^s = y_m^s - y_0$, for all m and $s \in \{l, h\}$, and then recover taxes by noting that $T(w_m^s) + y_0 = w_m^s - \Delta y_m^s$. This implies that the Lagrangian is given by

$$\begin{aligned}\mathcal{L}\left(\overline{w}, \{\Delta y_m^s\}, y_0\right) \quad = \quad & \left(L_I^l + L_I^h + K_I\right) \cdot G(y_0) \\ & +\alpha_l \cdot \int_0^{U^l - y_0} G(U^l - c) dF_l(c) + \alpha_h \cdot \int_0^{U^h - y_0} G(U^h - c) dF_h(c) \\ & +K \cdot \int_{\psi^*}^{\overline{\psi}} G((1-t) \cdot \Pi(\psi, t) - \xi) dO(\psi) \\ & +\gamma \cdot \left[\int \left(E_m^l (w_m^l - \Delta y_m^l) + E_m^h (w_m^h - \Delta y_m^h)\right) dm \right. \\ & \left. +t \cdot K \cdot \int_{\psi^*}^{\overline{\psi}} \Pi(\psi, t) dO(\psi) - y_0 \left(1 + K_I\right)\right]. \end{aligned} \tag{B.IX}$$

There main differences with respect to the first-order condition derived for Proposition II is that the planner takes partial derivatives rather than total derivatives, leaving Δy_m^s constant, for all m and $s \in \{l, h\}$ when choosing $\overline{w}$. This implies that the optimality condition of the minimum wage can be written as

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \overline{w}} \cdot \frac{1}{\gamma} \quad = \quad & \frac{\partial U^l}{\partial \overline{w}} \cdot L_A^l \cdot g_1^l + \frac{\partial U^h}{\partial \overline{w}} \cdot L_A^h \cdot g_1^h + K \cdot (1-t) \cdot \int_{\psi^*}^{\overline{\psi}} g_\psi \frac{\partial \Pi(\psi, t)}{\partial \overline{w}} dO(\psi) \\ & + \int \left(\frac{\partial E_m^l}{\partial \overline{w}} \left(T(w_m^l) + y_0\right) + E_m^l \frac{\partial w_m^l}{\partial \overline{w}}\right) dm \\ & + \int \left(\frac{\partial E_m^h}{\partial \overline{w}} \left(T(w_m^h) + y_0\right) + E_m^h \frac{\partial w_m^h}{\partial \overline{w}}\right) dm \\ & +t \cdot K \cdot \int_{\psi^*}^{\overline{\psi}} \frac{\partial \Pi(\psi, t)}{\partial \overline{w}} dO(\psi) - \frac{\partial K_I}{\partial \overline{w}} \cdot \left(t \cdot \Pi(\psi^*, t) + y_0\right). \end{aligned} \tag{B.X}$$

Since Δy_m^s is fixed, we have that $\partial U^s / \partial \overline{w} = p_\theta \cdot (\partial \theta_m^s / \partial \overline{w}) \cdot \Delta y_m^s$, $\partial E_m^s / \partial \overline{w} = p_\theta \cdot (\partial \theta_m^s / \partial \overline{w}) \cdot L_A^s + p \cdot \alpha_s \cdot f_s(U^s - y_0) \cdot (\partial U^s / \partial \overline{w})$, profit effects are given by equations (A.XXIV) and (A.XXV), and wage spillovers are given by equation (A.XXI), except for firms that are constrained by the minimum wage for which $\partial w_m^l / \partial \overline{w} = 1$. Assume that $\overline{w} = \min_m w_m^l$. Then profit effects are all given by equation (A.XXV). This implies that all behavioral responses are mediated by $\partial \theta_m^s / \partial \overline{w}$. Since Δy_m^s is fixed, changes in the minimum wage do not (mechanically) change applicants, so the only possibility of $\partial \theta_m^s / \partial \overline{w} \neq 0$ is that vacancies change. However, if $\overline{w}$ is set at the market level, then the wage, albeit fixed, is on the first-order condition of the firm and, therefore, vacancies are unaffected because of envelope arguments. Then, when $\overline{w} = \min_m w_m^l \equiv w_i^l$, $\partial \mathcal{L} / \partial \overline{w} = E_i^l > 0$, so it is always optimal to increase the minimum wage above the market level. Further increase generate first-order distortions because profit reactions are given by equation (A.XXIV), so vacancy posting changes and can generate additional general equilibrium effects.

Also, the first order condition with respect to y_0 yields

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial y_0} = \left(L_I^l + L_I^h + K_I\right) \cdot G'(y_0) + \alpha_l \cdot \int_0^{U^l - y_0} G'(U^l - c) dF_l(c) \\ + \alpha_h \cdot \int_0^{U^l - y_0} G'(U^l - c) dF_l(c) - \gamma \cdot (1 + K_I) \quad &= \quad 0,\end{aligned} \tag{B.XI}$$

after noting that $\partial U^s / \partial y_0 = 1$ when Δy_m^s is fixed, so $\partial L_I^s / \partial y_0 = \partial K_I / \partial y_0 = 0$. This implies that $\omega_0 \cdot g_0 + \omega_1^l \cdot g_1^l + \omega_1^h \cdot g_1^h = 1$, where $\omega_0 = \left(L_I^l + L_I^h + K_I\right)/(1+K_I)$, $\omega_1^s = L_A^s/(1+K_I)$, and $\omega_0 + \omega_1^l + \omega_1^h = 1$.

Finally, after simplifying terms, the first order condition with respect to t yields

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial t} \cdot \frac{1}{\gamma} \quad = \quad & L_A^l \cdot g_1^l \cdot \frac{\partial U^l}{\partial t} + L_A^h \cdot g_1^h \cdot \frac{\partial U^h}{\partial t} + K \cdot \int_{\psi^*}^{\overline{\psi}} g_\psi \left[-\Pi(\psi, t) + (1-t) \cdot \frac{\partial \Pi(\psi, t)}{\partial t} \right] dO(\psi) \\ & + \int \left(\frac{\partial E_m^l}{\partial t} \left(w_m^l - \Delta y_m^l \right) + E_m^l \cdot \frac{\partial w_m^l}{\partial t} \right) dm + \int \left(\frac{\partial E_m^h}{\partial t} \left(w_m^h - \Delta y_m^h \right) + E_m^h \cdot \frac{\partial w_m^h}{\partial t} \right) dm \\ & + t \cdot K \cdot \int_{\psi^*}^{\overline{\psi}} \frac{\partial \Pi(\psi, t)}{\partial t} dO(\psi) + K \cdot \int_{\psi^*}^{\overline{\psi}} \Pi(\psi, t) dO(\psi) \\ & - t \cdot K \cdot \Pi(\psi^*, t) \cdot o(\psi^*) \cdot \frac{\partial \psi^*}{\partial t} - y_0 \cdot \frac{\partial K_I}{\partial t} = 0.\end{aligned} \tag{B.XII}$$

From equations (A.I) and (A.II), it follows that wages and vacancies (and, therefore, employment and profits) decrease with t, because $\phi_{tn} \leq 0$. Together with Δy_m^s being fixed, this implies that tightness decreases and, therefore, $\partial U^s / \partial t < 0$ for $s \in \{l, h\}$. Reordering terms yields

$$\begin{aligned}K \cdot \int_{\psi^*}^{\overline{\psi}} (1 - g_\psi) \Pi(\psi, t) dO(\psi) \quad = \quad & - \left(L_A^l \cdot g_1^l \cdot \frac{\partial U^l}{\partial t} + L_A^h \cdot g_1^h \cdot \frac{\partial U^h}{\partial t} \right) \\ & - K \cdot \int_{\psi^*}^{\overline{\psi}} \frac{\partial \Pi(\psi, t)}{\partial t} \left[g_\psi (1-t) + t \right] dO(\psi) \\ & - \int \left(\frac{\partial E_m^l}{\partial t} \left(w_m^l - \Delta y_m^l \right) + E_m^l \cdot \frac{\partial w_m^l}{\partial t} \right) dm \\ & - \int \left(\frac{\partial E_m^h}{\partial t} \left(w_m^h - \Delta y_m^h \right) + E_m^h \cdot \frac{\partial w_m^h}{\partial t} \right) dm \\ & + \frac{\partial K_I}{\partial t} \left(\Pi(\psi^*, t) + y_0 \right).\end{aligned} \tag{B.XIII}$$

The right-hand side is positive,[5] which implies that

$$\int_{\psi^*}^{\overline{\psi}} \omega_\psi (1 - g_\psi) dO(\psi) > 0, \tag{B.XIV}$$

with $\omega_\psi = \Pi(\psi, t) / \int_{\psi^*}^{\overline{\psi}} \Pi(\psi, t) dO(\psi)$, so the profit-weighted average welfare weight on active capitalists is smaller than one. $\square$

[5]Provided the optimal income tax system is not giving generalized tax-based employment subsidies that are large enough to make the optimal corporate tax rate increase to decrease employment to more than compensate all the negative welfare and fiscal externalities, which would imply that the income tax system would not be optimal in the first place.

Proof of Proposition IV Abstracting from the income tax system and the firm-level entry decisions implies that $T(w) = -y_0$, for all w, which is funded by the corporate tax revenue. Equation (B.X) implies that increasing the minimum wage when the corporate tax rate is optimal is welfare improving if

$$\frac{\partial U^l}{\partial \overline{w}} \cdot L_A^l \cdot g_1^l + K_S \cdot \frac{\partial \Pi^S}{\partial \overline{w}} \cdot \left(g_K^S + t \cdot (1 - g_K^S)\right) > 0. \tag{B.XV}$$

where I used that $\partial U^h/\partial \overline{w} = 0$ and $\partial \Pi^M/\partial \overline{w} = 0$ because high-skill workers work in firms non-affected by the minimum wage. I omit arguments of the profit functions to simplify notation.

With no income taxes, $U^l = p^l(\theta^l) \cdot \overline{w} + y_0$ and $U^h = p^h(\theta^h) \cdot w^h + y_0$. Then, it follows that $dU^l = p_\theta \cdot d\theta^l \cdot \overline{w} + p^l(\theta^l) \cdot d\overline{w}$ and $dU^h = p_\theta \cdot d\theta^h \cdot w^h + p^h(\theta^h) \cdot dw^h$, so

$$\frac{\partial U^l}{\partial t} = p_\theta \cdot \frac{\partial \theta^l}{\partial t} \cdot \overline{w}, \tag{B.XVI}$$

$$\frac{\partial U^h}{\partial t} = p_\theta \cdot \frac{\partial \theta^h}{\partial t} \cdot w^h + p^h(\theta^h) \cdot \frac{\partial w^h}{\partial t}, \tag{B.XVII}$$

$$\frac{\partial U^l}{\partial \overline{w}} = p_\theta \cdot \frac{\partial \theta^l}{\partial \overline{w}} \cdot \overline{w} + p^l(\theta^l), \tag{B.XVIII}$$

$$\frac{\partial U^h}{\partial \overline{w}} = 0. \tag{B.XIX}$$

Differentiating the first order conditions of the firms with respect to vacancies we have that

$$\left(\widetilde{\phi}_{nk}^S \cdot dk_D - d\overline{w}\right) q^l + \left(\widetilde{\phi}_n^S - \overline{w}\right) \cdot q_\theta \cdot \left(\frac{\partial \theta^l}{\partial w} \cdot d\overline{w} + \frac{\partial \theta^l}{\partial U^l} \cdot dU^l\right) = \eta_{vv}^l \cdot dv^l, \tag{B.XX}$$

$$\left(\widetilde{\phi}_{nk}^M \cdot dk_D - dw^h\right) q^h + \left(\widetilde{\phi}_n^M - w^h\right) \cdot q_\theta \cdot \left(\frac{\partial \theta^h}{\partial w} \cdot dw^h + \frac{\partial \theta^h}{\partial U^h} \cdot dU^h\right) = \eta_{vv}^h \cdot dv^h, \tag{B.XXI}$$

where, for analytical simplicity, I assumed that the differential effects driven by the curvature of $\widetilde{\phi}$ are second-order (i.e., $\widetilde{\phi}_{nn}$ is small relative to the first-order effects). Recall also from the first order condition of manufacturing firms with respect to wages that $\left(\widetilde{\phi}_n^M - w^h\right) \cdot q_\theta \cdot \frac{\partial \theta^h}{\partial w^h} = q^h$. Then, (B.XXI) simplifies to

$$\widetilde{\phi}_{nk}^M \cdot dk_D \cdot q^h + \left(\widetilde{\phi}_n^M - w^h\right) \cdot q_\theta \cdot \frac{\partial \theta^h}{\partial U^h} \cdot dU^h = \eta_{vv}^h \cdot dv^h. \tag{B.XXII}$$

Finally, we also know that $\theta^l = K_S \cdot v^l / L_A^l$ and $\theta^h = K_M \cdot v^h / L_A^h$. Then:

$$d\theta^l = \frac{K_S}{L_A^l} \cdot dv^l - \frac{K_S \cdot v^l}{L_A^{l2}} \cdot f_l(U^l - y_0) \cdot dU^l, \tag{B.XXIII}$$

$$d\theta^h = \frac{K_M}{L_A^h} \cdot dv^h - \frac{K_M \cdot v^h}{L_A^{h2}} \cdot f_h(U^h - y_0) \cdot dU^h. \tag{B.XXIV}$$

First, consider the comparative statics with respect to t, i.e., equations (B.XVI) and (B.XVII). Setting

$d\overline{w} = 0$ and combining equations (B.XVI), (B.XX), and (B.XXIII) yields

$$\begin{aligned}
&\frac{\widetilde{\phi}_{nk}^{S} \cdot q^{l}}{\eta_{vv}^{l}} \cdot \frac{\partial k_D}{\partial t} + \left(\widetilde{\phi}_{n}^{S} - \overline{w}\right) \cdot \frac{q_\theta}{\eta_{vv}^{l}} \cdot \frac{\partial \theta^{l}}{\partial U^{l}} \cdot p_\theta \cdot \overline{w} \cdot \frac{\partial \theta^{l}}{\partial t} \\
&= \frac{\partial \theta^{l}}{\partial t} \cdot \frac{L_A^{l}}{K_S} \cdot \left(1 + \frac{K_S \cdot v^{l}}{L_A^{l2}} \cdot f_l(U_l - y_0) \cdot p_\theta \cdot \overline{w}\right),
\end{aligned} \tag{B.XXV}$$

which implies that $\partial\theta^l/\partial t = -a_t^l \cdot \varepsilon_{k,t}^S$, with $a_t^l > 0$ provided $\widetilde{\phi}_{nk}^S > 0$. Then, $\partial U^l/\partial t = -a_t^l \cdot p_\theta \cdot \overline{w} \cdot \varepsilon_{k,t}^S$. Then, $\left(\partial U^l/\partial t\right)/\partial\varepsilon_{k,t}^S < 0$.

Also, from equation (A.III), we know that $dw/dv > 0$ for firms that are not constrained by the minimum wage. Then, combining equations (B.XVII), (B.XXII), and (B.XXIV) yields

$$\begin{aligned}
&\frac{\widetilde{\phi}_{nk}^{M} \cdot q^{h}}{\eta_{vv}^{h}} \cdot \frac{\partial k_D}{\partial t} + \left(\widetilde{\phi}_{n}^{M} - w^{h}\right) \cdot \frac{q_\theta}{\eta_{vv}^{h}} \cdot \frac{\partial \theta^{h}}{\partial U^{h}} \cdot p_\theta \cdot w^{h} \cdot \frac{\partial \theta^{h}}{\partial t} \\
&= \frac{\partial \theta^{h}}{\partial t} \cdot \frac{L_A^{h}}{K_M} \cdot \left(1 + \frac{K_M \cdot v^{h}}{L_A^{h2}} \cdot f_h(U_h - y_0) \cdot p_\theta \cdot w^{h}\right) \cdot \frac{\left(1 - \left(\widetilde{\phi}_{n}^{M} - w^{h}\right) \cdot \frac{q_\theta}{\eta_{vv}^{h}} \cdot \frac{\partial \theta^{h}}{\partial U^{h}} \cdot p^{h}(\theta^{h}) \cdot \frac{\partial w^{h}}{\partial v^{h}}\right)}{\left(1 - \frac{v^{h}}{L_A^{h}} \cdot f_h(U^{h} - y_0) \cdot \frac{\partial w}{\partial v}\right)}
\end{aligned} \tag{B.XXVI}$$

The only term that has ambiguous sign is the last denominator. I assume it is positive, which economically implies that the increase in the corporate tax rate generates a decrease in posted vacancies that is attenuated by a change in the posted wage and is expected to happen when the density is negligible.[6] Then, this expression implies that $\partial\theta^h/\partial t = -a_t^h \cdot \varepsilon_{k,t}^M$, with $a_t^h > 0$ provided $\widetilde{\phi}_{nk}^M > 0$. Then,

$$\frac{\partial U^{h}}{\partial t} \quad = \quad -a_t^{h} \cdot \varepsilon_{k,t}^{M} \cdot \left(p_\theta \cdot w^{h} + p^{h}(\theta^{h}) \cdot \frac{\partial w^{h}}{\partial v^{h}} \cdot b\right), \tag{B.XXVII}$$

where $b = \frac{L_A^h}{K_M} \cdot \left(1 + \frac{K_M \cdot v^h}{L_A^{h2}} \cdot f_h(U_h - y_0) \cdot p_\theta \cdot w^h\right) \cdot \left(1 - \frac{v^h}{L_A^h} \cdot f_h(U^h - y_0) \cdot \frac{\partial w}{\partial v}\right) > 0$ under the assumption used above. Then, $\left(\partial U^h/\partial t\right)/\partial\varepsilon_{k,t}^M < 0$.

Finally, allowing $d\overline{w}$ to be non-zero, and combining equations (B.XVIII), (B.XX), and (B.XXIII) yields

$$\begin{aligned}
&\widetilde{\phi}_{nk}^{S} \cdot \frac{\partial k_D}{\partial \overline{w}} \cdot q^{l} - q^{l} + \left(\widetilde{\phi}_{n}^{S} - \overline{w}\right) \cdot q_\theta \cdot \frac{\partial \theta^{l}}{\partial w} + \left(\widetilde{\phi}_{n}^{S} - \overline{w}\right) \cdot q_\theta \cdot \frac{\partial \theta^{l}}{\partial U^{l}} \left(p_\theta \cdot \frac{\partial \theta^{l}}{\partial \overline{w}} \cdot \overline{w} + p^{l}(\theta^{l})\right) \\
&= \eta_{vv}^{l} \cdot \left(\frac{\partial \theta^{l}}{\partial \overline{w}} \cdot \frac{L_A^{l}}{K_S} \cdot \left(1 + \frac{K_S \cdot v^{l}}{L_A^{l2}} \cdot f_l(U^{l} - y_0) \cdot p_\theta \cdot \overline{w}\right) + \frac{v^{l}}{L_A^{l}} \cdot f_l(U^{l} - y_0) \cdot p^{l}(\theta^{l})\right).
\end{aligned} \tag{B.XXVIII}$$

Noting that $-q^l + \left(\widetilde{\phi}_n^S - \overline{w}\right) \cdot q_\theta \cdot \frac{\partial\theta^l}{\partial w} \leq 0$ because the firm is possibly deviating from the first order condition, then we have that $\partial\theta^l/\partial\overline{w} = -a_{\overline{w}}^l \cdot \varepsilon_{k,t}^S - b_{\overline{w}}^l$, with $b_{\overline{w}}^l > 0$ and $a_{\overline{w}}^l > 0$ provided $\widetilde{\phi}_{nk}^S > 0$. Then $\partial U^l/\partial\overline{w} = -\left(a_{\overline{w}}^l \cdot \varepsilon_{k,t}^S + b_{\overline{w}}^l\right) \cdot p_\theta \cdot \overline{w} + p^l(\theta)^l$. While the sign of $\partial U^l/\partial\overline{w}$ is ambiguous, it follows that

[6] Assuming the contrary would imply that when capital mobility is larger the distortion of the corporate tax rate is smaller because of a huge labor participation effect if the density is large at $U^h - y_0$.

$\left(\partial U^l/\partial \overline{w}\right)/\partial \varepsilon_{k,t}^S < 0$.

Now, consider the first order condition of the planner with respect to the corporate tax rate:

$$\begin{aligned}
\frac{\partial U^l}{\partial t}\cdot L_A^l \cdot g_1^l + \frac{\partial U^h}{\partial t}\cdot L_A^h \cdot g_1^h & \\
+K_S \cdot g_K^S \cdot \left(-\Pi^S + (1-t)\cdot \frac{\partial \Pi^S}{\partial t}\right) + K_M \cdot g_K^M \cdot \left(-\Pi^M + (1-t)\cdot \frac{\partial \Pi^M}{\partial t}\right) & \\
+K_S \cdot \Pi^S + K_M \cdot \Pi^M + K_S \cdot t \cdot \frac{\partial \Pi^S}{\partial t} + K_M \cdot t \cdot \frac{\partial \Pi^M}{\partial t} &= 0. \quad \text{(B.XXIX)}
\end{aligned}$$

Grouping terms yields

$$K_S \cdot \left(g_K^S + t\cdot(1-g_K^S)\right) = \frac{A + B - \left(g_K^M + t\cdot(1-g_K^M)\right)\cdot \gamma_t^M \cdot \varepsilon_{k,t}^M}{\gamma_t^S \cdot \varepsilon_{k,t}^S}, \quad \text{(B.XXX)}$$

where $A = K_S \cdot \Pi^S \cdot (1-g_K^S) + K_M \cdot \Pi^M \cdot (1-g_K^M)$ and $B = (\partial U^l/\partial t)\cdot L_A^l \cdot g_1^l + (\partial U^h/\partial t)\cdot L_A^h \cdot g_1^h$. Replacing (B.XXX) in (B.XV) gives

$$\frac{\partial U^l}{\partial \overline{w}}\cdot L_A^l \cdot g_1^l + \left(-\gamma_{\overline{w}}^S \cdot \varepsilon_{k,t}^S + \widetilde{\Pi}_{\overline{w}}^S\right)\cdot \frac{A + B - \left(g_K^M + t\cdot(1-g_K^M)\right)\cdot \gamma_t^M \cdot \varepsilon_{k,t}^M}{\gamma_t^S \cdot \varepsilon_{k,t}^S} > 0. \quad \text{(B.XXXI)}$$

Name the LHS $\mathcal{F}(\varepsilon_{k,t}^S, \varepsilon_{k,t}^M)$, so increasing the minimum wage is desirable if $\mathcal{F}(\varepsilon_{k,t}^S, \varepsilon_{k,t}^M) > 0$. Assuming the welfare weights are fixed (or that the effect of capital mobility on them are of second-order), note that:

$$\frac{\partial \mathcal{F}(\varepsilon_{k,t}^S, \varepsilon_{k,t}^M)}{\partial \varepsilon_{k,t}^S} = \frac{\partial\left(\partial U^l/\partial \overline{w}\right)}{\partial \varepsilon_{k,t}^S}\cdot L_A^l \cdot g_1^l - \frac{\gamma_{\overline{w}}^S}{\gamma_t^S}\cdot \frac{\partial B}{\partial \varepsilon_{k,t}^S} + \widetilde{\Pi}_{\overline{w}}^S \cdot \frac{\left(\frac{\partial B}{\partial \varepsilon_{k,t}^S} - C\cdot \gamma_t^S\right)}{\gamma_t^S \cdot \varepsilon_{k,t}^S}, \quad \text{(B.XXXII)}$$

where $C = \left(A + B - \left(g_K^M + t\cdot(1-g_K^M)\right)\cdot \gamma_t^M \cdot \varepsilon_{k,t}^M\right)/\gamma_t^S \cdot \varepsilon_{k,t}^S > 0$ following (B.XXX), provided the optimal corporate tax rate has an interior solution. Then, the sign of $\partial \mathcal{F}(\varepsilon_{k,t}^S, \varepsilon_{k,t}^M)/\partial \varepsilon_{k,t}^S$ is ambiguous: the first term is negative and the second and third a positive. On the other hand:

$$\frac{\partial \mathcal{F}(\varepsilon_{k,t}^S, \varepsilon_{k,t}^M)}{\partial \varepsilon_{k,t}^M} = \frac{\left(-\gamma_{\overline{w}}^S \cdot \varepsilon_{k,t}^S + \widetilde{\Pi}_{\overline{w}}^S\right)}{\gamma_t^S \cdot \varepsilon_{k,t}^S}\cdot \left(\frac{\partial B}{\partial \varepsilon_{k,t}^M} - \left(g_K^M + t\cdot(1-g_K^M)\right)\cdot \gamma_t^M\right) > 0, \quad \text{(B.XXXIII)}$$

which is unambiguously positive. □

Proof of Proposition V The objective function of the planner is given by

$$\begin{aligned}\mathcal{L}(\overline{w},\Delta y^l,\Delta y^h) &= (L_I^l+L_I^h)\cdot G(y_0)+\alpha_l\cdot\int_0^{U^l-y_0}G(U^l-c)dF_l(c)+\alpha_h\cdot\int_0^{U^h-y_0}G(U^h-c)dF_h(c)\\ &\quad +K_S\cdot G\left((1-t)\cdot\Pi^S\right)+K_M\cdot G\left((1-t)\cdot\Pi^M\right)\\ &\quad +\gamma\cdot\Big[E^l\cdot(\overline{w}-\Delta y^l)+E^h\cdot(w^h-\Delta y^h)-y_0+t\cdot\left(K_S\cdot\Pi^S+K_M\cdot\Pi^M\right)\Big],\end{aligned} \tag{B.XXXIV}$$

where $U^s=p^s(\theta^s)\cdot\Delta y^s-y_0$ and $E^s=p^s(\theta^s)\cdot L_A^s$, for $s\in\{l,h\}$. The first-order condition with respect to the minimum wage, after cancelling terms, is given by

$$\frac{\partial\mathcal{L}}{\partial\overline{w}}\cdot\frac{1}{\gamma} = \frac{\partial U^l}{\partial\overline{w}}\cdot L_A^l\cdot g_1^l+\frac{\partial E^l}{\partial\overline{w}}\cdot(\overline{w}-\Delta y^l)+E^l+K_S\cdot\frac{\partial\Pi^S}{\partial\overline{w}}\cdot\left((1-t)\cdot g_K^S+t\right). \tag{B.XXXV}$$

Since after-tax allocations are fixed by the optimal tax system, we have that

$$\frac{\partial U^l}{\partial\overline{w}} = p_\theta\cdot\frac{\partial\theta^l}{\partial\overline{w}}\cdot\Delta y^l, \tag{B.XXXVI}$$

$$\frac{\partial E^l}{\partial\overline{w}} = p_\theta\cdot\frac{\partial\theta^l}{\partial\overline{w}}\cdot L_A^l+p^l(\theta^l)\cdot\frac{\partial L_A^l}{\partial\overline{w}}=p_\theta\cdot\frac{\partial\theta^l}{\partial\overline{w}}\cdot L_A^l+\frac{p^l(\theta^l)\cdot L_A^l}{\overline{w}}\cdot\varepsilon_{L,\overline{w}}^l, \tag{B.XXXVII}$$

where $\varepsilon_{L,\overline{w}}^l$ is the participation elasticity with respect to changes in the minimum wage fixing after-tax allocations, given by $(\partial L_A^l/\partial\overline{w})/(\overline{w}/L_A^l)$. Since after-tax allocations are fixed, $\text{sgn}\left(\varepsilon_{L,\overline{w}}^l\right)=\text{sgn}\left(\partial U^l/\partial\overline{w}\right)=\text{sgn}\left(\partial\theta^l/\partial\overline{w}\right)$. Since changes in the minimum wage have no direct effect on after-tax allocations, changes in participation will only be driven by responses in vacancies which in turn affect tightness and employment probabilities.

Since $\theta^l=(K_S\cdot v^l)/L_A^l$, we have that

$$\frac{\partial\theta^l}{\partial\overline{w}}=\left(\frac{K_S}{L_A^l}\cdot\frac{\partial v^l}{\partial\overline{w}}-\frac{K_S\cdot v^l}{L_A^{l2}}\cdot\frac{\partial L_A^l}{\partial\overline{w}}\right)=\left(\frac{K_S}{L_A^l}\cdot\frac{\partial v^l}{\partial\overline{w}}-\frac{K_S\cdot v^l}{L_A^l\cdot\overline{w}}\cdot\varepsilon_{L,\overline{w}}^l\right). \tag{B.XXXVIII}$$

Differentiating the first-order condition for vacancies in firms paying the minimum wage yields

$$-q^l+\left(\phi_n^S-\overline{w}\right)\cdot q_\theta\cdot\left(\frac{\partial\theta^l}{\partial w}+\frac{\partial\theta^l}{\partial U^l}\cdot\frac{\partial U^l}{\partial\overline{w}}\right)=\eta_{vv}^l\cdot\frac{\partial v^l}{\partial\overline{w}}. \tag{B.XXXIX}$$

If the minimum wage is set at the market wage, then the firm is also in the first order condition for wages. That implies that $q_l=\left(\phi_n^S-\overline{w}\right)\cdot q_\theta\cdot(\partial\theta^l/\partial w)$. Replacing this and the expression for $\partial U^l/\partial\overline{w}$ in the previous expression yields

$$\left(\phi_n^S-\overline{w}\right)\cdot q_\theta\cdot\frac{\partial\theta^l}{\partial\overline{w}}\cdot\Delta y^l=\eta_{vv}^l\cdot\frac{\partial v^l}{\partial\overline{w}}. \tag{B.XL}$$

Then, replacing in (B.XXXVIII), we conclude that

$$\frac{\partial\theta^l}{\partial\overline{w}} \cdot \left(1 - \frac{K_S}{L_A^l \cdot \eta_{vv}^l} \cdot \left(\phi_n^S - \overline{w}\right) \cdot q_\theta \cdot \Delta y^l\right) = -\frac{K_S \cdot v^l}{L_A^l \cdot \overline{w}} \cdot \varepsilon_{L,\overline{w}}^l. \tag{B.XLI}$$

Then, $\operatorname{sgn}\left(\varepsilon_{L,\overline{w}}^l\right) = -\operatorname{sgn}\left(\partial\theta^l/\partial\overline{w}\right)$. Together with the previous conclusion that $\operatorname{sgn}\left(\varepsilon_{L,\overline{w}}^l\right) = \operatorname{sgn}\left(\partial\theta^l/\partial\overline{w}\right)$, I conclude that when $\overline{w}$ is equal to the market wage, then $\partial\theta^l/\partial\overline{w} = \varepsilon_{L,\overline{w}}^l = 0$. Intuitively, the minimum wage generates no distortion in vacancy postings because the firm is optimizing, and also generates no distortion in labor supply since after-tax allocations are fixed. This argument no longer holds when the minimum wage is larger than the market wage, since in that case $q_l > (\phi_l - \overline{w}) \cdot q_\theta \cdot (\partial\theta^l/\partial w)$, generating therefore a negative effect on vacancies. Note that when $\overline{w}$ is set at the market-wage, the envelope theorem implies that $\partial\Pi^S/\partial\overline{w} = q_\theta \cdot \left(\partial\theta^l/\partial\overline{w}\right) \cdot v^l \cdot \left(\phi_n^S - \overline{w}\right)$. So $\partial\theta^l/\partial\overline{w} = 0$ implies $\partial\Pi^S/\partial\overline{w} = 0$. Replacing this in equation (B.XXXV) we conclude that, when the minimum wage equals the market wage, $\partial\mathcal{L}/\partial\overline{w} = \gamma \cdot E^l > 0$, so increasing the minimum wage above the market level is welfare improving when the tax system is optimal.

Regarding low-skill workers after-tax allocations, the first order condition is given by

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial\Delta y^l} \cdot \frac{1}{\gamma} &= \frac{\partial U^l}{\partial\Delta y^l} \cdot L_A^l \cdot g_1^l + \frac{\partial E^l}{\partial\Delta y^l} \cdot \left(\overline{w} - \Delta y^l\right) - E^l \\ &\quad + K_S \cdot \frac{\partial\Pi^S}{\partial\Delta y^l} \cdot \left((1-t) \cdot g_K^S + t\right). \end{aligned} \tag{B.XLII}$$

Holding fixed the minimum wage, we have that

$$\frac{\partial U^l}{\partial\Delta y^l} = p_\theta \cdot \frac{\partial\theta^l}{\partial\Delta y^l} \cdot \Delta y^l + p^l(\theta^l), \tag{B.XLIII}$$

$$\frac{\partial E^l}{\partial\Delta y^l} = p_\theta \cdot \frac{\partial\theta^l}{\partial\Delta y^l} \cdot L_A^l + p^l(\theta^l) \cdot \frac{\partial L_A^l}{\partial\Delta y^l} = p_\theta \cdot \frac{\partial\theta^l}{\partial\Delta y^l} \cdot L_A^l + \frac{p^l(\theta^l) \cdot L_A^l}{\Delta y^l} \cdot \varepsilon_{L,\Delta}^l, \tag{B.XLIV}$$

where $\varepsilon_{L,\Delta}^l$ is the participation elasticity with respect to changes in the after-tax allocations holding the minimum wage fixed, given by $(\partial L_A^l/\partial \overline{w})/(\overline{w}/L_A^l)$. The first order condition can be written as

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \Delta y^l}\cdot\frac{1}{\gamma} &= \left(p_\theta\cdot\frac{\partial\theta^l}{\partial\Delta y^l}\cdot\Delta y^l+p^l(\theta^l)\right)\cdot L_A^l\cdot g_1^l+\left(p_\theta\cdot\frac{\partial\theta^l}{\partial\Delta y^l}\cdot L_A^l+\frac{p^l(\theta^l)\cdot L_A^l}{\Delta y^l}\cdot\varepsilon_{L,\Delta}^l\right)\cdot(\overline{w}-\Delta y^l)\\
&\quad -E^l+K_S\cdot q_\theta\cdot\frac{\partial\theta^l}{\partial\Delta y^l}\cdot v^l\cdot\left(\phi_n^S-\overline{w}\right)\cdot\left((1-t)\cdot g_K^S+t\right),\\
&= p_\theta\cdot\frac{\partial\theta^l}{\partial\Delta y^l}\cdot L_A^l\cdot\left(\Delta y^l\cdot g_1+\overline{w}-\Delta y^l\right)+E^l\cdot\left(g_1+\varepsilon_{L,\Delta}^l\cdot\frac{\overline{w}-\Delta y^l}{\Delta y^l}-1\right)\\
&\quad +K_S\cdot q_\theta\cdot\frac{\partial\theta^l}{\partial\Delta y^l}\cdot v^l\cdot\left(\phi_n^S-\overline{w}\right)\cdot\left((1-t)\cdot g_K^S+t\right),\\
&= -\frac{p_\theta\cdot\theta^l\cdot L_A^l\cdot\varepsilon_{\theta,\Delta}^l}{\Delta y^l}\cdot\left(\Delta y^l\cdot g_1+\overline{w}-\Delta y^l\right)+E^l\cdot\left(g_1+\varepsilon_{L,\Delta}^l\cdot\frac{\overline{w}-\Delta y^l}{\Delta y^l}-1\right)\\
&\quad -\frac{K_S\cdot v^l\cdot q_\theta\cdot\theta^l\cdot\left(\phi_n^S-\overline{w}\right)\cdot\varepsilon_{\theta,\Delta}^l}{\Delta y^l}\cdot\left((1-t)\cdot g_K^S+t\right),
\end{aligned}\tag{B.XLV}$$

where $\varepsilon_{\theta,\Delta}^l=-\left(\partial\theta^l/\partial\Delta y^l\right)/\left(\Delta y^l/\theta^l\right)>0$ is the elasticity of tightness to changes in after-tax allocations holding the minimum wage fixed, with $\partial\theta^l/\partial\Delta y^l<0$ as shown below.

Noting that $\Delta y^l=(1-\tau_l)\cdot\overline{w}$ and $\overline{w}-\Delta y^l=\tau_l\cdot\overline{w}$ implies that

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \Delta y^l}\cdot\frac{1}{\gamma} &= -p_\theta\cdot\theta^l\cdot L_A^l\cdot\varepsilon_{\theta,\Delta}^l\cdot\left(g_1+\frac{\tau_l}{1-\tau_l}\right)+E^l\cdot\left(g_1+\varepsilon_{L,\Delta}^l\cdot\frac{\tau_l}{1-\tau_l}-1\right)\\
&\quad -\frac{K_S\cdot v^l\cdot q_\theta\cdot\theta^l\cdot\left(\phi_n^S-\overline{w}\right)\cdot\varepsilon_{\theta,\Delta}^l}{(1-\tau_l)\cdot\overline{w}}\cdot\left((1-t)\cdot g_K^S+t\right).
\end{aligned}\tag{B.XLVI}$$

To see whether a negative τ_l is optimal when $\overline{w}$ is fixed at its optimal value, I evaluate equation (B.XLVI) at $\tau_l=0$. Also, since the matching function has constant returns to scale, it follows from Euler's theorem that $p^l(\theta^l)>\theta^l\cdot p_\theta$. Then, at $\tau_l=0$,

$$\frac{\partial \mathcal{L}}{\partial \Delta y^l}>E^l\cdot\gamma\cdot\left[-g_1\cdot\varepsilon_{\theta,\Delta}^l+g_1-1+c\cdot\varepsilon_{\theta,\Delta}^l\cdot\left((1-t)\cdot g_K^S+t\right)\right],\tag{B.XLVII}$$

where $c=-\left(q_\theta\cdot\theta^l\cdot\left(\phi_n^S-\overline{w}\right)\right)/\left(\overline{w}\cdot q(\theta^l)\right)\in(0,1)$ since $q(\theta^l)>-q_\theta\cdot\theta$ and $\overline{w}>\phi_n^S-\overline{w}$. Then, a sufficient condition for having negative marginal tax rates for low-skill workers (i.e., increasing Δy^l is welfare improving when $\tau_l=0$) is given by

$$-g_1\cdot\varepsilon_{\theta,\Delta}^l+g_1-1+c\cdot\varepsilon_{\theta,\Delta}^l\cdot\left((1-t)\cdot g_K^S+t\right)\quad>\quad 0,\tag{B.XLVIII}$$

which happens when $g_1>\left(1-c\cdot\varepsilon_{\theta,\Delta}^l\cdot\left((1-t)\cdot g_K^S+t\right)\right)/\left(1-\varepsilon_{\theta,\Delta}^l\right)$, where I assumed $\varepsilon_{\theta,\Delta}^l<1$.

To show that $\partial\theta^l/\partial\Delta y^l<0$, note that $\theta^l=(K_S\cdot v^l)/L_A^l$, so we have that

$$\frac{\partial\theta^l}{\partial\Delta y^l}=\left(\frac{K_S}{L_A^l}\cdot\frac{\partial v^l}{\partial\Delta y^l}-\frac{K_S\cdot v^l}{L_A^{l2}}\cdot\frac{\partial L_A^l}{\partial\Delta y^l}\right)=\left(\frac{K_S}{L_A^l}\cdot\frac{\partial v^l}{\partial\Delta y^l}-\frac{K_S\cdot v^l}{L_A^l\cdot\Delta y^l}\cdot\varepsilon_{L,\Delta}^l\right).\tag{B.XLIX}$$

Differentiating the first-order condition for vacancies in firms paying the minimum wage yields

$$\left(\phi_n^S - \overline{w}\right) \cdot q_\theta \cdot \left(\frac{\partial \theta^l}{\partial w} \cdot \frac{\partial \overline{w}}{\partial \Delta y^l} + \frac{\partial \theta^l}{\partial U^l} \cdot \frac{\partial U^l}{\partial \Delta y^l}\right) = \eta_{vv}^l \cdot \frac{\partial v^l}{\partial \Delta y^l}. \tag{B.L}$$

Since the planner holds fixed $\overline{w}$ when varying Δy^l, the previous expression can be simplified to

$$\frac{\left(\phi_n^S - \overline{w}\right) \cdot q_\theta}{\eta_{vv}^l} \cdot \left(\frac{\partial \theta^l}{\partial \Delta y^l} \cdot \Delta y^l + \frac{p^l(\theta^l)}{p_\theta}\right) = \frac{\partial v^l}{\partial \Delta y^l}. \tag{B.LI}$$

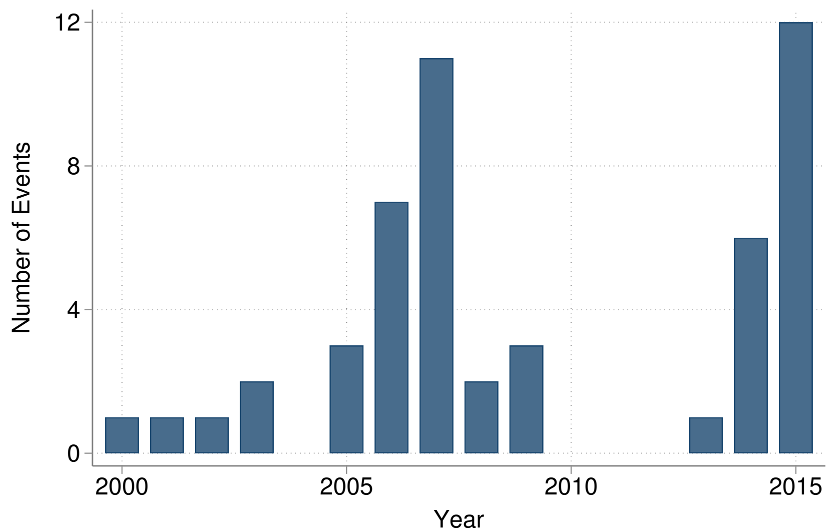
Replacing in equation (B.XLIX) yields

$$\begin{aligned} \frac{\partial \theta^l}{\partial \Delta y^l} \cdot \left(1 - \frac{\left(\phi_n^S - \overline{w}\right) \cdot q_\theta \cdot \Delta y^l \cdot K_S}{\eta_{vv}^l \cdot L_A^l}\right) &= \frac{\left(\phi_n^S - \overline{w})\right) \cdot q_\theta \cdot p^l(\theta^l) \cdot K_S}{\eta_{vv}^l \cdot L_A^l \cdot p_\theta} - \frac{\theta^l}{\Delta y^l} \cdot \varepsilon_{L,\Delta}^l, \\ \iff \frac{\partial \theta^l}{\partial \Delta y^l} \cdot \left(1 - \frac{\left(\phi_n^S - \overline{w}\right) \cdot q_\theta \cdot \Delta y^l \cdot \theta^l}{\eta_{vv}^l \cdot v^l}\right) &= \frac{\theta^l}{\Delta y^l} \cdot \left(\frac{\left(\phi_n^S - \overline{w}\right) \cdot q_\theta \cdot p^l(\theta^l) \cdot \Delta y^l}{\eta_{vv}^l \cdot v^l \cdot p_\theta} - \varepsilon_{L,\Delta}^l\right), \end{aligned} \tag{B.LII}$$

which implies that $\partial \theta^l / \partial \Delta y^l < 0$ provided $\varepsilon_{L,\Delta}^l > 0$. $\square$

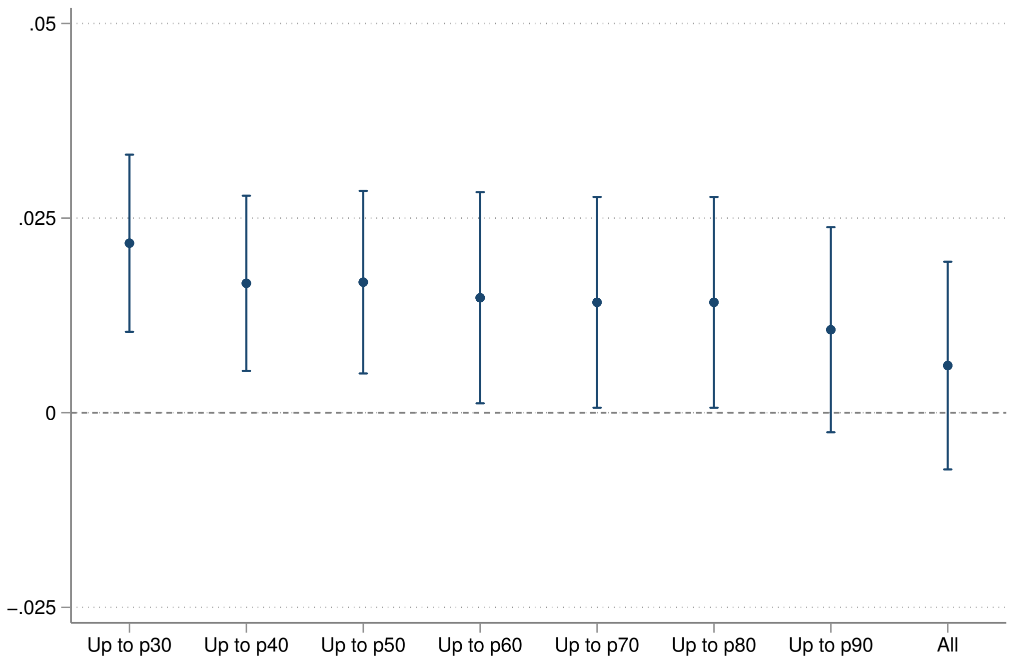
C Additional figures and tables

Figure C.I: State-level events by year



Notes: This figure plots the annual frequency of state-level minimum wage increases classified as events following Cengiz et al. (2019, 2022). Data on minimum wages is taken from Vaghul and Zipperer (2016). A state-level hourly minimum wage increase above the federal level is classified as an event if the increase is of at least \$0.25 (in 2016 dollars) in a state with at least 2% of the working population affected, where the affected population is computed using the NBER Merged Outgoing Rotation Group of the CPS, treated states do not experience other events in the three years previous to the event, and the event-timing allows to observe the outcomes from three years before to four years after.

Figure C.II: Minimum wage effects on low-skill workers' welfare: change in percentile considered



Notes: This figure plots the estimated β coefficient with its corresponding 95% confidence intervals from equation (18) using the average pre-tax wage of active low-skill workers including the unemployed as dependent variable. Each coefficient comes from a different regression where the sufficient statistic is computed using different percentiles to truncate the sample of employed low-skill workers when computing the average wage. Low-skill workers are defined as not having a college degree. All regressions include year-by-event fixed effects. Standard errors are clustered at the state level, and regressions are weighted by state-by-year average population.

Table C.I: List of Events

State	Events (year)	Total	State	Events (year)	Total
Alabama	-	0	Montana	2007	1
Alaska	2003, 2015	2	Nebraska	2015	1
Arizona	2007	1	Nevada	2006	1
Arkansas	2006, 2015	2	New Hampshire	-	0
California	2007, 2014	2	New Jersey	2006, 2014	2
Colorado	2007, 2015	2	New Mexico	2008	1
Connecticut	2009, 2015	2	New York	2005, 2013	2
Delaware	2000, 2007, 2014	3	North Carolina	2007	1
District of Columbia	2014	1	North Dakota	-	0
Florida	2005, 2009	2	Ohio	2007	1
Georgia	-	0	Oklahoma	-	0
Hawaii	2002, 2015	2	Oregon	2003	1
Idaho	-	0	Pennsylvania	2007	1
Illinois	2005	1	Rhode Island	2006, 2015	2
Indiana	-	0	South Carolina	-	0
Iowa	2008	1	South Dakota	2015	1
Kansas	-	0	Tennessee	-	0
Kentucky	-	0	Texas	-	0
Louisiana	-	0	Utah	-	0
Maine	-	0	Vermont	2009, 2015	2
Maryland	2015	1	Virginia	-	0
Massachusetts	2001, 2007, 2015	3	Washington	2007	1
Michigan	2006, 2014	2	West Virginia	2006, 2015	2
Minnesota	2014	1	Wisconsin	2006	1
Mississippi	-	0	Wyoming	-	0
Missouri	2007	1			

Notes: This table details the list of events considered in the event-studies. Data on minimum wages is taken from Vaghul and Zipperer (2016). A state-level hourly minimum wage increase above the federal level is classified as an event if the increase is of at least \$0.25 (in 2016 dollars) in a state with at least 2% of the working population affected, where the affected population is computed using the NBER Merged Outgoing Rotation Group of the CPS, treated states do not experience other events in the three years previous to the event, and the event-timing allows to observe the outcomes from three years before to four years after.

D Simulation appendix

Functional forms To simulate the model, I impose the following structure. Matching functions are given by $\mathcal{M}^s(L^s, V^s) = \delta_{0s} L^{s\delta_{1s}} V^{s1-\delta_{1s}}$, for $s \in \{l, h\}$. Revenue functions are given by $\widetilde{\phi}^I(k,n) = \psi^I \cdot k^{\beta_k^I} \cdot n^{\beta_n^I}$, for $I \in \{S, M\}$. The vacancy cost functions are given by $\eta^s(v) = \frac{\kappa_{0s} v^{1+\kappa_{1s}}}{1+\kappa_{1s}}$, so the marginal cost is given by $\kappa_{0s} v^{\kappa_{1s}}$, for $s \in \{l, h\}$. The outside option distribution is distributed uniform with upper bound λ_s, for $s \in \{l, h\}$.

Calibration To simulate the model, I need to calibrate the parameters. For a subset of parameters, I take values from the related literature (*fixed parameters*). The rest I choose them to match empirical moments (*calibrated parameters*). Whenever relevant and possible, I use values for 2019 (last year of my sample) to better approximate current policy analysis. Monetary values are in 2022 dollars (August).

Table D.I: Fixed parameters

Parameters	Value	Source
$\{\alpha_l, \alpha_h\}$	$\{0.68, 0.32\}$	CPS
$\{\beta_n^S, \beta_k^S\}$	$\{0.65, 0.14\}$	BEA, Lamadon et al. (2022)
$\{\beta_n^M, \beta_k^M\}$	$\{0.44, 0.35\}$	BEA, Lamadon et al. (2022)
$r^*(1-t^*)$ $(I = S)$	0.032	Piketty and Zucman (2014), Bachas et al. (2022)
$r^*(1-t^*)$ $(I = M)$	0.052	Piketty and Zucman (2014), Bachas et al. (2022)
t	0.2	US statutory corporate tax rate
$\{y_0, \tau\}$	$\{15.92, 0.276\}$	Piketty et al. (2018)
ζ	1	-

Notes: All monetary values are in thousands of dollars of 2019.

Fixed parameters: Table D.I summarizes the fixed parameters. I set $\{\alpha_l, \alpha_h\} = \{0.68, 0.32\}$, based on the distribution of skill within the working age population in the CPS Basic files. To compute factor shares, I use data from the BEA tables on the Composition of Gross Output by Industry and define the labor share (LS) as compensation of employees over the sum of compensation of employees and gross operating surplus. I do this for each of the industries used in the empirical analysis of the groups "exposed services" and "manufacturing". Then, I define $\beta_n^I = b \cdot LS$ and $\beta_k^I = b \cdot (1 - LS)$, for $I \in \{S, M\}$, where b is a returns to scale parameter, which I set equal to 0.79 based on Lamadon et al. (2022). This yields $\{\beta_n^S, \beta_k^S\} = \{0.65, 0.14\}$ and $\{\beta_n^M, \beta_k^M\} = \{0.44, 0.35\}$. To calibrate the foreign return to capital, I use the fact that the ratio of global capital to global output is around 500%, and the global capital share of output is around 30%, so the global pre-tax return is around 30%/500% = 6% (Piketty and Zucman, 2014). And the global capital tax rate is around 30% (Bachas et al., 2022), so the global after-tax return is around 4.2%. To accommodate differential capital mobility by differential transportation costs paid in terms of lower returns, I assume that the global after-tax return is 3.2% for $I = S$ and 5.2% for $I = M$. The parameter calibration is under a given tax system, which I define as follows. I use $t = 20\%$, which is the statutory corporate tax rate, and calibrate the given income tax system as follows. I use Piketty et al. (2018) files and estimate linear regressions of taxes paid (post-tax incomes minus pre-tax incomes, including all taxes and transfers apportioned) over pre-tax incomes, restricting to working-age units whose total income is almost exclusively composed by labor income and whose annual incomes are lower than \$250,000. The relationship is surprisingly linear, being the current tax system reasonably approximated by a universal lump-sum of almost \$16,000 and a flat income tax rate of 27.6%. Finally, I set $\zeta = 1$ so the social welfare function is logarithmic.

Table D.II: Calibrated parameters

Panel (a): Moments

Moment	Source	Data	Model
Unemployment rates ($s = \{l, h\}$)	CPS	$\{0.049, 0.024\}$	$\{0.046, 0.054\}$
Job-filling rates ($I = \{S, M\}$)	JOLTS	$\{0.825, 0.774\}$	$\{0.752, 0.831\}$
Ratio employment to establishments ($I = \{S, M\}$)	QCEW	$\{9.90, 29.89\}$	$\{9.94, 25.86\}$
Annual pre-tax earnings ($s = \{l, h\}$)	CPS	$\{13.20, 82.42\}$	$\{13.20, 76.85\}$
Labor force participation ($s = \{l, h\}$)	CPS	$\{0.570, 0.737\}$	$\{0.583, 0.675\}$
Profit per establishment ($I = \{S, M\}$)	BEA, QCEW	$\{198.08, 314.64\}$	$\{199.94, 345.96\}$
Average markdown ($s = \{l, h\}$)	Berger et al. (2022)	$\{0.72, 0.72\}$	$\{0.516, 0.794\}$

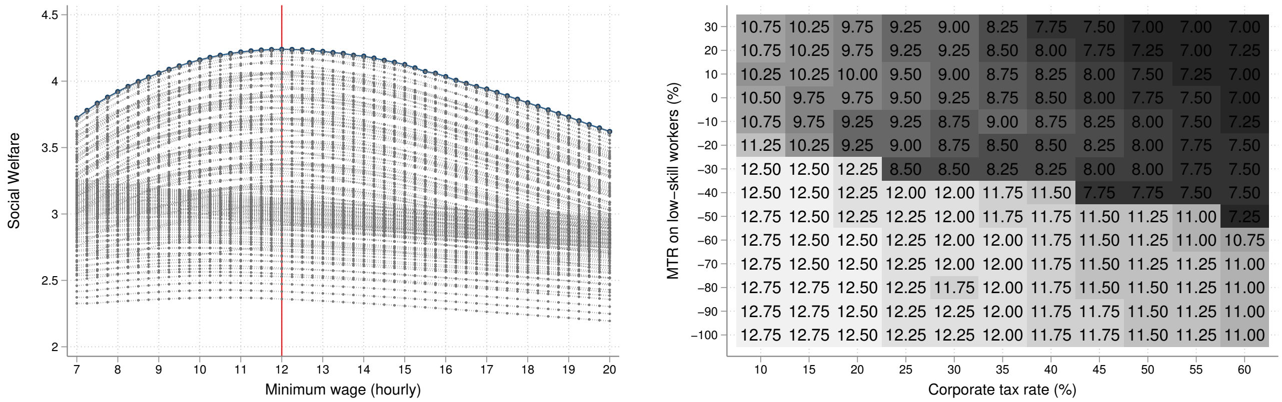
Panel (b): Parameters

Parameter	Value	Parameter	Value
δ_{0l}	0.85	δ_{0h}	0.92
δ_{1l}	0.51	δ_{1h}	0.79
λ_l	15.62	λ_h	77.97
K_S	0.038	K_M	0.008
ψ^S	31.46	ψ^M	36.78
κ_{0l}	0.727	κ_{0h}	0.239
κ_{1l}	0.987	κ_{1h}	1.233

Notes: All monetary values are in thousands of dollars of 2019.

Calibrated parameters: Table D.II summarizes the moments matched and the calibrated parameters. I solve the model to calibrate the following parameters by matching the following moments, separately for low-skill workers in exposed services and high-skill workers in manufacturing. I match skill-specific unemployment rates computed using the CPS and industry-specific job-filling rates (hires over postings) computed using JOLTS data to discipline the matching function parameters, $\{\delta_{0l}, \delta_{01}, \delta_{1l}, \delta_{1h}\}$. I match industry-specific ratios of establishments to employment computed using the QCEW files to discipline the mass of capitalists, $\{K_S, K_M\}$. I match skill-specific labor force participation rates computed using the CPS to discipline the upper bounds of the opportunity cost distribution, $\{\lambda_l, \lambda_h\}$. Finally, I match the average profit per establishment computed using BEA data, the skill-specific average pre-tax annual earnings computed using the CPS, and the average wage markdown estimated by Berger et al. (2022) to discipline the productivity and vacancy creation functions, $\{\psi^S, \psi^M, \kappa_{0l}, \kappa_{0h}, \kappa_{1l}, \kappa_{1h}\}$.

Figure D.I: Simulation results



(a) Social welfare versus minimum wage

(b) Optimal minimum wage given taxes

Notes: This figure presents the results of the simulation exercises using the calibration procedures described in this appendix.

Exercise and results With the calibrated parameters, I proceed to solve the model for different combinations of tax parameters and minimum wages. The policy parameters are reduced to $\{\tau_l, \tau_h, t, \overline{w}\}$, with y_0 recovered using the budget constraint. To simplify the analysis, given its second-order relevance, I fix $\tau_h = 0.3$ and focus the attention on τ_l, t, and $\overline{w}$. I solve the model for 154 permutations of tax parameters and 53 hourly minimum wage values and compute social welfare for each combination. Figure D.I summarizes the results. Panel (a) plots social welfare against hourly minimum wages given tax parameters. Each gray line represents one of the 154 tax combinations. The blue line represents the social welfare envelope (i.e., the tax system that maximizes social welfare given a minimum wage). Panel (b) shows the optimal minimum wage (i.e., the minimum wage that maximizes social welfare) for each of the 154 tax combinations.

There are three messages from Panel (a). First, given taxes, social welfare follows a concave trajectory against minimum wages. That is, social welfare increases with the minimum wage until a point that starts decreasing. This is explained by the fact that wage effects tend to dominate employment effects at low minimum wages, but the effect is reverted at some point. Second, different tax systems yield very different levels of social welfare. Third, the envelope suggests that the optimal minimum wage is far from the market level. The exact number ($12 dollars the hour) should be not taken literally given the simplicity of the exercise and, if anything, should be considered a lower bound given the efficiency properties of the model. Interestingly, at the optimal minimum wage of $12 dollars the hour, the tax system consists on an aggressive EITC (with $\tau_l = -100\%$) and a corporate tax rate of 35%, suggesting that the joint optimum uses all instruments in tandem.

Similarly, there are three messages from Panel (b). First, the optimal minimum wage varies with the tax system considered. That is, there is vast heterogeneity in the turning points of each of the gray lines plotted in Panel (a). Second, optimal minimum wages seem to be larger when the EITC is larger (in magnitude), and when the corporate tax rate is lower. This supports the intuition developed throughout the paper, which suggests that minimum wages complement tax-based transfers to low-skill workers and substitute profit redistribution based on corporate tax rates. Third, together with Panel (a), it is suggested that social welfare is maximized when both the minimum wage and the corporate tax rate are set at “intermediate” values. Since the distortions of each policy are increasing, the planner benefits from redistribute profits using both instruments, rather than just using larger corporate tax rates or large minimum wages.

Appendix Bibliography

Aaronson, D., E. French, I. Sorkin, and T. To (2018). Industry dynamics and the minimum wage: A putty-clay approach. *International Economic Review 59*(1), 51–84.

Acemoglu, D. and J. S. Pischke (1999). The structure of wages and investment in general training. *Journal of Political Economy 107*(3), 539–572.

Allegretto, S. and M. Reich (2018). Are local minimum wages absorbed by price increases? Estimates from internet-based restaurant menus. *ILR Review 71*(1), 35–63.

Ashenfelter, O. and Š. Jurajda (2022). Minimum wages, wages, and price pass-through: The case of McDonald's Restaurants. *Journal of Labor Economics 40*(S1), S179–S201.

Bachas, P., M. H. Fisher-Post, A. Jensen, and G. Zucman (2022). Globalization and factor income taxation. *Working Paper*.

Berger, D. W., K. F. Herkenhoff, and S. Mongey (2022). Minimum wages and welfare. *Working Paper*.

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

Bosch, M. and M. Manacorda (2010). Minimum wages and earnings inequality in urban Mexico. *American Economic Journal: Applied Economics 2*(4), 128–49.

Butschek, S. (2022). Raising the bar: Minimum wages and employers' hiring standards. *American Economic Journal: Economic Policy*.

Cengiz, D., A. Dube, A. Lindner, and D. Zentler-Munro (2022). Seeing beyond the trees: Using machine learning to estimate the impact of minimum wages on labor market outcomes. *Journal of Labor Economics 40*(S1), S203–S247.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics 134*(3), 1405–1454.

Clemens, J., L. B. Kahn, and J. Meer (2021). Dropouts need not apply? The minimum wage and skill upgrading. *Journal of Labor Economics 39*(S1), S107–S149.

Coviello, D., E. Deserranno, and N. Persico (2021). Minimum wage and individual worker productivity: Evidence from a large US retailer. *Journal of Political Economy*.

Dube, A., T. W. Lester, and M. Reich (2016). Minimum wage shocks, employment flows, and labor market frictions. *Journal of Labor Economics 34*(3), 663–704.

Dustmann, C., A. Lindner, U. Schönberg, M. Umkehrer, and P. Vom Berge (2022). Reallocation effects of the minimum wage: Evidence from Germany. *Quarterly Journal of Economics 137*(1), 267–328.

Gittings, R. K. and I. M. Schmutte (2016). Getting handcuffs on an octopus: Minimum wages, employment, and turnover. *ILR Review 69*(5), 1133–1170.

Haanwinckel, D. and R. R. Soares (2021). Workforce composition, productivity, and labor regulations in a compensating differentials theory of informality. *Review of Economic Studies 88*(6), 2970–3010.

Harasztosi, P. and A. Lindner (2019). Who pays for the minimum wage? *The American Economic Review 109*(8), 2693–2727.

Hurst, E., P. Kehoe, E. Pastorino, and T. Winberry (2022). The Distributional Impact of the Minimum Wage in the Short and Long Run. *Working Paper*.

Jardim, E., M. C. Long, R. Plotnick, E. Van Inwegen, J. Vigdor, and H. Wething (2022). Minimum-wage increases and low-wage employment: Evidence from seattle. *American Economic Journal: Economic Policy 14*(2), 263–314.

Ku, H. (2022). Does minimum wage increase labor productivity? Evidence from piece rate workers. *Journal of Labor Economics 40*(2), 325–359.

Lamadon, T., M. Mogstad, and B. Setzler (2022). Imperfect competition, compensating differentials and rent sharing in the US labor market. *American Economic Review 112*(1), 169–212.

Leung, J. H. (2021). Minimum wage and real wage inequality: Evidence from pass-through to retail prices. *Review of Economics and Statistics 103*(4), 754–769.

MaCurdy, T. (2015). How effective is the minimum wage at supporting the poor? *Journal of Political Economy 123*(2), 497–545.

Mayneris, F., S. Poncet, and T. Zhang (2018). Improving or disappearing: Firm-level adjustments to minimum wages in China. *Journal of Development Economics 135*, 20–42.

Meghir, C., R. Narita, and J.-M. Robin (2015). Wages and informality in developing countries. *American Economic Review 105*(4), 1509–46.

Pérez, J. (2020). The minimum wage in formal and informal sectors: Evidence from an inflation shock. *World Development 133*, 104999.

Piketty, T., E. Saez, and G. Zucman (2018). Distributional national accounts: Methods and estimates for the United States. *The Quarterly Journal of Economics 133*(2), 553–609.

Piketty, T. and G. Zucman (2014). Capital is back: Wealth-income ratios in rich countries 1700–2010. *The Quarterly Journal of Economics 129*(3), 1255–1310.

Renkin, T., C. Montialoux, and M. Siegenthaler (2022). The pass-through of minimum wages into US retail prices: Evidence from supermarket scanner data. *Review of Economics and Statistics 104*(5), 890–908.

Riley, R. and C. R. Bondibene (2017). Raising the standard: Minimum wages and firm productivity. *Labour Economics 44*, 27–50.

Robinson, J. (1933). *The economics of imperfect competition*. Springer.

Ruffini, K. (2021). Higher wages, service quality, and firm profitability: Evidence from nursing homes and minimum wage reforms. *The Review of Economics and Statistics*.

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

Shapiro, C. and J. E. Stiglitz (1984). Equilibrium unemployment as a worker discipline device. *The American Economic Review 74*(3), 433–444.

Sorkin, I. (2015). Are there long-run effects of the minimum wage? *Review of Economic Dynamics 18*(2), 306–333.

Vaghul, K. and B. Zipperer (2016). Historical state and sub-state minimum wage data. *Washington Center for Equitable Growth*.