Minimum Wages and Optimal Redistribution: The Role of Firm Profits*

Damián Vergara
Princeton University

Abstract

This paper analyzes the minimum wage as a redistributive tool. Empirically, I find that the average US state-level minimum wage increase in the 1997-2019 period raised low-skill workers' earnings at the expense of firm profits. Motivated by this fact, I characterize the desirability of the minimum wage in a setting with firm profits and optimal corporate and labor income taxes and transfers, considering both a stylized neoclassical labor market and a directed search model of the labor market. I show that a minimum wage is superfluous when the corporate tax is non-distortionary. A binding minimum wage is desirable, however, when it redistributes profits more efficiently than the corporate tax. This state of affairs prevails in two empirically plausible scenarios: when capital mobility keeps corporate taxes low and when firms capture low-wage income subsidies such as the Earned Income Tax Credit.

*Email: damianvergara@princeton.edu. First version: September 2021. This version: October 2023. I am especially grateful to Danny Yagan, Pat Kline, and Emmanuel Saez, for their continuous guidance and encouragement throughout this project. I also thank Alan Auerbach, Nano Barahona, Sydnee Caldwell, David Card, Patricio Domínguez, Peter Ganong, Cecile Gaubert, Andrés González, Henrik Kleven, Attila Lindner, Pablo Muñoz, Cristóbal Otero, Demian Pouzo, Michael Reich, Marco Rojas, Jesse Rothstein, Joel Slemrod, Dustin Swonder, Ivan Werning, Harry Wheeler, Gabriel Zucman, my discussants Euiyoung Jung, Benjamin Glass, and Thomas Winberry, and seminar participants at Bocconi University, the GRAPE-IMD Online Seminar, the IIPF 2023 Annual Congress, Northwestern University, the NTA 114th Annual Conference on Taxation, the OECD (CFE and ELS/JAI divisions), the Online Public Finance Workshop for Graduate Students, Princeton University, the XIII RIDGE Forum Workshop on Public Economics, UC Berkeley, UCLA, Universidad Adolfo Ibáñez, University of Chicago (Booth and Harris), University of Michigan, and Yale University 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 titles "When do Minimum Wages Increase Social Welfare? A Sufficient Statistics Analysis with Taxes and Transfers" and "Minimum Wages and Optimal Redistribution". Usual disclaimers apply.

"The simplest expedient which can be imagined for keeping the wages of labor up to the desirable point would be to fix them by law; the ground of decision being, not the state of the labor market, but natural equity; to provide that the workmen shall have reasonable wages, and the capitalist reasonable profits."

John Stuart Mill - *Principles of Political Economy* (1884).

1 Introduction

Governments use income taxes and transfers to redistribute to low-income individuals. However, those taxes and transfers can be distortionary, yielding equity-efficiency tradeoffs (Mirrlees, 1971). Motivated by the vast literature that suggests that minimum wages increase incomes of low earners with limited reductions in employment (Manning, 2021a), this paper asks whether a minimum wage can relax the government's tradeoffs and enable more efficient redistribution than income taxes and transfers alone.

Economists have long debated this question. Stigler (1946) argued that a minimum wage is inefficient relative to income-based taxes and transfers given its effects on employment. More generally, the production efficiency results of Diamond and Mirrlees (1971a,b) provide a strong case against the use of distortions on production inputs as part of optimal policies for revenue and redistribution. Not surprisingly, attempts to formally assess the desirability of the minimum wage when income taxes and transfers are available have failed to provide strong support for the minimum wage. Its desirability is either absent or reliant on knife-edge assumptions, missing instruments, or particular labor market frictions (e.g., Allen, 1987; Guesnerie and Roberts, 1987; Marceau and Boadway, 1994; Hungerbühler and Lehmann, 2009; Lee and Saez, 2012; Cahuc and Laroque, 2014; Gerritsen and Jacobs, 2020; Lavecchia, 2020).

The extensive literature mentioned above, however, abstracts from the existence of firm profits, either by assuming zero profit conditions or non-distortionary profit taxation. Both conditions are difficult to motivate in practice. Profits have substantially increased in recent decades, affecting the distribution of incomes between labor and capital (Karabarbounis and Neiman, 2014; Piketty and Zucman, 2014; Piketty et al., 2018; Stansbury and Summers, 2020). Moreover, the increase in profits has not been accompanied by an increase in effective corporate taxes. If anything, capital mobility and international tax competition have made it more difficult to tax capital incomes (Devereux et al., 2008, 2021; Bachas et al., 2023). As a result, the design of efficient policies to redistribute profits is a pressing policy issue.

This paper departs from the recent literature on optimal minimum wages by formally including firm profits in the distributional problem, giving rise to a new set of interactions between the tax system and the minimum wage. The inclusion of profits is shown to be important because the minimum wage works as an indirect tax on profits and, therefore, relates not only to the income tax system but also to the corporate tax. The idea that the minimum wage can affect profits is not new. Mill (1884) suggested that a minimum wage was the simplest way to redistribute profits to raise the incomes of low earners, and recent empirical evidence from outside the US suggests that the minimum wage has incidence on corporate profits (Draca et al., 2011; Harasztosi and Lindner, 2019; Drucker et al., 2021). The contribution of this paper is, therefore, to formalize this intuition and better understand the conditions under which complementing the tax system with a minimum wage is socially desirable.

To motivate the importance of profits, I follow Cengiz et al. (2019, 2022) and exploit US state-level variation in minimum wages in the 1997-2019 period to show that minimum wages in the US have increased average pre-tax earnings of low-skill workers (including employment effects), have decreased government expenses on income maintenance benefits, and have decreased average profits in exposed industries (food and accommodation, retail trade, and low-skill health services). As a consequence, I show that the labor share has increased in exposed industries as a consequence of the minimum wage. These margins configure the main tradeoffs explored throughout the paper. The minimum wage redistributes profits to wages, helping the government to relax the budget constraint at the cost of potential disemployment effects.

This distributional tension is studied by modeling a utilitarian social planner that chooses the minimum wage, the labor income tax-and-transfer schedule, and the corporate tax rate, to maximize social welfare. I present the key results in two models of the labor market. First, I consider a simple neoclassical model with limited heterogeneity and a frictionless labor market. Second, I produce full optimal policy results in a rich non-neoclassical model with directed search and two-sided heterogeneity. While the neoclassical model allows me to present the fundamental intuitions of the analysis in a simple way, the directed search model considers additional empirically relevant mechanisms through which the minimum wage can affect welfare, at the time of providing results that are more appropriate for quantitative exercises. Importantly, in both models, the decentralized equilibrium is efficient. This allows me to focus on redistribution by ruling out *Pigouvian* rationales for the minimum wage motivated by, for example, inefficient monopsony power (Robinson, 1933), information asymmetries (Burdett and Mortensen, 1998), hold-up (Acemoglu, 2001), or misallocation (Berger et al., 2022; Dustmann et al., 2022).

The neoclassical framework with perfect competition considers equally productive workers that make extensive margin labor supply decisions, and capitalists that allocate capital between domestic firms with decreasing returns to scale and a foreign investment with a fixed after-tax return. While stylized, this simple framework contains the main ingredients for the analysis: positive firm profits, corporate tax distortions on pre-tax outcomes, worker-level behavioral responses, and general equilibrium effects.

In this model, a minimum wage redistributes profits to workers, yielding an equity gain when the planner values workers more than firm owners. However, the minimum wage generates efficiency costs that are proportional to its employment effects. The natural question that follows is: if the planner can use taxes and transfers, can similar equity gains be achieved more efficiently by only using the tax system? For example, the planner could use the corporate tax to fund transfers to workers that replicate the minimum wage allocation. In this context, I show that the desirability of the minimum wage under optimal taxes is a positive function of the distortions induced by the corporate tax. That is, while the tax system is possibly more efficient than the minimum wage in isolation, complementing it with a binding minimum wage can make redistribution more efficient when corporate taxes are distortionary.

To see why, consider the allocation with optimal taxes but no minimum wage. The planner can implement the following reform package: a small minimum wage, combined with a reduction in transfers to workers that keeps workers' consumption constant, and a corporate tax cut that keeps employment constant by offsetting the minimum wage effect on labor demand. This reform does not affect workers' welfare while generating fiscal savings from reduced workers' transfers and fiscal costs from reduced corporate tax revenue. Hence, the net fiscal externality determines whether the reform is welfare-improving. If the corporate tax induces *large* distortions (relative to the minimum wage), then a *small* corporate tax cut will be sufficient to compensate for the labor demand shock, so the negative fiscal externality of the reform will be *small*. This makes the reform more likely to be desirable: corporate tax distortions make corporate tax cuts attractive. On the contrary, if the corporate tax is close to non-distortionary, its cut will have to be massive to keep employment constant, thus making the negative fiscal externality sizable and the case for the minimum wage dubious. In the limit, when corporate taxes do not distort employment, the minimum wage becomes superfluous regardless of its employment effects.

Further intuition follows from benchmarking the minimum wage to an in-work benefit like the Earned Income Tax Credit (EITC), as it is usually done in the public debate. Suppose the government redistributes to workers using an EITC. This transfer has to be funded by taxes on other inputs, such as profits. Also, labor supply effects can push equilibrium wages downward (Rothstein, 2010; Zurla, 2021; Gravoueille, 2023). The minimum wage prevents wages from decreasing and allows the transfer to workers to be partially paid by firm profits. This positive fiscal externality implies that the minimum wage can be interpreted as a tax on profits. When corporate taxes distort pre-tax profits, they cannot fully correct this incidence distortion, so it becomes an open question of which combination of minimum wages and corporate taxes is optimal for taxing profits. Since the marginal efficiency cost of taxation is usually increasing in the tax rate (Auerbach and Hines, 2002), the combination that minimizes the aggregate distortions of the redistribution of profits will possibly consider both minimum wages *and* corporate taxes.

Consistent with recent evidence (Garrett et al., 2020; Curtis et al., 2022; Kennedy et al., 2023), the model suggests that the corporate tax distortion increases with capital intensity. Building on that intuition, I show that the minimum wage can be desirable as a kind of industry-specific corporate tax when unaffected industries are particularly capital intensive (and, therefore, responsive to corporate taxes). To see why, consider the US case, where the minimum wage mostly affects labor-intensive services industries, contrary to corporate taxes that affect all industries including capital-intensive sectors such as manufacturing. In these situations, governments can benefit from using the minimum wage to tax profits in the affected industries, relaxing distortions in unaffected industries. This suggests that the minimum wage can be used to address the consequences of international tax competition, which have prevented governments from enforcing large effective corporate taxes because of international capital mobility.

One caveat of the simple model is that it oversimplifies the effects of the minimum wage on the labor market. The increasing consensus is that labor markets have frictions and complex competitive structures (Manning, 2021b; Card, 2022) that mediate the wage and employment effects – even affecting workers who earn more than the minimum wage – usually in the presence of firm-level heterogeneity and endogenous entry of firms. Motivated by this reflection, I reproduce the analysis in a more general model of the labor market that accommodates more realistic predictions of minimum wage effects and provides results that, while less stylized, are better suited for quantitative analyses.

The model 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, wage and employment spillovers to non-minimum wage jobs, and firm profits incidence. A population of workers with heterogeneous skills and costs of participating in the labor market decides whether to enter the labor market and which jobs to apply to. A corresponding population of capitalists with heterogeneous productivities and technologies decides whether to create firms, how many vacancies to post, and attaches a wage to those vacancies. Minimum wages 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. As mentioned above, decentralized allocations are shown to be constrained efficient (as in standard directed search models, e.g. Moen, 1997) so the analysis keeps the attention on the redistributive role of the minimum wage.

The main conclusions of the neoclassical analysis hold in the extended model. The higher the distortions of the corporate tax, the more likely a binding minimum wage is to be desirable, especially when unaffected sectors are particularly capital intensive. In addition, the richness of the model allows me to perform two additional exercises that complement the theoretical analysis: a sufficient statistics analysis of a small minimum wage increase that uses the reduced form estimates, and a simple numerical exercise that informs a global analysis of the optimal minimum wage. Both exercises support the general message of this paper. Under plausible conditions, optimal redistribution takes the form of a "three-legged stool": targeted income taxes and transfers, positive corporate taxes, and a binding minimum wage.

Related literature This paper contributes to the normative analysis of the minimum wage in frameworks with optimal taxes by incorporating profits into the distributional equation. As far as I am aware, this paper is the first to explore the interaction between the minimum wage and the corporate tax. The (inconclusive) previous literature either imposes zero profit conditions or assumes that profits can be taxed with no meaningful efficiency costs, so it focuses on the interaction between the minimum wage and the income tax system. Allen (1987) and Guesnerie and Roberts (1987) consider two-type models with intensive margin responses à la Stiglitz (1982) and argue that the minimum wage is superfluous when non-linear income taxation is available. Marceau and Boadway (1994) and Boadway and Cuff (2001) overturn these results by including participation costs and a continuum of types, respectively. Gerritsen (2023) provides stronger support for the minimum wage when preference heterogeneity induces dispersion in hours worked conditional on wage. The rest of the literature has mostly focused on extensive margin responses. Hungerbühler and Lehmann (2009) and Lavecchia (2020) use random search models and focus on congestion inefficiencies that cannot be completely addressed by the income tax system. Lee and Saez (2012) use a perfectly competitive model with two skill types and find that the minimum wage increases the efficiency of the tax system when rationing is efficient, a result contested by Cahuc and Laroque (2014) who use a neoclassical monopsonistic model and conclude that the minimum wage is superfluous when there is a continuum of skill types. Gerritsen and Jacobs (2020) argue that the desirability of the minimum wage relies on the incentives it generates on skill formation to insure against unemployment.[1,2]

A different literature studies the welfare consequences of the minimum wage using quantitative analyses based on rich structural general equilibrium models that abstract from the optimal tax question. Flinn (2006), Wu (2021), Ahlfeldt et al. (2023), and Drechsel-Grau (2023) focus on efficiency rationales motivated by labor market imperfections, while Berger et al. (2022) and Hurst et al. (2023) consider both efficiency and redistribution. I see this paper (and the ones referenced above) as complementary to this literature as it provides additional qualitative insights on the tradeoffs involved in the optimal design of the minimum wage, while the structural literature allows for richer quantitative explorations that include additional margins such as dynamics, strategic interactions, or spatial distortions.

This paper also adds to different sub-literatures on optimal redistribution. First, there is a theoretical literature that explores deviations from production efficiency (Diamond and Mirrlees, 1971a,b). Naito (1999) developed an argument for production inefficiency (extended in Saez, 2004, Scheuer, 2014, Gomes et al., 2018, and Costinot and Werning, 2023) based on technological constraints and missing instruments. The argument for production inefficiency in this paper differs from this tradition since the focus is on the existence of profits whose taxation is costly. This paper also contributes to the literature that explores whether the combination of different policy instruments can improve the efficiency of redistribution (e.g., Atkinson and Stiglitz, 1976; Saez, 2002; Gaubert et al., 2020; Ho and Pavoni, 2020; Ferey, 2022) by exploring the interaction between the minimum wage and the tax system. This paper also adds to the analysis of redistributive policies in labor markets with frictions (e.g., Hungerbühler et al., 2006; Stantcheva, 2014; Sleet and Yazici, 2017; Kroft et al., 2020; Bagger et al., 2021; Mousavi, 2022; Craig, 2023; Doligalski et al., 2023; Hummel, 2023), and to the analysis of redistribution between capital and labor income (e.g., Scheuer, 2014; Eeckhout et al., 2021; Hummel, 2022; Atesagaoglu and Yazici, 2023).

[1] Allen (1987), Boadway and Cuff (2001), Lee and Saez (2012), Gerritsen and Jacobs (2020), and Lavecchia (2020) work in frameworks with no firm profits in equilibrium. Guesnerie and Roberts (1987), Marceau and Boadway (1994), and Gerritsen (2023) allow for profits whose taxation is non-distortionary. Cahuc and Laroque (2014) allows for profits whose taxation only distorts the extensive margin (i.e., does not affect wages or employment).

[2] Dworczak et al. (2021) indirectly analyzes the redistributive consequences of the minimum wage by studying redistribution through markets and price controls using mechanism design techniques.

This paper also contributes to the policy discussion around the consequences of international tax competition. Existing proposals advocate for international coordination to erode incentives for capital mobility through global minimum taxes and information exchange (e.g., Johannesen, 2022; Devereux et al., 2023; Zucman, 2023). This paper highlights the role of the minimum wage as a tax on profits, which becomes more relevant when exposed industries are less sensitive to changes in the corporate tax.

Finally, the paper also adds to the vast positive literature on the minimum wage. The directed search model contributes to a theoretical literature that tries to rationalize minimum wage evidence. See, for example, Engbom and Moser (2022), Haanwinckel (2023), Vogel (2023), and the structural literature referenced above. Also, the empirical results presented as motivation 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 effects on wages and employment (Manning, 2021a). Results on income maintenance transfers and other fiscal outcomes complement the evidence presented in Reich and West (2015) and Dube (2019). Finally, the empirical results on profits and the labor share are in line with the findings of Draca et al. (2011), Harasztosi and Lindner (2019), and Drucker et al. (2021).

Structure of the paper Section 2 presents empirical evidence to motivate the tradeoffs studied throughout the paper. Section 3 proceeds with the policy analysis using a stylized neoclassical model of the labor market. Section 4 extends the analysis to a model with richer worker- and firm-level heterogeneity and directed search. Section 5 briefly discusses the limitations of the analysis and suggests avenues for future research. Finally, Section 6 concludes. All proofs are presented in Appendix B.1.

2 Empirical effects of minimum wages on workers and firm profits

To motivate the relevance of the tradeoffs discussed throughout the paper, I provide empirical evidence on the effects of minimum wages on employment, wages, transfers, and profits. In what follows, I provide a general description of the empirical strategy and the data. Appendix A contains additional details on the estimated models, data sources, and sample restrictions, as well as additional results.

Empirical strategy The analysis closely follows Cengiz et al. (2019, 2022). I use state-level variation in minimum wages to estimate stacked event studies. State-level minimum wage data is taken from Vaghul and Zipperer (2016) for the 1997-2019 period. An event is defined as a state-level real hourly minimum wage increase of at least \$0.25 (in 2016 dollars) in a state with at least 2% of the employed population affected. These restrictions are imposed to focus on minimum wage increases that are likely to have effects on the labor market. 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 allows me to observe the outcomes from three years before to four years after. This results in 50 “valid” state-level events.[3]

With these events, I estimate stacked event studies, which address multiple treatment challenges and potential biases driven by treatment effect heterogeneity (Cengiz et al., 2019, 2022; Gardner, 2021; Baker et al., 2022; Dube et al., 2023). I implement the stacked event studies 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:

$$\log Y_{ite} = \sum_{\tau=-3}^{4} \beta_\tau D_{i\tau e} + \alpha_{ie} + \gamma_{cd(i)te} + X_{it}'\rho_e + \epsilon_{ite}, \tag{1}$$

where i, t, and e index state, year, and event, respectively, Y_{ite} is an outcome of interest, $D_{i\tau e}$ are event indicators with τ the distance from the event (in years), α_{ie} are state-by-event fixed effects, $\gamma_{cd(i)te}$ are census division-by-year-by-event fixed effects, and X_{it} are time-varying controls that include small state-level minimum wage increases and binding federal minimum wage increases, whose effect is allowed to vary by event e. The inclusion of time fixed effects that vary by census division allows me to better control for time-varying confounders that differentially affect states and industries while limiting the variation used for identification. Standard errors are clustered at the state level and regressions are weighted by the state-by-year average total population. When the outcome varies at the state-by-industry level, I 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.[4]

I also report pooled difference-in-difference estimations to summarize the average treatment effect in the post-period:

$$\log Y_{ite} = \beta T_{ie}\text{Post}_{te} + \alpha_{ie} + \gamma_{cd(i)te} + X_{it}'\rho_e + \epsilon_{ite}, \tag{2}$$

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 (1). Finally, to provide estimates of elasticities $d\log Y_{ite}/d\log \text{MW}_{ite}$, I also report IV coefficients from analog regressions that instrument $\log \text{MW}_{ite}$ with $T_{ie}\text{Post}_{te}$.[5]

[3]See Figure A.1 and Table A.1 for the state and year distribution of the events considered. As in Cengiz et al. (2019, 2022), 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.

[4]Table 2 shows that results hold in specifications with year-by-event and census region-by-year-by-event fixed effects. However, as shown in Appendix A, while the effects on workers' outcomes and transfers are very consistent across models, specifications based on industry-by-state data display pre-trends in some cases when using the less flexible time fixed effects, suggesting that controlling by census division time trends is the preferred specification.

[5]For a graphical representation of the first stage, see Figure A.2.

Outcomes and data Outcome variables are state or state-by-industry annual aggregates for the period 1997-2019. I focus on low-skill wages and employment, transfers, profits, and the labor share.

I use the NBER Merged Outgoing Rotation Group of the CPS to compute average pre-tax hourly wages and the Basic CPS monthly files to compute employment rates at the state-by-year level for low-skill workers. Low-skill workers are defined as not having a college degree. To report an overall effect on low-skill workers that includes wage and employment effects, I compute the average wage of active workers including the unemployed, which equals the average wage conditional on employment times the employment rate. I drop individuals aged 15 or less, self-employed, veterans, and whose hourly wage is in the upper half of the wage distribution when employed.[6] I use data on income maintenance benefits from the Bureau of Economic Analysis (BEA) to proxy for transfers disbursed to low-skill workers at the state-by-year level. This variable mainly includes Security Income (SSI) benefits, the EITC, the Additional Child Tax Credit, and Supplemental Nutrition Assistance Program (SNAP) benefits, among other minor assistance benefits. Absent firm-level microdata, I compute a measure of average profits at the industry-by-state-by-year level. Since aggregate profits are too noisy to study, I focus on profits per establishment and the number of establishments. I use the Gross Operating Surplus (GOS) estimates from the BEA as a proxy for state-by-industry aggregate profits and divide them by the average number of private establishments reported at the state-by-industry level in the Quarterly Census of Employment and Wages (QCEW). Noting that minimum wage workers are not evenly distributed across industries, I group industries into two large groups: exposed and non-exposed industries. Exposed industries include food and accommodation, retail trade, and low-skill health services.[7] Finally, I compute the labor share at the state-by-industry level, computed using standard formulas based on BEA data on GOS, taxes on inputs and imports net of subsidies, and compensation of employees.

Table 1 presents descriptive statistics of the non-stacked sample for the period 1997-2019 (51 states $\times$ 23 years). All values are annual and in 2016 dollars. The average annual pre-tax income of low-skill workers (including the unemployed) is $19,396. Average income maintenance benefits per working-age individual are $1,051 which represents around 5% of low-skill workers' pre-tax income. Average pre-tax profits per establishment are substantially larger than disposable incomes for workers: in exposed industries, the average pre-tax profit per establishment is almost 9 times the average pre-tax income of low-skill workers, suggesting equity gains from redistributing profits. Also, exposed industries are much more labor-intensive than non-exposed industries, with average labor shares of 0.67 and 0.45, respectively.

[6] Figure A.7 shows that the results are robust to relaxing this restriction except when including the top decile of the wage distribution, which attenuates the wage effects. This result is consistent with wage spillovers that are decreasing in the distance from the minimum wage, implying that the very top wages should not be affected by minimum wage reforms.

[7] A large empirical literature in the US characterizes food and accommodation and retail as the main exposed industries; see, for example, Dube et al. (2010) and Cengiz et al. (2019). Low-skill health and social services, such as the nursing home sector, are also highly exposed to the minimum wage; see, for example, Gandhi and Ruffini (2022) and Ruffini (Forthcoming).

Results Figures 1 and 2 plot the estimated coefficients $\{\beta_\tau\}_{\tau=-3}^{4}$ of equation (1) with their corresponding 95% confidence intervals. Table 2 reports the corresponding estimated β coefficient of equation (2) and the implied elasticity with respect to the change in the minimum wage.

Figure 1 reports effects on low-skill workers. Panel (a) of Figure 1 shows the results on the log average wage of low-skill workers including the unemployed (equal to the product of the average wage and the employment rate). When factoring both wage and employment effects, state-level minimum wage increases have raised the average wages of low-skill workers (including the low-skill workers who are unemployed or earn more than the minimum wage). The implied elasticity ranges between 0.11 and 0.15, meaning that a 1% increase in the state-level minimum wage leads to a 0.11%-0.15% increase in average pre-tax wages of low-skill workers. Table 2 and Figure A.4 show that there are no detectable effects on low-skill employment rates. That finding implies that the overall effect on low-skill workers is explained by an increase in the wage conditional on employment, which I show directly in Appendix A (Figure A.4 and Table A.3).[8] The Appendix further shows that the overall effect is homogenous across subgroups of low-skill workers (Figure A.6) and is not reproduced in the high-skill worker population (Figures A.3 and A.4, and Tables A.4 and A.5). These results are consistent with the vast literature that consistently finds positive wage effects with "elusive" employment effects (see Manning, 2021a for a review).[9]

Panel (b) of Figure 1 shows the results on the log income maintenance benefits per working-age individual. Income maintenance benefits per working-age individual have decreased after state-level minimum wage increases, with an implied elasticity that ranges between -0.35 and -0.45. This finding suggests that higher pre-tax earnings lead to smaller means-tested transfers, thus generating a positive fiscal externality for the government that attenuates the pre-tax gains of low-skill workers. This result is consistent with the empirical findings of Reich and West (2015) and Dube (2019) for the US context and with the micro-simulations made by Giupponi et al. (Forthcoming) in the UK context.[10]

Figure 2 reports effects on industry-level profits and labor shares. Panel (a) of Figure 2 shows the results on the log average profit per establishment. The average profit per establishment has decreased in exposed industries (food and accommodation, retail trade, and low-skill health services) after minimum wage reforms, with no effect on non-exposed industries, with an implied elasticity that ranges between -0.49 and -0.55. Table 2 and Figure A.9 show that there is no effect on the number of establishments, suggesting that the profit effect is driven by an intensive margin response. This result is consistent with the findings in Draca et al. (2011) for the UK, Harasztosi and Lindner (2019) for Hungary, and Drucker et al. (2021) for Israel. While these results should be interpreted with caution since they are based on

[8] Figure A.5 and Table A.3 also show null effects on average hours worked and labor force participation rates.

[9] While consistent, my results differ from Cengiz et al. (2019, 2022) in two dimensions. First, I focus on a broader group (low-skill workers) that is not exclusively composed of minimum wage workers. Second, I provide results that focus on the combined effect on wages and employment rather than the pure employment effect.

[10] Figure A.8 and Table A.6 show that medical benefits and gross federal income taxes do not respond to minimum wages, reinforcing the idea that worker-level fiscal externalities are mediated by targeted transfers based on pre-tax income levels.

aggregate data, they suggest that there is non-trivial profit incidence in exposed industries.[11]

Finally, Panel (b) of Figure 2 shows the results on the log labor share. The labor share increases in exposed industries, with no effect on non-exposed industries, with an implied elasticity that ranges between 0.13 and 0.16. This result provides a summary of the within-firm distributional impacts of the minimum wage. By increasing low-skill workers' earnings at the expense of firm profits, the minimum wage redistributes from firm owners to workers. This distributional tension is theoretically explored in the rest of the paper, by formally considering the efficiency costs and the fiscal externalities of the minimum wage, and its comparison with tax-and-transfer schemes that can replicate similar redistribution schemes.

3 Optimal minimum wage in a frictionless labor market

I start the formal analysis using a stylized model of the labor market with perfect competition. This approach generates sharp results that capture the main intuition of the analysis at the cost of delivering unrealistic predictions regarding the labor market effects of minimum wage reforms. Section 4 extends the analysis to a more general model of the labor market.

3.1 Setup

There is a population of equally productive workers normalized to 1 and a fixed number, N, of symmetric capitalists. Each capitalist owns capital and allocates it between a domestic firm with decreasing returns to scale and a foreign investment with a fixed after-tax return. Domestic firms are price takers in both the product and labor markets. There are no frictions in the labor market. The model focuses on extensive margin labor supply responses; hence, the concepts of wage, income, and earnings are used indistinctly.

Workers Workers are endowed with a scalar parameter $c \in \mathcal{C} = [0, C] \subset \mathbb{R}$, which is distributed with cdf F and pdf f. c represents the (dollar-valued) cost of participating in the labor market, which admits different interpretations such as the disutility of labor supply or other opportunity costs such as home production. Then, if a worker of type c works, she gets utility $u_1 = y_1 - c = w - T_1 - c$, where $y_1 = w - T_1$ is after-tax income (consumption), w is the wage, and T_1 is the total taxes net of transfers paid by employed workers. If a worker of type c does not work, she gets utility $u_0 = y_0 = -T_0$, where $-T_0 \geq 0$ is a government transfer. Workers work if $u_1 \geq u_0$, that is, if $\Delta y = y_1 - y_0 = w - T_1 + T_0 \geq c$. Then, aggregate labor supply is given by $L^S(w, T_0, T_1) = F(w - T_1 + T_0)$.

[11] The effect on firm owners can, in principle, generate fiscal externalities in other parts of the tax system. Figure A.10 and Table A.7 provide little evidence of these alternative fiscal externalities. I estimate no effects on business income per income tax return or on dividend income per income tax return, as reported in the Statistics of Income (SOI) state-level tables, and also no effects on taxes on production and imports net of subsidies, as reported by the BEA.

Capitalists Each capitalist is endowed with capital and a domestic production function with decreasing returns to scale denoted by $\widetilde{\phi}(l,k)$, with l employment, k capital, $\widetilde{\phi}_l > 0$, $\widetilde{\phi}_k > 0$, and $\widetilde{\phi}_{lk} \geq 0$. Profits are taxed with a linear corporate tax rate, t, and the output price is normalized to 1, so domestic after-tax profits are given by $(1-t)\widetilde{\pi}(l,k) = (1-t)\left(\widetilde{\phi}(l,k) - wl\right)$.

Capitalists decide whether to allocate capital to the domestic firm or to a foreign investment that has a fixed after-tax return, $\widetilde{r}^*$, taking w as given. Assuming an interior solution, the problem is solved in two stages where the capitalist allocates k given l and then optimizes l. Let $k(l,w,t)$ be the optimal domestic capital given l. Then, let $\phi(l,w,t) = \widetilde{\phi}(l,k(l,w,t))$ and $(1-t)\pi(l,w,t) = (1-t)\widetilde{\pi}(l,k(l,w,t))$, with $\phi_{ll} < 0$, $\phi_t \leq 0$, and ϕ_w ambiguous (depends on the capital-labor substitution patterns of $\widetilde{\phi}(l,k)$).[12] The problem of the domestic firm is, then, given by $\max_l (1-t)\pi(l,w,t) = \max_l (1-t)\left[\phi(l,w,t) - wl\right]$. The first-order condition states that the marginal revenue of labor equals the wage, $\phi_l(l,t,w) = w$, which implicitly defines a labor demand function, $l^D = l^D(w,t)$. I assume that the labor demand elasticities with respect to wages and corporate taxes, $\eta_w = -(\partial l^D(w,t)/\partial w)(w/l) > 0$ and $\eta_t = -(\partial l^D(w,t)/\partial t)(t/l) > 0$, respectively, are finite. Total labor demand is then given by $L^D(w,t) = Nl^D(w,t)$.

Because of decreasing returns to scale, firms have positive profits. Let the optimized after-tax profits (firms' value function) be denoted by $(1-t)\Pi$, where $\Pi = \Pi(w,t) = \pi(l(w,t),w,t)$. Note that, when there is a minimum wage, $\Pi_t = \phi_t$ and $\Pi_w = \phi_w - l$, which give form to pre-tax profits elasticities, $\epsilon_w = -\Pi_w(w/\Pi) > 0$ and $\epsilon_t = -\Pi_t(t/\Pi) > 0$, which are also assumed to be finite.[13]

The main innovation of this framework is that changes in the domestic corporate tax rate can generate distortions in pre-tax profits through distortions in labor demand, as captured by $\phi(l,t,w)$ and η_t. Minimum wages, $\overline{w}$, can also affect revenues on top of the direct effect on labor costs, thus implicitly affecting η_w. Both distortions are not isomorphic, since changes in $\overline{w}$ distort employment through movements along the demand curve, while changes in t distort employment through shifts in the demand curve.

Equilibrium The market equilibrium is given by $L^S(w,T_0,T_1) = L^D(w,t) = L$, with L total employment. The market clearing condition pins down the equilibrium wage, w, given (T_0,T_1,t). Changes in (T_0,T_1) generate shifts in the labor supply, while changes in t generate shifts in labor demand. In the presence of a binding minimum wage, $\overline{w}$, there can be excess labor supply, so the equilibrium condition depends on the rationing mechanism that allocates workers to employment (see below).

Planner's problem The social planner chooses the tax system, (T_0,T_1,t), and the minimum wage, $\overline{w}$, to maximize a (generalized) utilitarian social welfare function (SWF). I assume that the planner does

[12]The capital allocation problem, while intuitive, is instrumental to generating corporate tax distortions on pre-tax profits. The analysis can be extended to other sources of corporate tax distortions on pre-tax profits. For example, Appendix B.3 proposes a model of optimal effort exerted by firm owners that microfounds a similar reduced-form revenue function.

[13]The envelope theorem rules out effects on profits through changes in l. With no minimum wage, $\Pi_t = \phi_t + (\phi_w - l)(\partial w/\partial t)$.

not observe c and, therefore, is restricted to second-best allocations: the pair (T_0, T_1) is a non-linear income tax schedule that only depends on earnings. The choice of a linear corporate tax rate follows the same principle since, in the absence of firm-level heterogeneity, lump-sum taxes on profits would be non-distortionary. Also, a linear tax on profits better approximates the actual implementation of corporate taxation in practice.[14] I assume that the planner only observes domestic profits and, therefore, abstracts from the capital returns generated by foreign investments. Then, the SWF is given by:

$$SWF = (1-L)\omega_L G(y_0) + \int_{\mathcal{C}_1} \omega_L G(y_1 - c) dF(c) + N\omega_k G\left((1-t)\Pi\right), \tag{3}$$

where $\mathcal{C}_1 = \{c \in \mathcal{C} : \textit{individual is working}\}$, and G is an increasing and concave function that, together with the vector $\{\omega_L, \omega_k\}$ of exogenous Pareto weights on workers and capitalists, summarizes social preferences for redistribution. $\mathcal{C}_1$ is endogenous to policy choices, so the integration limits incorporate the incentive compatibility constraints. When workers are on their labor supply curve, $\mathcal{C}_1 = [0, \Delta y]$. This may not be the case when there is a binding minimum wage, since there may be involuntarily unemployed workers. The rationing mechanism, then, determines $\mathcal{C}_1$ when there is excess labor supply.

Assuming no exogenous expending requirement, the government budget constraint is given by:

$$(1-L)T_0 + LT_1 + tN\Pi = 0. \tag{4}$$

Let γ be the budget constraint multiplier. The average social marginal welfare weights (SMWWs) of non-employed workers, employed workers, and capitalists, are defined as $g_0 = \omega_L G'(y_0)/\gamma$, $g_1 = \omega_L \int_{\mathcal{C}_1} G'(y_1 - c) dF(c)/(L\gamma)$, and $g_k = \omega_k G'\left((1-t)\Pi\right)/\gamma$, respectively. SMWWs represent the social value of the marginal utility of consumption normalized by the social cost of raising funds, thus measuring the social value of redistributing one dollar uniformly across a group of individuals. At the optimum, the planner is indifferent between giving one more dollar to an individual i or having g_i more dollars of public funds.

Rationing As in Lee and Saez (2012), the analysis in this section assumes efficient rationing: workers that involuntarily lose their jobs because of $\overline{w}$ are those with the larger values of c. This assumption simplifies the exposition and derivations but is not critical for the qualitative results discussed below. More importantly, the analysis developed in the next section does not rely on any rationing assumption.

3.2 Optimal minimum wage with no tax system

To develop intuition, I first consider a case with no tax system. Let w^* denote the market wage with no minimum wage. The government chooses the minimum wage, $\overline{w}$, to maximize (3). Panels (a) and (b) of

[14] Non-linear taxation of profits has shown to be challenging in the real world because firms can reorganize their structure – for example, by splitting large firms into several small firms – to avoid the progressivity of the corporate tax. See, for example, Onji (2009), Best et al. (2015), Agostini et al. (2018), Bachas and Soto (2021), and Lobel et al. (Forthcoming).

Figure 3 illustrate the welfare effects of a minimum wage reform. Panel (a) shows the welfare effects of a marginal increase in the minimum wage from w^* to $w^* + d\overline{w}$. The gray area represents workers' surplus and the pink area represents firm profits. The reform generates two effects. First, there is a transfer from the firm (profits) to the employed workers (wages), which is represented by the yellow rectangle. This transfer is valued by the social planner whenever workers' incomes are more valued than firm profits. Second, there is a decrease in employment along the demand curve, from L to $L - d\overline{L}$, which generates a welfare loss equal to the red triangle. Because of efficient rationing, this welfare loss is second-order (rationed workers are indifferent between working and not working).

Panel (b) shows the welfare effects of a marginal change in the minimum wage departing from an already binding minimum wage, that is, a minimum wage increase from $\overline{w}$ to $\overline{w} + d\overline{w}$, with $\overline{w} > w^*$. In this case, the welfare losses for the displaced workers are no longer second-order. Then, to assess the desirability of further increases in the minimum wage, the social planner has to compare the welfare gains for employed workers, the welfare losses for firm owners, and the welfare losses for displaced workers. A situation where rationing is not efficient (e.g., is uniform) looks closer to Panel (b) even in cases where $\overline{w} = w^*$, since involuntarily unemployed workers are not guaranteed to have second-order surplus losses.

The following proposition formalizes this argument in the model with N capitalists. Define as $\underline{w}$ the value of c of the marginally employed worker, implicitly defined by $L^D(\overline{w}) = L^S(\underline{w})$. When $\overline{w} \leq w^*$, workers are on their labor supply and, therefore, $\overline{w} = \underline{w}$. When $\overline{w} > w^*$, there is excess of labor supply so $\overline{w} > \underline{w}$. Therefore, $\overline{w} - \underline{w}$ is a measure of the employment rent of the marginally employed worker.

Proposition 1. *Assume that rationing is efficient. Increasing the minimum wage, $\overline{w}$, just above the market wage, w^*, is welfare-improving if:*

$$Lg_1 - N\Pi\frac{\epsilon_w}{\overline{w}}g_k > 0. \tag{5}$$

Further increases of $\overline{w}$ are welfare-improving if:

$$Lg_1 - L\frac{\eta_w}{\overline{w}}g_u - N\Pi\frac{\epsilon_w}{\overline{w}}g_k > 0, \tag{6}$$

where $g_u = \omega_L\left(G\left(\overline{w} - \underline{w}\right) - G(0)\right)/\gamma$ is the social value of the utility loss of rationed workers.

Proposition 1 states that a binding minimum wage is desirable when the planner values employed workers' utility more than capitalists' profits. An increase in $\overline{w}$ generates a welfare gain of g_1 for the mass of L workers earning the minimum wage, $\overline{w}$. Likewise, an increase in $\overline{w}$ generates a welfare loss of $g_k(d\Pi/d\overline{w}) = -g_k\Pi\epsilon_w/\overline{w}$ for each of the N capitalists. These forces characterize a critical g_1/g_k ratio equal to $\epsilon_w N\Pi/L\overline{w}$ such that using the minimum wage for redistribution is desirable. The desirability of further increases in $\overline{w}$ also depends on the labor demand elasticity – which mediates the employment effect – and the degree of excess of supply – which mediates the welfare impact of the employment loss.[15]

3.3 Is the minimum wage desirable when the planner can also choose taxes?

The previous analysis formalizes the intuition that the minimum wage is a tool for redistributing profits to workers, with potential costs in employment. If taxes are available, however, the planner can mimic the same redistribution with a transfer to workers funded by corporate tax revenue. Is the minimum wage a relevant redistributive policy when the planner can freely choose the tax system?

Panels (c) and (d) of Figure 3 illustrate the effects of the proposed tax reform. As illustrated in Panel (c), any tax-based attempt for redistributing to employed workers (i.e., a decrease in T_1) is accompanied by an increase in labor supply. The increase in labor supply pushes the equilibrium wage down, thus increasing firm profits (Rothstein, 2010; Zurla, 2021; Gravoueille, 2023). Then, the increase in consumption for employed workers induced by the transfer (yellow rectangle) generates an efficiency cost (green rectangle) that possibly leads to incremental profits. Formally, to generate an increase of Δy_1 in workers' consumption, the government has to increase the transfer in $-dT_1 = \Delta y_1 + (w^* - dw)$.

The planner could tax these incremental profits by increasing the corporate tax. This, however, induces distortions in labor demand that make the transfer program less efficient. As illustrated in Panel (d), increasing the corporate tax from t_0 to t_1 shifts labor demand downwards (red-lined area). Then, the planner induces additional efficiency costs when using the corporate tax rate to correct for the wage effects of the transfer. Importantly, these distortions are present even in the absence of labor supply effects because non-distortionary transfers also have to be funded by costly corporate tax revenue.

A minimum wage can correct the distortions depicted in Panel (c) while avoiding the additional distortions depicted in Panel (d) by preventing wages from decreasing, thus forcing the firm to pay for the efficiency cost of the increased labor supply. The minimum wage, therefore, indirectly works as a tax on profits. However, as it is illustrated in Panels (a) and (b), it does so by imposing its own distortions on employment. Then, the desirability of the minimum wage will depend on the relative size of the labor demand and profit distortions induced by the corporate tax and the minimum wage. The following proposition formalizes this intuition.

Proposition 2. *Consider the allocation induced by the optimal tax system with no minimum wage, (T_0^*, T_1^*, t^*). Adding a binding minimum wage is desirable if:*

$$1 > \frac{t^* N \Pi}{L\overline{w}} \left((1 - g_k)\frac{\eta_w}{\eta_t} + \left(1 + \frac{g_k(1-t^*)}{t^*}\right)\left(\epsilon_w - \epsilon_t \frac{\eta_w}{\eta_t}\right)\right). \tag{7}$$

Proposition 2 states that a minimum wage can be a useful complement to the optimal tax system if

[15] Deviations from efficient rationing possibly imply a non-zero g_u even when $\overline{w} = w^*$. Then, the desirability of the minimum wage under more general rationing schemes probably takes the form of equation (6) even when departing from w^*.

the distortions induced by the minimum wage, evaluated at the allocation induced by the optimal tax system, are smaller than the corresponding distortions induced by the corporate tax.[16]

To understand the intuition of the result, assume that $\omega_k = 0$ and, therefore, $g_k = 0$. In this case, the planner only cares about workers' welfare and, therefore, all decisions around profit taxation are driven by efficiency motives. When $g_k = 0$, equation (7) reduces to:

$$1 > \frac{t^* N \Pi}{L\overline{w}} \left(\frac{\eta_w}{\eta_t}(1-\epsilon_t) + \epsilon_w \right). \tag{8}$$

This expression shows that the desirability of the minimum wage is increasing in ϵ_t and decreasing in ϵ_w. Likewise, if $\epsilon_t < 1$, the desirability of the minimum wage is also increasing in η_t and decreasing in η_w.[17] The planner can tax profits using the minimum wage to relax the efficiency costs of the corporate tax. Since using the minimum wage generates costs on employment and corporate tax revenue, the strategy is only valuable if those distortions are small relative to the efficiency costs imposed by the corporate tax.

To develop more intuition, consider the following polar cases. If the corporate tax is non-distortionary – which in terms of the model implies that $\phi_t = 0$ –, then $\eta_t = \epsilon_t = 0$ and, therefore, the condition depicted in equation (7) is never true. That is, when corporate taxes induce no distortions and the planner implements the optimal value (that could, eventually, reach $t^* = 100\%$ if $\omega_k \to 0$), the minimum wage is always dominated by the optimal tax system. By the contrary, if the employment distortions of the minimum wage are negligible – which in terms of the model can be approximated by $\phi_w = 0$ –, then $\eta_w = 0$ and $\epsilon_w = l(w/\Pi)$ and, therefore, the condition depicted in equation (7) is reduced to $1 > g_k$. Since any distortion in the corporate tax implies that, at the optimum, $1 > g_k$, then complementing the optimal tax system with a minimum wage is desirable if the corporate tax is distortionary. Likewise, no distortions in the corporate tax rate would imply that $g_k = 1$ at the optimum, thus making the minimum wage superfluous under an optimal tax system even if employment effects are negligible.[18]

When are the distortions different? The analysis suggests that the desirability of the minimum wage depends on the relation between (η_w, ϵ_w) and (η_t, ϵ_t). While both sets of distortions are not isomorphic, they could be "correlated": production functions that make the minimum wage distortionary could also make the corporate tax distortionary, and vice versa. Then, comparing both distortions could be

[16]To further develop intuition, Appendix B.1 provides two proofs of Proposition 2. The first is based on a perturbation argument around the initial optimum that does not rely on efficient rationing, similar in spirit to the proof strategy of Kaplow (2006). The second is derived from first principles by directly solving the optimization problem of the planner.

[17]If, at the optimum with $g_k = 0$, $\epsilon_t > 1$, there are mechanic corporate revenue gains from decreasing the corporate tax. If, however, on top of that $T_0^* > T_1^*$ (i.e., there are in-work benefits), the employment increase generates a negative fiscal externality that needs to be compensated with increases in the minimum wage to enforce the budget constraint. In line with that, in Appendix B.1, I show that, at the optimum, $\epsilon_t < 1$ if $T_0^* < T_1^*$.

[18]The redundancy of the minimum wage when t is non-distortionary is based on the assumption that the planner implements t^*. Even in the absence of distortions, if the planner implements a sub-optimal corporate tax rate (because, for example, of political economy constraints), there may be room for using the minimum wage to tax profits.

fallacious if either both are large or small but a case with substantial differences between them is unlikely.

To develop intuition on this concern, I consider a parametric example where $\widetilde{\phi}(l,k) = \psi l^{\alpha} k^{\beta}$, with $\psi > 0$, and $\alpha + \beta = b < 1$, so $1 - b$ is a measure of returns to scale and α and β represent the output elasticities of labor and capital so α/b and β/b measure factor intensities in the production function. In Appendix B.2, I derive closed-form expressions for the relevant elasticities using this parametric structure and find that: $\epsilon_t > \epsilon_w$ if $\beta/\alpha > (1-t^*)/t^*$; $\eta_t > \eta_w$ if $\beta > 1 - t^*$; and equation (8) is reduced to $\beta > (1-t^*)(1-b)$. That is, the minimum wage is more likely to be a useful complement to an optimal tax system when the implied technology has a degree of capital intensity above some threshold.

Intuitively, capital intensity is what makes the corporate tax rate distortionary, since it implies that the capital allocation problem matters more in the determination of profits. This conclusion is consistent with recent empirical literature (Garrett et al., 2020; Curtis et al., 2022; Kennedy et al., 2023). Then, the larger the degree of capital intensity, the larger the size of the corporate tax distortion. On the contrary, the minimum wage distortion works mostly through the increase in labor costs. Then, if labor intensity is small, the minimum wage has nuanced impacts on employment and pre-tax profits.[19]

The minimum wage as an industry-specific profit tax In the US, minimum wage workers are concentrated in labor-intensive industries such as food and accommodation and retail trade (see Table 1). In what follows I argue that, in the presence of firm-level heterogeneity, the minimum wage has the potential to generate additional fiscal benefits to the planner when corporate taxes cannot be designed to be sector-specific, especially when unaffected sectors are particularly capital-intensive.

To understand why, consider the following extension of the model. Workers and firms described above constitute a segmented low-skill labor market of low-skill workers working in firms of a low-skill sector. Denote the low-skill market primitives and equilibrium objects, x, by x^l. Assume a second segmented labor market of high-skill workers working in firms of a high-skill sector – with primitives and equilibrium objects denoted by x^h. The policy objects of the planner consist of (T_0, T_1, T_2, t), where T_1 are net taxes paid by low-skill employed workers and T_2 are net taxes paid by high-skill workers.[20] The key assumption that motivates the subsequent analysis is that the planner is constrained to impose a unique linear corporate tax, t, that affects all firms regardless of their industry.

The intuition is the following. Assume without loss of generality that $w^h > w^l$, so $\overline{w}$ is not binding in the high-skill sector. If the planner wants to tax profits using the corporate tax, it has to choose a tax rate that affects all sectors equally. If the high-skill sector is very capital intensive, the corporate tax

[19]This analysis is incomplete because t^* is also a (possibly non-linear) function of β. The effect of β on t^* is in principle ambiguous since β affects both the distortions and the pre-tax distribution (equity motives). Appendix C.1 provides a numerical example with optimal taxes that is consistent with the intuition above: complementing the optimal tax system with a minimum wage is desirable only when β is larger than a well-defined threshold.

[20]Because of the segmented markets assumption, the planner's problem does not involve cross-skill incentive compatibility constraints because high-skill workers cannot work in the low-skill industry and vice versa.

is particularly distortionary in the high-skill sector. Then, the planner can alleviate distortions in the high-skill sector by taxing part of the low-skill sector's profits with the minimum wage, which allows the planner to marginally decrease the corporate tax at the optimum. That is, the positive fiscal externality on the high-skill sector makes a case for using the minimum wage as a sector-specific corporate tax. The following proposition formalizes this intuition. To simplify exposition, I consider the case where $\omega_k = 0$.

Proposition 3. *Consider the two-sector model with segmented labor markets and let $\omega_k = 0$ and $w^l < w^h$. Consider the allocation induced by the optimal tax system with no minimum wage, $(T_0^*, T_1^*, T_2^*, t^*)$. Adding a binding minimum wage is desirable if:*

$$1 > \frac{t^* N^l \Pi^l}{L^l \overline{w}} \left(\frac{\eta_w^l}{\eta_t^l}(1-\epsilon_t^l) + \epsilon_w^l \right) + \frac{1}{L^l \overline{w}} \frac{\eta_w^l}{\eta_t^l} \left(t^* N^h \Pi^h (1-\epsilon_t^h) - L^h \eta_t^h \Delta T_2^* \right), \tag{9}$$

where $\Delta T_2^* = T_2^* - T_0^*$.

The first part of equation (9) coincides with equation (8), so it replicates the argument developed above. The second part describes the fiscal externality of the high-skill sector. The benefit of using the minimum wage is increasing in ϵ_t^h and, if $\Delta T_2^* > 0$ (which corresponds to a situation in which high-skill workers pay more taxes than the non-employed), it is also increasing in η_t^h. That is, the more distortionary the corporate tax is in the high-skill sector (which is a function of its capital intensity), the more likely that complementing the optimal tax system with a minimum wage is socially desirable since the minimum wage allows the planner to relax the corporate tax distortions on non-exposed industries.

This result suggests that, in a case like the US where minimum wage workers are concentrated in industries with relatively low capital intensity such as food and accommodation and retail trade, a binding minimum wage can help to alleviate corporate tax distortions in unaffected industries such as manufacturing. Intuitively, the minimum wage plays the role of a sector-specific corporate tax that helps the government minimize capital distortions in the sectors that are more affected by capital mobility.

4 Optimal minimum wage in a more general model of the labor market

The analysis so far suggests that the desirability of the minimum wage depends on the distortions of the corporate tax. One caveat of the analysis above is that it is based on a model of the labor market that oversimplifies the effects of the minimum wage on the labor market. In the real world, labor markets have frictions that mediate the wage and employment effects of the minimum wage. Moreover, there is substantial firm-level heterogeneity in terms of wages, profits, and exposure to minimum wage policies, and there may be entry and exit of firms as a response to taxes and labor market regulations.

This section extends the analysis to a more general, non-neoclassical, model of the labor market with positive profits that can accommodate limited employment effects and spillovers to non-minimum wage jobs after minimum wage increases. The decentralized equilibrium of the proposed model remains (constrained) efficient despite the frictions. Hence, the analysis in this section also abstracts from efficiency rationales and maintains the focus on the redistributive properties of the minimum wage.

4.1 Model of the labor market

The model is static and features two populations. On one side, there is a population of workers that is heterogeneous in two dimensions: skills and costs of participating in the labor market. For simplicity, I assume workers are either low-skill or high-skill. On the other side, there is a population of capitalists with heterogeneous productivities and technologies. 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*.[21] 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 possessing 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 can be written as:

$$p_m^s = \frac{\mathcal{M}^s(L_m^s, V_m^s)}{L_m^s} = \mathcal{M}^s(1, \theta_m^s) \equiv p^s(\theta_m^s), \tag{10}$$

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 can be written as:

$$q_m^s = \frac{\mathcal{M}^s(L_m^s, V_m^s)}{V_m^s} = \mathcal{M}^s\left(\frac{1}{\theta_m^s}, 1\right) \equiv q^s(\theta_m^s), \tag{11}$$

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. Neither workers nor firms internalize that their behavior affects equilibrium tightness, so they take p_m^s and q_m^s as given when making their decisions.

[21]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.

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, with $\alpha_l + \alpha_h = 1$. As in Section 3, 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 and is distributed with conditional cdf F_s and pdf f_s.

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

$$u_1(s,c) \quad = \quad \max_m \{p_m^s y_m^s + (1 - p_m^s) y_0\} - c, \tag{12}$$

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 $1 - p_m^s$.

Individuals take p_m^s as given but it is endogenously determined by the aggregate application behavior. This implies that, in equilibrium, all markets yield the same expected utility, i.e., $p_{i_1}^s y_{i_1}^s + (1 - p_{i_1}^s) y_0 = p_{i_2}^s y_{i_2}^s + (1 - p_{i_2}^s) y_0 = \max_m \{p_m^s y_m^s + (1 - p_m^s) y_0\}$, for all pairs (i_1, i_2); 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 tradeoff between wages and employment probabilities because it is more difficult to get a job in sub-markets that pay higher wages.

In what follows, I define $U^s \equiv \max_m \{p_m^s y_m^s + (1 - p_m^s) y_0\}$ so $u_1(s, c) = U^s - c$. Workers participate in the labor market if $1\{u_1(s, c) \geq u_0\} = 1\{U^s - y_0 \geq c\}$, which implies that the mass of active workers of skill s is given by $L_A^s = \alpha_s F_s(U^s - y_0)$. 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 of 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 $\partial \theta_m^s / \partial U^s > 0$.[22] This result simplifies the analysis below since it implies that, conditional on wages, equilibrium behavior can be summarized by the scalars U^s without needing to characterize the continuous sequence of θ_m^s.

[22]Since $U^s = p^s(\theta_m^s)(w_m^s - T(w_m^s)) + (1 - p^s(\theta_m^s)) y_0$, then $dU^s = p_\theta^s d\theta_m^s y_m^s + p_m^s (1 - T'(w_m^s)) dw_m^s$. Recalling that $p_\theta^s > 0$ and asssuming $T'(w_m^s) < 1$ yields the result.

Capitalists Instead of assuming a fixed (discrete) number of capitalists, I now consider a continuous population normalized to K. Each capitalist draws a parameter $\psi \in \Psi = \left[\underline{\psi}, \overline{\psi}\right] \subset \mathbb{R}^+$ that represents firm productivity. Capitalists also draw a technology $j \in \mathcal{J} = \{1, 2, ..., J\}$, that may induce, for example, different firm-level factor shares and returns to scale. Let o_j and O_j be the conditional pdf and cdf of ψ given j, respectively. The pmf of j is denoted by σ_j, with $\sum_{j \in \mathcal{J}} \sigma_j = 1$.

Capitalists observe (ψ, j) 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 corporate tax rate, t, so a firm of type (ψ, j) that hires (n^l, n^h) workers generates revenue equal to $\phi^j(\psi, n^l, n^h, t)$, with ϕ^j twice differentiable, $\phi^j_\psi > 0$, $\phi^j_{n^s} > 0$ and $\phi^j_{n^s n^s} \leq 0$, $\phi^j_t \leq 0$, and $\phi^j_{tn^s} \leq 0$, for $s \in \{l, h\}$. As in Section 3, allowing the revenue function to depend on t accommodates, in a reduced-form fashion, the fact that corporate tax rates can distort pre-tax profits.[23]

Firms choose skill-specific wages, w^s, and vacancies, v^s. While firms take the job-filling probabilities as given, they internalize that paying higher wages increases the job-filling probabilities. Using the workers' equilibrium condition, I rewrite job-filling probabilities as $\tilde{q}^s(w^s, U^s) = q(\theta^s(w^s, U^s))$, with $\tilde{q}^s_w = q^s_\theta(\partial\theta^s/\partial w^s) > 0$, so $n^s = \tilde{q}^s(w^s, U^s)v^s$. 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 of a capitalist of type (ψ, j) are given by revenue net of labor costs:

$$\begin{aligned}\pi^j\left(w^l, w^h, v^l, v^h; \psi, t\right) \quad = \quad & \phi^j\left(\psi, \tilde{q}^l(w^l, U^l)v^l, \tilde{q}^h(w^h, U^h)v^h, t\right) \\ & - \left(w^l\tilde{q}^l(w^l, U^l)v^l + \eta^l(v^l)\right) - \left(w^h\tilde{q}^h(w^h, U^h)v^h + \eta^h(v^h)\right). \end{aligned} \tag{13}$$

Denote the value function by $\Pi^j(\psi, t) = \max_{w^l, w^h, v^l, v^h} \pi^j\left(w^l, w^h, v^l, v^h; \psi, t\right)$. Then, after-tax profits are given by $(1-t)\Pi^j(\psi, t)$. 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 $(1-t)\Pi^j(\psi, t) \geq \xi + y_0$. Since, conditional on j, profits are increasing in productivity, the entry rule defines a j-specific productivity threshold, ψ^*_j, implicitly determined by $(1-t)\Pi^j(\psi^*_j, t) = \xi + y_0$ such that j-type capitalists create firms only if $\psi \geq \psi^*_j$. Then, the mass of active capitalists is given by $K_A = K\sum_{j \in \mathcal{J}} \sigma_j \left(1 - O_j(\psi^*_j)\right)$. Likewise, the mass of inactive capitalists, K_I, is given by $K_I = K\sum_{j \in \mathcal{J}} \sigma_j O_j(\psi^*_j)$, with $K_A + K_I = K$.

Conditional on (ψ, j), firms are homogeneous. Then, the solution to the firm's problem can be characterized by functions $w^s(\psi, j)$ and $v^s(\psi, j)$.[24] Then, m indexes sub-markets as well as the (ψ, j) values of capitalists that create firms, so $w^s_m = w^s\left(\widetilde{\psi}, \widetilde{j}\right)$, $v^s_m = v^s\left(\widetilde{\psi}, \widetilde{j}\right)$, and $V^s_m = Kv^s\left(\widetilde{\psi}, \widetilde{j}\right)\sigma_{\widetilde{j}}o_{\widetilde{j}}\left(\widetilde{\psi}\right)$, for some $\left(\widetilde{\psi}, \widetilde{j}\right) \in \left\{(\psi, j) \in \Psi \times \mathcal{J} : \psi \geq \psi^*_j\right\}$.[25]

[23] In Appendix B.3 I adapt the capital allocation problem discussed in Section 3 where capitalists that own capital allocate it between a domestic firm and a foreign investment to the model presented in this section.

[24] Appendix B.3 derives the first-order conditions and shows that firm-heterogeneity leads to dispersion in wages conditional on skill, with wages *marked down* relative to the marginal productivities. The wage markdown is governed by the wage-dependent job-filling probabilities – which emulate firm-specific labor supply elasticities – and the vacancy creation costs.

[25] There could be more than one (ψ, j) pair yielding the same w^s_m. In those cases, firms' FOCs imply that they will

Equilibrium The notion of equilibrium in the labor market is formally described in Appendix B.3. The equilibrium objects are U^l, U^h, $\left\{\psi_j^*\right\}_{j=1}^{J}$, and the sub-market skill-specific wages, vacancies, and applicants, (w_m^s, v_m^s, L_m^s), for all m and s, and solve the firms' first order conditions, the worker- and firm-level participation constraints, and the across sub-market equilibrium condition on workers' applications.

Discussion Before introducing a minimum wage to the model, I highlight some relevant properties of the proposed framework. Some caveats of the model are discussed in Section 5.

- *Directed search:* Directed search models generate efficient outcomes in terms of search and posting behavior (Moen, 1997; Wright et al., 2021). That is, these models don't exhibit inefficient mixes of applicants and vacancies as can happen in random search models (e.g., Hosios, 1990; Mangin and Julien, 2021). In Appendix B.3, I show that the proposed model maintains this property.[26]

- *Monopsony power:* While search and posting behavior is efficient, the model admits monopsony power through wage-dependent job-filling probabilities that have a similar spirit to the standard monopsony intuition of upward-slopping firm-specific labor supply curves (Robinson, 1933; Card et al., 2018). In the model, these elasticities are endogenous since firms internalize that paying higher wages leads to more applicants, so wages are *marked down* relative to marginal productivities.[27]

- *Rationing:* Different from the model presented in Section 3, this framework does not impose efficient rationing. On the contrary, the allocation of applicants to jobs is independent of c (conditional on participation), which looks closer to a uniform rationing assumption. However, the implicit rationing assumption is second-order for the welfare analysis below since the social planner will maximize the sum of expected utilities, which are ex-ante equal for all workers within a skill type.

- *Low-wage labor markets:* The equilibrium of the model is also 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).

4.2 Introducing a minimum wage

I introduce a minimum wage, $\overline{w}$, to explore how the model predictions speak to the related literature.

also post the same vacancies since the functions q^s and η^s do not vary with j (see Appendix B.3). For those cases, define $\mathcal{I}(s,m) \subset \left\{(\psi, j) \in \Psi \times \mathcal{J} : \psi \geq \psi_j^*\right\}$ as the set of combinations that optimally post the same wage w for workers of skill s, w_m^s. Let ι index elements in $\mathcal{I}(s,m)$. Then, $V_m^s = K \sum_{\iota \in \mathcal{I}(s,m)} v^s\left(\widetilde{\psi}_\iota, \widetilde{j}_\iota\right) \sigma_{\widetilde{j}_\iota} o_{\widetilde{j}_\iota}\left(\widetilde{\psi}_\iota\right)$.

[26] 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.

[27] Appendix B.3 shows that the standard markdown equation can be derived from the firm's first-order conditions.

Workers Consider first the sub-population of low-skill workers. In equilibrium, $U^l = p^l(\theta^l_m)y^l_m + (1 - p^l(\theta^l_m))y_0$, for all sub-markets m. Let i_1 be the sub-market for which the minimum wage is binding, so $w^l_{i_1} = \overline{w}$. Differentiating yields:

$$\frac{dU^l}{d\overline{w}} \quad = \quad p^l_\theta \frac{d\theta^l_{i_1}}{d\overline{w}}(\overline{w} - T(\overline{w}) - y_0) + p^l(\theta^l_{i_1})(1 - T'(\overline{w})). \tag{14}$$

Since $p^s(\theta^l_{i_1}) > 0$, and assuming $T'(\overline{w}) < 1$, $dU^l/d\overline{w} = d\theta^l_{i_1}/d\overline{w} = 0$ is not a feasible solution to equation (14). This implies that changes in $\overline{w}$ necessarily affect the equilibrium values of U^l, $\theta^l_{i_1}$, or both.

An increase in the minimum wage mechanically makes minimum wage jobs more attractive for low-skill workers. This effect is captured by the second term in the right-hand-side of equation (14). The increased attractiveness attracts new applicants toward minimum-wage sub-markets thus pushing $\theta^l_{i_1}$ downwards until the across sub-market equilibrium is restored. This decreases the employment probability, as captured by the first term in the right-hand-side of equation (14). The overall effect is ambiguous, depending on whether the wage or the employment effects dominate.

Changes in $\overline{w}$ also affect low-skill sub-markets for which the minimum wage is not binding. Let i_2 be a sub-market for which the minimum wage is not binding, so $w^l_{i_2} > \overline{w}$. Differentiating yields:

$$\frac{dU^l}{d\overline{w}} \quad = \quad p^l_\theta \frac{d\theta^l_{i_2}}{d\overline{w}}(w^l_{i_2} - T(w^l_{i_2}) - y_0) + p^l(\theta^l_{i_2})(1 - T'(w^l_{i_2}))\frac{dw^l_{i_2}}{d\overline{w}}. \tag{15}$$

Equation (14) suggests that the left-hand-side of equation (15) is unlikely to be zero, implying that $\theta^l_{i_2}$ or $w^l_{i_2}$ 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 effect is captured by the first term in the right-hand-side of equation (15). Second, firms can also respond to changes in applicants. The potential wage response is captured in the second term in the right-hand-side of equation (15). Changes in vacancy posting implicitly enter the terms $d\theta^l_m/d\overline{w}$ of equations (14) and (15). The overall effect is also ambiguous.

Changes in U^l also affect labor market participation. Recall that $L^l_A = \alpha_l F_l(U^l - y_0)$, so $dL^l_A/d\overline{w} = \alpha_l f_l(U^l - y_0)\left(dU^l/d\overline{w}\right)$. Then, if $dU^l/d\overline{w} > 0$, minimum wage hikes increase labor market participation. The behavioral response, however, is scaled by $f_l(U^l)$, which may be negligible. This may result in positive impacts on expected utilities with modest participation effects at the aggregate level.

Now consider the sub-population of high-skill workers. If $\min_m\{w^h_m\} > \overline{w}$, equilibrium effects for high-skill workers take the form of equation (15). 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^j. 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 Appendix B.3 provides expressions for the effects of minimum wage changes on firms' outcomes. 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 taking 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 an 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 effect increases the expected value of posting a low-skill vacancy and helps to attenuate potential disemployment effects driven by increased labor costs.

Firms for which the minimum wage does not bind also react by adapting their posted wages and vacancies to changes in their sub-market tightness. The analytical expression for the wage spillover has an ambiguous sign but directly depends on the change in sub-market tightness. Since wages and vacancies are positively correlated at the firm and skill level, firms also change their posted vacancies, thus generating potential employment spillovers.

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 Recent evidence suggests that minimum wage hikes generate positive wage effects with "elusive" disemployment effects (Manning, 2021a). The model can rationalize positive wage effects with limited employment and participation effects through the equilibrium changes in applications: labor costs are attenuated given the increase in the mass of applicants. The empirical literature also finds that minimum wages generate spillovers to non-minimum wage jobs, both within and between firms (Cengiz et al., 2019; Derenoncourt et al., 2022; Dustmann et al., 2022; Engbom and Moser, 2022; Giupponi and Machin, 2023; Vogel, 2023; Forsythe, Forthcoming). The same application responses 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. Finally, the model can also accommodate negative effects on firm profits, as documented by Draca et al. (2011), Harasztosi and Lindner (2019), Drucker et al. (2021), and Section 2.

4.3 Planner's problem

As in Section 3, I assume that the planner does not observe (c, ψ, j) and, therefore, constrains the policy choice to second-best incentive-compatible policy schemes that depend on earnings and profits. Following related literature that includes matching frictions (Kroft et al., 2020; Lavecchia, 2020), the planner is assumed to maximize a (generalized) utilitarian SWF based on expected utilities:

$$\begin{aligned} SWF &= \left(L_I^l + L_I^h + K_I\right)\omega_L G(y_0) + \alpha_l \int_0^{U^l - y_0} \omega_L G(U^l - c) dF_l(c) \\ &\quad + \alpha_h \int_0^{U^h - y_0} \omega_L G(U^h - c) dF_h(c) + K \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \omega_K G\left((1-t)\Pi^j(\psi, t) - \xi\right) dO_j(\psi), \end{aligned} \quad (16)$$

where $(\overline{w}, T, t)$ are the policy parameters and G and (ω_L, ω_K) are defined as in Section 3. 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_j^*. The first term of equation (16) accounts for the utility of inactive workers and inactive capitalists. The second and third terms of equation (16) 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.[28]

Likewise, the natural extension of the budget constraint defined in equation (4) is given by:

$$\begin{aligned} \left(L_I^l + L_I^h + K_I + \rho^l L_A^l + \rho^h L_A^h\right) y_0 &= \int \left(E_m^l T(w_m^l) + E_m^h T(w_m^h)\right) dm \\ &\quad + tK \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \Pi^j(\psi, t) dO_j(\psi), \end{aligned} \quad (17)$$

where $E_m^s = p_m^s L_m^s$ is the mass of employed workers of skill s in sub-market m and ρ^s is the skill-specific unemployment rate given by $\left(L_A^s - \int E_m^s dm\right)/L_A^s$. Equation (17) establishes that the transfer paid to individuals with no market income must be funded by the tax collection on employed workers and active capitalists. Finally, as in Section 3, if γ is the budget constraint multiplier, the SMWWs of inactive workers, active workers of skill type s, and active capitalists of type (ψ, j) are defined as $g_0 = \omega_L G'(y_0)/\gamma$, $g_1^s = \alpha_s \omega_L \int_0^{U^s - y_0} G'(U^s - c) dF_s(c)/\gamma L_A^s$, and $g_\psi^j = \omega_K G'((1-t)\Pi^j(\psi, t) - \xi)/\gamma$, respectively.

4.4 Optimal minimum wage with no tax system

As in Section 3, I start developing intuition with the simple case with no tax system.

Proposition 4. *In the absence of taxes, increasing the minimum wage is welfare-improving if:*

$$\frac{dU^l}{d\overline{w}} L_A^l g_1^l + \frac{dU^h}{d\overline{w}} L_A^h g_1^h + K \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} g_\psi^j \frac{d\Pi^j(\psi)}{d\overline{w}} dO_j(\psi) \quad > \quad 0. \quad (18)$$

Proposition 4 generalizes Proposition 1 by illustrating that a small increase in the minimum wage can affect the relative welfare of active low-skill workers (first term), active high-skill workers (second

[28] Appendix B.3 provides further intuition by relating equation (16) with the average welfare by group.

term), and active capitalists (third term). Depending on the change in utility for the different groups ($dU^s/d\overline{w}$ and $d\Pi^j(\psi)/d\overline{w}$), the social value of those changes (g_1^s and g_ψ^j), and the size of the groups (L_A^s and $K\sum_{j\in\mathcal{J}}\sigma_j o_j(\psi)$), increasing the minimum wage may be desirable or not for the social planner.

There are three main differences between Proposition 1 and Proposition 4. First, Proposition 4 includes high-skill workers in the distributional equation, which allows the minimum wage to affect the relative welfare between low- and high-skill workers on top of affecting the relative welfare between workers and capitalists. More generally, the model in Section 3 displays no wage dispersion. In the current analysis, heterogeneous impacts by pre-reform wage levels matter for the welfare consequences of the minimum wage. Second, the inclusion of firm-level heterogeneity delivers heterogenous SMWWs within capitalists, which implies that identifying who are the firm owners exposed to minimum wage changes also matters for the welfare assessment of the policy. The differential exposure to minimum wage workers is an empirical question that possibly follows from the joint distribution of ψ and j.

Third, and more importantly, the worker-level welfare impacts are measured at the skill level: what matters for assessing the welfare impacts of the minimum wage are the average effects on the expected utility of low-skill workers, which comprise workers employed at different wage rates (including the minimum wage) and also unemployed workers. Formally, recall that, in the absence of taxes, $U^s = p_m w_m$. Multiplying both sides by the sub-market mass of applicants, L_m^s, and integrating over m, yields:

$$U^s = \frac{\int E_m^s w_m^s dm}{L_A^s} = (1-\rho^s)\mathbb{E}_m[w_m^s] + \rho^s \cdot 0 = (1-\rho^s)\mathbb{E}_m[w_m^s], \tag{19}$$

where $\mathbb{E}_m[w_m^s] = \int \nu_m^s w_m^s dm$ is the average wage of employed workers, with $\nu_m^s = E_m^s / \int E_m^s dm$ and $\int \nu_m^s dm = 1$. Then, U^s is equal to the average wage of active workers including the unemployed.[29] $dU^s/d\overline{w}$ captures all general equilibrium effects that affect workers' utility, including direct and spillover effects on wages, employment, and participation. In the public debate, there is an unsettled discussion about the appropriate way of weighting these different effects. The proposed framework offers an avenue for aggregating all effects into a single elasticity, given the focus on expected utilities.[30] Note that rationing assumptions do not affect the desirability of the minimum wage because the ex-ante rent measure from labor market participation internalizes the risk of unemployment.

4.5 Incorporating fiscal externalities

Before proceeding to the analysis where the planner jointly chooses taxes and the minimum wage, I explore the minimum wage desirability under fixed taxes. The idea is to extend Proposition 4 to incorporate

[29] In the case with taxes, Appendix B.3 shows that U^s is equal to the average pre-tax wage of active workers including the unemployed net of their average tax liabilities.

[30] 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 B.3 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.

fiscal externalities. When unmodeled constraints restrict the scope for simultaneous tax reforms, this may be the policy-relevant scenario for assessing the desirability of the minimum wage.

Proposition 5. *If taxes are fixed, increasing the minimum wage is welfare-improving if:*

$$\begin{aligned}
&\frac{dU^l}{d\overline{w}}L_A^l g_1^l + \frac{dU^h}{d\overline{w}}L_A^h g_1^h + K(1-t)\sum_{j\in\mathcal{J}}\sigma_j\int_{\psi_j^*}^{\overline{\psi}} g_\psi^j \frac{d\Pi^j(\psi,t)}{d\overline{w}}dO_j(\psi)\\
&+\int\left(\frac{dE_m^l}{d\overline{w}}\Big(T(w_m^l)+y_0\Big)+E_m^l T'(w_m^l)\frac{dw_m^l}{d\overline{w}}\right)dm\\
&+\int\left(\frac{dE_m^h}{d\overline{w}}\Big(T(w_m^h)+y_0\Big)+E_m^h T'(w_m^h)\frac{dw_m^h}{d\overline{w}}\right)dm\\
&+tK\sum_{j\in\mathcal{J}}\sigma_j\int_{\psi_j^*}^{\overline{\psi}}\frac{d\Pi^j(\psi,t)}{d\overline{w}}dO_j(\psi)-\frac{dK_I}{d\overline{w}}\left(t\Pi^{marg}+y_0\right) \quad > \quad 0,
\end{aligned} \tag{20}$$

where $\Pi^{marg} = \Pi^j(\psi_j^*, t) = (\xi + y_0)/(1-t)$ *is the pre-tax profit of the marginal capitalists.*

The first line in equation (20) reproduces the same welfare tradeoff described in Proposition 3. The second to fourth lines in equation (20) summarize the fiscal externalities on both sides of the market.

The second line describes the fiscal externalities on low-skill labor markets. The first term shows that, if low-skill employment changes, there is a substitution between tax collection from employed workers, $T(w_m^l)$, and transfers paid to unemployed individuals, y_0. 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. 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. Then, firm-level fiscal externalities seem particularly relevant in the current state of international tax competition. 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.

Sufficient statistics One appealing feature of Proposition 5 is its suitability for empirical assessments. Group sizes are observed in the data, welfare weights can be imputed (or calibrated based on observed post-tax incomes), and welfare changes and fiscal externalities can be estimated in a "sufficient statistics" fashion (Chetty, 2009; Kleven, 2021). In particular, the terms U^s can be constructed using data in wages, employment, tax liabilities, and participation, as suggested by equations (19) and Appendix B.3. In fact, the worker-level results presented in Figure 1 (Panel (a)) correspond to estimates of a pre-tax version of $dU^l/d\overline{w}$. Likewise, results on income maintenance benefits (Panel (b) of Figure 1) proxy for worker-level fiscal externalities and results on profits per establishment (Panel (a) of Figure 2) can be used to estimate aggregate proxies of $d\Pi^j(\psi,t)/d\overline{w}$ and its corresponding fiscal externalities.

Appendix D provides an example of this calibration exercise. I use the reduced-form results presented in Section 2 to compute, under different assumptions, the minimum welfare weight on low-skill workers that would justify increasing the minimum wage, which essentially consists of finding conditions on g_1^l such that equation (20) is positive. The analysis makes clear the importance of the distributional dimension when assessing the desirability of the minimum wage. Total gains for low-skill workers are of similar magnitude to total losses for exposed capitalists, and the net fiscal externality (fiscal savings in transfers minus fiscal losses in corporate tax revenue) is close to zero. Then, in the absence of preferences for redistribution, the policy is not far away from breaking even. However, when preferences for redistribution enter the analysis, the change in profits only affects the fiscal externality but plays a negligible role in the welfare effect because average post-tax profits are several times larger than average post-tax incomes of low-skill workers and, therefore, there are substantial equity gains from redistributing profits to workers. In this scenario, increasing the minimum wage is unambiguously desirable because the distinction between winners and losers is aligned with the planner's preferences.

4.6 Optimal minimum wage with optimal taxes

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 $\max_i w_i^l < \min_i w_i^h$, or that the social planner can implement skill-specific income tax schedules.[31]

Proposition 6. *If taxes are optimal, increasing the minimum wage is welfare-improving if:*

$$\begin{aligned}
\frac{\partial U^l}{\partial \overline{w}} L_A^l g_1^l + \frac{\partial U^h}{\partial \overline{w}} L_A^h g_1^h + K(1-t) \sum_{j\in\mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} g_\psi^j \frac{\partial \Pi^j(\psi,t)}{\partial \overline{w}} dO_j(\psi) \\
+ \int \left(\frac{\partial E_m^l}{\partial \overline{w}} \left(T(w_m^l) + y_0 \right) + E_m^l \frac{\partial w_m^l}{\partial \overline{w}} \right) dm \\
+ \int \left(\frac{\partial E_m^h}{\partial \overline{w}} \left(T(w_m^h) + y_0 \right) + E_m^h \frac{\partial w_m^h}{\partial \overline{w}} \right) dm \\
+tK \sum_{j\in\mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \frac{\partial \Pi^j(\psi,t)}{\partial \overline{w}} dO_j(\psi) - \frac{\partial K_I}{\partial \overline{w}} \left(t\Pi^{marg} + y_0 \right) \quad > 0 \quad .
\end{aligned} \tag{21}$$

where $\Pi^{marg} = \Pi^j(\psi_j^*, t) = (\xi + y_0)/(1-t)$ *is the pre-tax profit of the marginal capitalists. 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 (profit-weighted) average SMMW among active capitalists is below 1.*

[31] These assumptions allow me to solve the planner's problem by pointwise maximization. Note that these assumptions increase the relative attractiveness of the tax system. In fact, skill-specific income taxes mimic the instruments considered in Diamond and Mirrlees (1971a,b), which are more flexible than the ones considered, for example, in Naito (1999).

Proposition 6 extends Proposition 2 to the non-neoclassical model of the labor market. Note that Proposition 6 looks similar to Proposition 5. However, when taxes are optimized together with the minimum wage, how the minimum wage affects welfare and generates fiscal externalities changes. This is reflected in three important differences between equations (20) and (21).

First, all relevant elasticities are *micro* rather than *macro* elasticities (Scheuer and Werning, 2017; Landais et al., 2018b,a; Kroft et al., 2020; Lavecchia, 2020), which I denote by partial (rather than total) derivatives.[32] Macro elasticities (Propositions 4 and 5) internalize all general equilibrium effects of the minimum wage, while micro elasticities (Proposition 6) are partial equilibrium elasticities that mute some of these effects because, at the joint optimum, the minimum wage moves in tandem with taxes. To see why, recall that $U^s = p^s_m y^s_m + (1 - p^s_m)y_0 \equiv p^s_m \Delta y^s_m + y_0$, with $\Delta y^s_m = y^s_m - y_0$. When taxes are fixed, both Δy^s_m and p^s_m react to minimum wage changes. However, at the joint optimum, an increase in the minimum wage is paired with changes in taxes for low-skill workers, possibly leaving consumption fixed. Then, the minimum wage directly affects workers' welfare mainly through changes in the employment probabilities driven by changes in vacancy posting.[33] This logic also applies to employment and profits.

Second, optimal taxes affect the levels of the SMWWs. If, relative to a baseline tax system, redistribution from capitalists to workers is socially desirable, the optimal tax system in the background will possibly transfer resources between groups in the desired direction, thus pushing g^l_1 down and g^j_ψ up. Then, the equity motives from using the minimum wage are reduced. Note, however, that they are not completely eroded since the distortions of the tax system put limits on the redistributive possibilities. In particular, corporate tax distortions prevent the corporate tax system from fully redistributing profits, so the average SMWW of active capitalists is smaller than 1.

Third, the effect of the minimum wage on the distribution of pre-tax wages relaxes the government budget constraint by allowing the planner to change the tax liabilities while leaving consumption constant. In simple words, the effects on the pre-tax wage distribution generate fiscal gains for the planner because they switch the burden of redistribution from the government to firms. That is, under optimal taxes, the planner can use the minimum wage as an indirect tax on profits. This effect is captured by the terms $E^s_m(\partial w^s_m/\partial \overline{w})$ which, for the minimum wage sub-market, is equal to low-skill employment ($\partial w^l_m/\partial \overline{w} = 1$). To develop intuition, consider a marginal increase in the subsidy to low-skill workers. As in Panels (c) and (d) of Figure 3, this reform increases labor supply, so firms react by lowering pre-tax wages (Rothstein, 2010; Zurla, 2021; Gravoueille, 2023). The minimum wage mutes this behavioral response, making the

[32] This distinction between micro and macro differs from the one debated in the labor supply elasticity literature, which relates to notions of aggregation and time horizon (e.g., Chetty et al., 2011; Keane and Rogerson, 2015; Kleven et al., 2023).

[33] The direct welfare effects on workers are proportional to the (presumably negative) employment effects. If $U^s = p^s_m \Delta y^s_m + y_0$, multiplying by L^s_m and integrating over m yields $(U^s - y_0)L^s_A = \int E^s_m \Delta y^s_m dm$. Then, if Δy^s_m is fixed:

$$\frac{\partial U^s}{\partial \overline{w}} \left(L^s_A + (U^s - y_0) f^s(U^s - y_0)\right) = \int \frac{\partial E^s_m}{\partial \overline{w}} \Delta y^s_m dm. \tag{22}$$

transfer to low-skill workers less costly. This reform cannot be exactly mimicked by the corporate tax since it distorts pre-tax profits: at the optimum, the average SMWW of active capitalists is below 1.[34]

On a high level, this intuition coincides with Proposition 2. A non-linear tax system is, in isolation, more efficient for redistribution than a minimum wage, especially in the presence of wage heterogeneity. However, complementing the optimal tax scheme with a minimum wage can make tax-based redistribution more efficient when corporate taxes are distortionary through the generation of positive fiscal externalities: it forces firms to pay for part of the transfers to workers, allowing decreases in the corporate tax at the cost of employment effects. The desirability of the minimum wage at the joint optimum, then, depends on how these two forces balance; the assessment of equation (21) is ultimately a quantitative question.

4.7 Restricted model and applications

The number of objects that are affected by general equilibrium forces limits the analytical insights that can be obtained from Proposition 6. To make progress in this regard and benchmark the results with the conclusions of Section 3, I consider a restricted version of the model with limited heterogeneity that allows for greater analytical tractability in the non-neoclassical model presented in this section.

The restrictions are summarized in Assumption 1 below. I consider two fixed populations of inframarginal capitalists, $j \in \{S, M\}$, where firms of type S only hire low-skill workers and firms of type M only hire high-skill workers.[35] Conditional on j firms are homogeneous although ψ can vary between j. The idea is to represent firms in low-skill "services" ($j = S$) and high-skill "manufacturing" ($j = M$), that possibly vary in both their exposure to minimum wages and their capital intensity and mobility.[36]

Assumption 1. *There are two fixed populations of inframarginal capitalists indexed by $j = \{S, M\}$ with sizes $K_j = K\sigma_j$, where capitalists of type $j = S$ – "services" – only employ low-skill workers and capitalists of type $j = M$ – "manufacturing" – only employ high-skill workers. Their respective production functions are given by $\phi^S(\psi^S, n^l, t)$ and $\phi^M(\psi^M, n^h, t)$, with $\{\psi^S, \psi^M\}$ fixed, $\widetilde{\phi}^j_{nk} > 0$, and $\phi^j_{nn} \approx 0$, for $j \in \{S, M\}$. Their respective SMWWs are denoted by g^S_K and g^M_K.*

With the restricted model, I consider two applications. First, I explore conditions under which it is optimal to have a binding minimum wage that is complemented by a tax-based in-work benefit to low-skill workers (i.e., a negative marginal tax rate like the EITC). This analysis is summarized in the following proposition.

[34] 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.

[35] The abstraction from the firm entry margin is motivated by the evidence presented in Table 2.

[36] The restricted version of the model also shuts down second-order effects in the production function. This restriction is made only for analytical simplicity and, if anything, works against the desirability of the minimum wage: if the minimum wage affects employment, $\phi^k_{nn} < 0$ helps firms because the marginal product of labor increases, thus attenuating the employment effect. With matching frictions, decreasing returns to scale are not needed for generating profits.

Proposition 7. *Let Assumption 1 hold. Consider the allocation induced by the optimal tax system with no minimum wage. Let $\varepsilon^l_{\theta,\overline{w}}$ denote the elasticity of low-skill labor market tightness with respect to changes in the minimum wage when after-tax allocations are fixed.*

(i) If $\varepsilon^l_{\theta,\overline{w}} \to 0$ when $\overline{w}$ is set at the market level, having a binding minimum wage is optimal if $g^S_K < 1$.

(ii) Under the optimal binding minimum wage, the optimal marginal tax rate on employed low-skill workers is negative if:

$$g^l_1 \quad > \quad \frac{1 - C\varepsilon^l_{\theta,\Delta}\left[(1-t)g^S_K + t\right]}{1 - B\varepsilon^l_{\theta,\Delta}}, \tag{23}$$

where $B \in (0,1)$, $C \in (0,1)$, and $\varepsilon^l_{\theta,\Delta}$ is (the absolute value of) the elasticity of low-skill labor market tightness with respect to changes in low-skill net-of-tax wage when the minimum wage is fixed.

Recall from Proposition 2 that, in the competitive model, when employment effects are negligible, a binding minimum wage under optimal taxes is desirable when corporate taxes are distortionary. The first part of Proposition 7 reproduces the exact same intuition in the non-neoclassical setting: when the distortions on vacancy posting are negligible (which, in turn, imply negligible distortions on labor market tightness and, therefore, employment), the binding minimum wage is desirable if the SMWW of exposed capitalists, g^S_K, is smaller than one, which is true whenever the corporate tax is distortionary. When vacancy distortions are not negligible, the fiscal externality has to compensate for both the decrease in profits and the overall effects of the congestion externality, implicitly requiring an even smaller SMWW on affected capitalists (or, alternatively, an even larger corporate tax distortion).

The second part of the proposition provides conditions under which the optimal minimum wage is complemented by an in-work transfer (i.e., a negative marginal tax rate like the EITC) to employed low-skill workers. The condition specifies a threshold on g^l_1 that depends on the effects of the transfer on labor market tightness.[37] Since the EITC generates an increase in labor supply and wages are fixed, firms react by decreasing vacancies, thus generating a congestion externality. This effect is captured by $\varepsilon^l_{\theta,\Delta}$ which, as illustrated in the denominator, generates a market-level inefficiency that makes the critical g^l_1 higher. The congestion effect also generates an increase in profits that may relax the critical g^l_1. If the planner values redistribution toward firms, the increase in after-tax profits is socially valuable. Even if $g^S_K = 0$, the increase in profits generates additional corporate tax revenue that is valued by the planner.

The second application concerns the interaction between the minimum wage and the corporate tax rate when the corporate tax distortions differ between firms. If services and manufacturing not only differ in their exposure to $\overline{w}$ but also in their capital intensity and capital mobility, there can be additional policy

[37] If $\varepsilon^l_{\theta,\Delta} \to 0$, equation (23) is reduced to $g^l_1 > 1$, which is the standard result on the EITC desirability in frictionless labor markets with multiple skill types and extensive margin responses (Lee and Saez, 2012; Piketty and Saez, 2013). This result can be thought of as an extension of Lee and Saez (2012) on the complementarity of the minimum wage and in-work benefits to a more general labor market framework with search and matching frictions

interactions as the ones explored with the neoclassical model in Proposition 3. Appendix B.3 shows that, under the capital allocation problem used throughout the paper to microfound the dependence of ϕ^j on t, the profit elasticity with respect to the corporate tax rate, ϵ_t^j, can be written as a positive linear function of the (industry-specific) degree of capital mobility denoted by $\varepsilon_{k,t}^j = (\partial k^j/\partial t)(t/k^j)$, for $j \in \{S, M\}$. The next proposition explores the implications of $\varepsilon_{k,t}^S \neq \varepsilon_{k,t}^M$ for the optimality of the minimum wage. For analytical simplicity, the result abstracts from the income tax system.

Proposition 8. *Let Assumption 1 hold, and assume that there is no income tax system.*

(i) The marginal social welfare of increasing $\overline{w}$ when t is optimal is increasing in $\varepsilon_{k,t}^M$.

(ii) The effect of $\varepsilon_{k,t}^S$ on the optimal $\overline{w}$ is ambiguous.

Recall from Section 3 that, in the competitive model, the desirability of the minimum wage was increasing in the profit elasticity of the non-exposed industry (Proposition 3), with profit elasticities increasing in capital intensity. Proposition 8 reproduces the same intuition in the non-neoclassical setting: when capital mobility increases in the non-exposed industry, the desirability of the minimum wage increases. Conversely, the effect of capital mobility on exposed industries is ambiguous.

The intuition is as follows. Capital mobility has a negative impact on the optimal t. This increases the optimal $\overline{w}$ because both policies redistribute profits. The attractiveness of $\overline{w}$ is particularly strong if capital mobility arises in the non-exposed industry because $\overline{w}$ does not distort that sector. However, capital mobility in the exposed industry decreases the optimal $\overline{w}$ because of similar distortions on domestic capital. This implies that increasing $\varepsilon_{k,t}^M$ unambiguously increases optimal $\overline{w}$, because it increases the distortion of t without affecting the distortion of $\overline{w}$, while increasing $\varepsilon_{k,t}^S$ has an ambiguous effect on the optimal $\overline{w}$ given the two forces that work in opposite directions.

This result suggests that the minimum wage can be interpreted as an industry-specific corporate tax that minimizes distortions related to capital mobility: the minimum wage can tax profits in exposed industries without distorting capital in non-exposed industries. Since capital mobility and the corresponding distortions of the corporate tax are larger in capital-intensive industries (Garrett et al., 2020; Curtis et al., 2022; Kennedy et al., 2023), the case for a binding minimum wage is stronger when minimum wage workers are concentrated in low-mobility labor-intensive industries, such as the US.

Numerical exercise Finally, the restricted model allows me to perform a suggestive numerical exercise that illustrates the intuitions developed throughout the paper. I calibrate the model to match empirical moments of the US labor market and compute the optimal minimum wage under different tax systems.[38] All the details about the simulations, including calibration and results, are presented in Appendix C.2.

[38] I consider the restrictions described in Assumption 1, except for the restriction on second-order effects that was made only for analytical simplicity (i.e., the numerical exercise allows for decreasing returns to scale).

Three conclusions arise from the exercise. First, given a tax system, social welfare is generally a globally concave function of the minimum wage, so the model generates an interior solution for its optimal value. This result is explained by the fact that wage effects tend to dominate employment effects at low levels because vacancy posting distortions are small near the decentralized equilibrium, but employment effects become larger as the minimum wage departs from the market level. This result, while expected and present in many 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 and is also larger when the corporate tax rate is smaller. Third, the joint optimum seems to use all policies in tandem. Among the cases considered, the optimal policy consists of 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. While the minimum wage and the corporate tax partially work as substitutes, the planner prefers to use them both because each policy's distortion is increasing in its level.

While this exercise should be taken only as suggestive since, as discussed in the next section, the model omits elements that could push the optimal minimum wage in either direction, these results support the idea that the minimum wage can increase the efficiency of tax-based redistribution: optimally combining all instruments can lead to larger social welfare.

5 Caveats and open questions

This section discusses a non-exhaustive list of limitations of the analysis and related open questions that, hopefully, will encourage additional research on the topic.

- *Pigouvian rationales for using the minimum wage:* This paper works with efficient frameworks to focus on the redistributive properties of the minimum wage. Related literature has studied rationales for the minimum wage to solve market inefficiencies such as inefficient monopsony power or misallocation. Extending the optimal policy framework to allow for labor market inefficiencies can shed light on policy tradeoffs or complementarities when dealing with both objectives simultaneously.

- *Avoidance, evasion, and administration costs:* In the real world, income tax schedules are not perfectly enforced and are costly to administrate because of, for example, tax evasion, tax avoidance, and imperfect benefit take up.[39] In fact, one of the main forces that affect effective corporate taxation is corporate tax avoidance (e.g., Hines and Rice, 1994; Auerbach and Slemrod, 1997; Desai

[39]See Andreoni et al. (1998), Slemrod and Yitzhaki (2002), Currie (2006), Kopczuk and Pop-Eleches (2007), Kleven et al. (2011), Chetty et al. (2013), Bhargava and Manoli (2015), Goldin (2018), Cranor et al. (2019), Finkelstein and Notowidigdo (2019), Guyton et al. (2021), and Linos et al. (2022). Abstracting from tax evasion also rules out some complementarities between the minimum wage and the tax system. For example, if workers underreport their incomes, then the minimum wage can increase tax collection by setting a floor on reported labor income (Bíró et al., 2022; Feinmann et al., 2023).

and Dharmapala, 2009; Cooper et al., 2016; Garcia-Bernardo et al., 2022; Tørsløv et al., 2023). I abstract from tax evasion and avoidance in my analysis. Minimum wages can also be difficult to enforce (Stansbury, 2021; Clemens and Strain, 2022), so a more general analysis should consider the relative enforcement and administrative costs of the two instruments.

- *Targeting advantage of the tax system:* An unmodeled benefit of the tax system is that it can target families instead of individuals, and also can tag additional variables (such as family size). This flexibility rarely applies to the minimum wage (Stigler, 1946). This omission is not central when exploring the profit taxation angle of the minimum wage and its interaction with the corporate tax. However, a better understanding of the interaction between the minimum wage and the EITC would benefit from considering more general forms of tagging in the tax and transfer system.

- *Social preferences for pre-distribution:* I assume that the social value of after-tax allocations does not depend on the composition between pre-tax incomes and transfers. However, recent evidence suggests that individuals favor redistribution in terms of pre-tax incomes – and, especially, through increases in the minimum wage – relative to redistribution through taxes and transfers (McCall, 2013; Kuziemko et al., 2015; Saez, 2021). Such social preferences could be incorporated by generalizing the SMWWs to non-utilitarian social preferences (Saez and Stantcheva, 2016).

- *Heterogeneous minimum wages:* National minimum wages may coexist with industry- or region-specific minimum wages (e.g., Dube and Lindner, 2021; Card and Cardoso, 2022). My theoretical results provide a first-order approximation to understand the rationale of such schemes, for example, by assessing if the welfare tradeoffs vary across regions. The analysis, however, is incomplete because heterogeneous minimum wages may induce additional behavioral responses, such as migration responses (e.g., Ahlfeldt et al., 2018, 2023; Monras, 2019; Gaubert et al., 2020; Simon and Wilson, 2021; Pérez, 2022; Todd and Zhang, 2022). A comprehensive assessment of these minimum wage schemes would require modeling these additional distortions.

- *Extensions to the model:* Finally, the model could be extended in several dimensions. One set of extensions consists of including more margins of adjustment to the minimum wage, for example, 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., 2022; Emanuel and Harrington, 2022; Ku, 2022; Ruffini, Forthcoming). Additionally, the model could be extended in several dimensions to include, for example, dynamics, informality, and non-wage amenities. See Appendix B.4 for a discussion on how these omissions are likely to affect the welfare analysis.

6 Conclusions

Despite being a popular and widespread used policy, 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 outcomes such as profits 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.

Using different theoretical frameworks, this paper concludes that the minimum wage should be understood as an alternative tax on profits. As such, its desirability increases when corporate taxes are distortionary, which tends to happen when technologies are capital-intensive. When industries exposed to minimum wage workers are particularly labor-intensive – as is the case in the US – the desirability of the minimum wage is enhanced, as it allows industry-specific corporate taxation that taxes more heavily labor-intensive services with the use of the minimum wage while allowing decreases in general corporate taxes that especially benefit capital-intensive industries.

A general message of the paper is that there are rationales to complement tax-based redistribution with minimum wages. Social planners can benefit from using all instruments together to make redistribution more efficient. Under plausible conditions, optimal redistribution takes the form of a "three-legged stool": targeted income taxes and transfers, positive corporate taxes, and a binding minimum wage.

Bibliography

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

Agostini, Claudio, Eduardo Engel, Andrea Repetto, and Damián Vergara, "Using small businesses for individual tax planning: evidence from special tax regimes in Chile," *International Tax and Public Finance*, 2018, *25*, 1449–1489.

Ahlfeldt, Gabriel, Duncan Roth, and Tobias Seidel, "The regional effects of Germany's national minimum wage," *Economics Letters*, 2018, *172*, 127–130.

_ , _ , and _ , "Optimal minimum wages in spatial economies," *Working Paper*, 2023.

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

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

Andreoni, James, Brian Erard, and Jonathan Feinstein, "Tax compliance," *Journal of Economic Literature*, 1998, *36* (2), 818–860.

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

Atesagaoglu, Orhan Erem and Hakki Yazici, "Optimal Taxation of Capital in the Presence of Declining Labor Share," *Working Paper*, 2023.

Atkinson, Anthony Barnes and Joseph Stiglitz, "The design of tax structure: Direct versus indirect taxation," *Journal of Public Economics*, 1976, *6* (1-2), 55–75.

Auerbach, Alan and James Hines, "Taxation and economic efficiency," in "Handbook of Public Economics," Vol. 3, Elsevier, 2002, pp. 1347–1421.

_ and Joel Slemrod, "The economic effects of the Tax Reform Act of 1986," *Journal of Economic Literature*, 1997, *35* (2), 589–632.

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

_ , Matthew Fisher-Post, Anders Jensen, and Gabriel Zucman, "Globalization and Factor Income Taxation," *Working Paper*, 2023.

Bagger, Jesper, Espen Moen, and Rune Vejlin, "Equilibrium Worker-Firm Allocations and the Deadweight Losses of Taxation," *Working Paper*, 2021.

Baker, Andrew, David Larcker, and Charles Wang, "How Much Should We Trust Staggered Difference-In-Differences Estimates?," *Journal of Financial Economics*, 2022, *144* (2), 370–395.

Berger, David, Kyle Herkenhoff, and Simon Mongey, "Minimum wages and welfare," *Working Paper*, 2022.

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

Best, Michael, Anne Brockmeyer, Henrik Kleven, Johannes Spinnewijn, and Mazhar Waseem, "Production versus revenue efficiency with limited tax capacity: Theory and evidence from Pakistan," *Journal of Political Economy*, 2015, *123* (6), 1311–1355.

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

Bíró, Anikó, Daniel Prinz, László Sándor et al., "The minimum wage, informal pay and tax enforcement," *Journal of Public Economics*, 2022, *215* (1), 104728.

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

Burdett, Kenneth and Dale Mortensen, "Wage differentials, employer size, and unemployment," *International Economic Review*, 1998, pp. 257–273.

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

Caldwell, Sydnee and Nikolaj Harmon, "Outside options, bargaining, and wages: Evidence from coworker networks," *Working Paper*, 2019.

Card, David, "Who set your wage?," *American Economic Review*, 2022, *112* (4), 1075–1090.

_ , Ana Rute Cardoso, Joerg Heining, and Patrick Kline, "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

_ and _ , "Wage Flexibility under Sectoral Bargaining," *Journal of the European Economic Association*, 2022, *20* (5), 2062–2097.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The effect of minimum wages on low-wage jobs," *The Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

_ , _ , _ , and David Zentler-Munro, "Seeing Beyond the Trees: Using Machine Learning to Estimate the Impact of Minimum Wages on Labor Market Outcomes," *Journal of Labor Economics*, 2022, *40* (S1), S203–S247.

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

_ , Adam Guren, Day Manoli, and Andrea Weber, "Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins," *American Economic Review: Papers and Proceedings*, 2011, *101* (3), 471–475.

_ , John Friedman, and Emmanuel Saez, "Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings," *American Economic Review*, 2013, *103* (7), 2683–2721.

Clemens, Jeffrey and Michael Strain, "Understanding "wage theft": Evasion and avoidance responses to minimum wage increases," *Labour Economics*, 2022, p. 102285.

Cooper, Michael, John McClelland, James Pearce, Richard Prisinzano, Joseph Sullivan, Danny Yagan, Owen Zidar, and Eric Zwick, "Business in the United States: Who owns it, and how much tax do they pay?," *Tax Policy and the Economy*, 2016, *30* (1), 91–128.

Costinot, Arnaud and Iván Werning, "Robots, trade, and Luddism: A sufficient statistic approach to optimal technology regulation," *The Review of Economic Studies*, 2023, *90* (5), 2261–2291.

Coviello, Decio, Erika Deserranno, and Nicola Persico, "Minimum wage and individual worker productivity: Evidence from a large US retailer," *Journal of Political Economy*, 2022, *130* (9), 2227–2500.

Craig, Ashley, "Optimal Income Taxation with Spillovers from Employer Learning," *American Economic Journal: Economic Policy*, 2023, *15* (2), 82–125.

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

Currie, Janet, "The take-up of social benefits," in Alan Auerbach, David Card, and John Quigley, eds., *Public Policy and the Income Distribution*, Russell Sage, 2006, pp. 80–148.

Curtis, Mark, Daniel Garrett, Eric Ohrn, Kevin Roberts, and Juan Carlos Suárez-Serrato Serrato, "Capital investment and labor demand: Evidence from 21st century tax policy," *Working Paper*, 2022.

Derenoncourt, Ellora, Clemens Noelke, David Weil, and Bledi Taska, "Spillover effects from voluntary employer minimum wages," *Working Paper*, 2022.

Desai, Mihir and Dhammika Dharmapala, "Corporate tax avoidance and firm value," *The Review of Economics and Statistics*, 2009, *91* (3), 537–546.

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

_ , Johanna Paraknewitz, and Martin Simmler, "Empirical evidence on the global minimum tax: what is a critical mass and how large is the substance-based income exclusion?," *Fiscal Studies*, 2023, *44* (1), 9–21.

_ , Michael Keen, Alan Auerbach, Paul Oosterhuis, John Vella, and Wolfgang Schön, *Taxing profit in a global economy*, Oxford University Press, 2021.

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

_ and _ , "Optimal taxation and public production II: Tax rules," *The American Economic Review*, 1971, *61* (3), 261–278.

Doligalski, Pawel, Abdoulaye Ndiaye, and Nicolas Werquin, "Redistribution with performance pay," *Journal of Political Economy Macroeconomics*, 2023, *1* (2), 371–402.

Draca, Mirko, Stephen Machin, and John Van Reenen, "Minimum wages and firm profitability," *American Economic Journal: Applied Economics*, 2011, *3* (1), 129–51.

Drechsel-Grau, Moritz, "Employment and Reallocation Effects of Higher Minimum Wages," *Working Paper*, 2023.

Drucker, Lev, Katya Mazirov, and David Neumark, "Who pays for and who benefits from minimum wage increases? Evidence from Israeli tax data on business owners and workers," *Journal of Public Economics*, 2021, *199*, 104423.

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

— **and Attila Lindner**, “City limits: What do local-area minimum wages do?,” *Journal of Economic Perspectives*, 2021, *35* (1), 27–50.

— **, Daniele Girardi, Oscar Jorda, and Alan Taylor**, “A local projections approach to difference-in-differences event studies,” *Working Paper*, 2023.

— **, William Lester, and Michael Reich**, “Minimum wage effects across state borders: Estimates using contiguous counties,” *The Review of Economics and Statistics*, 2010, *92* (4), 945–964.

Dustmann, Christian, Attila Lindner, Uta Schönberg, Matthias Umkehrer, and Philipp Vom Berge, “Reallocation effects of the minimum wage: Evidence from Germany,” *Quarterly Journal of Economics*, 2022, *137* (1), 267–328.

Dworczak, Piotr, Scott Duke Kominers, and Mohammad Akbarpour, “Redistribution through markets,” *Econometrica*, 2021, *89* (4), 1665–1698.

Eeckhout, Jan, Chunyang Fu, Wenjian Li, and Xi Weng, “Optimal Taxation and Market Power,” *Working Paper*, 2021.

Emanuel, Natalia and Emma Harrington, “Firm frictions and the payoffs of higher pay: Labor supply and productivity responses to a voluntary firm minimum wage,” *Working Paper*, 2022.

Engbom, Niklas and Christian Moser, “Earnings inequality and the minimum wage: Evidence from Brazil,” *American Economic Review*, 2022, *112* (12), 3803–3847.

Feinmann, Javier, Maximiliano Lauletta, and Roberto Rocha, “Payments under the table: Evidence from Brazil,” *Working Paper*, 2023.

Ferey, Antoine, “Redistribution and unemployment insurance,” *Working Paper*, 2022.

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

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

Forsythe, Eliza, “The Effect of Minimum Wage Policies on the Wage and Occupational Structure of Establishments,” *Journal of Labor Economics*, Forthcoming.

Gandhi, Ashvin and Krista Ruffini, “Minimum wages and employment composition,” *Working Paper*, 2022.

Garcia-Bernardo, Javier, Petr Janskỳ, and Gabriel Zucman, “Did the Tax Cuts and Jobs Act Reduce Profit Shifting by US Multinational Companies?,” *Working Paper*, 2022.

Gardner, John, "Two-stage differences in differences," *Working Paper*, 2021.

Garrett, Daniel, Eric Ohrn, and Juan Carlos Suárez-Serrato, "Tax policy and local labor market behavior," *American Economic Review: Insights*, 2020, *2* (1), 83–100.

Gaubert, Cecile, Patrick Kline, Damián Vergara, and Danny Yagan, "Place-Based Redistribution," *Working Paper*, 2020.

Gerritsen, Aart, "To redistribute or to predistribute? The minimum wage versus income taxation when workers differ in both wages and working hours," *Working Paper*, 2023.

_ **and Bas Jacobs**, "Is a minimum wage an appropriate instrument for redistribution?," *Economica*, 2020, *87* (347), 611–637.

Giupponi, Giulia and Stephen Machin, "Company wage policy in a low-wage labor market," *Working Paper*, 2023.

_ **, Robert Joyce, Attila Lindner, Tom Waters, Thomas Wernham, and Xiaowei Xu**, "The Employment and Distributional Impacts of Nationwide Minimum Wage Changes," *Journal of Labor Economics*, Forthcoming.

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

Gomes, Renato, Jean Marie Lozachmeur, and Alessandro Pavan, "Differential taxation and occupational choice," *The Review of Economic Studies*, 2018, *85* (1), 511–557.

Gravoueille, Maxime, "Wage and employment effects of wage subsidies," *Working Paper*, 2023.

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

Guyton, John, Patrick Langetieg, Daniel Reck, Max Risch, and Gabriel Zucman, "Tax evasion at the top of the income distribution: Theory and evidence," *Working Paper*, 2021.

Haanwinckel, Daniel, "Supply, Demand, Institutions, and Firms: A Theory of Labor Market Sorting and the Wage Distribution," *Working Paper*, 2023.

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

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

Harasztosi, Péter and Attila Lindner, "Who Pays for the minimum Wage?," *The American Economic Review*, 2019, *109* (8), 2693–2727.

Hines, James and Eric Rice, "Fiscal paradise: Foreign tax havens and American business," *The Quarterly Journal of Economics*, 1994, *109* (1), 149–182.

Ho, Christine and Nicola Pavoni, "Efficient child care subsidies," *American Economic Review*, 2020, *110* (1), 162–199.

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

Hummel, Albert Jan, "Monopsony power, income taxation and welfare," *Working Paper*, 2022.

_ , "Unemployment and tax design," *Working Paper*, 2023.

Hungerbühler, Mathias and Etienne Lehmann, "On the optimality of a minimum wage: New insights from optimal tax theory," *Journal of Public Economics*, 2009, *93* (3-4), 464–481.

_ , _ , Alexis Parmentier, and Bruno Van der Linden, "Optimal redistributive taxation in a search equilibrium model," *The Review of Economic Studies*, 2006, *73* (3), 743–767.

Hurst, Erik, Patrick Kehoe, Elena Pastorino, and Thomas Winberry, "The Macroeconomic Dynamics of Labor Market Policies," *Working Paper*, 2023.

Johannesen, Niels, "The Global Minimum Tax," *Journal of Public Economics*, 2022, *212*, 104709.

Kaplow, Louis, "On the undesirability of commodity taxation even when income taxation is not optimal," *Journal of Public Economics*, 2006, *90* (6-7), 1235–1250.

Karabarbounis, Loukas and Brent Neiman, "The global decline of the labor share," *The Quarterly Journal of Economics*, 2014, *129* (1), 61–103.

Keane, Michael and Richard Rogerson, "Reconciling micro and macro labor supply elasticities: A structural perspective," *Annual Review of Economics*, 2015, *7* (1), 89–117.

Kennedy, Patrick, Christine Dobridge, Paul Landefeld, and Jake Mortenson, "The Efficiency-Equity Tradeoff of the Corporate Income Tax: Evidence from the Tax Cuts and Jobs Act," *Working Paper*, 2023.

Kleven, Henrik, "Sufficient Statistics Revisited," *Annual Review of Economics*, 2021, *13* (1), 515–538.

_ , Claus Thustrup Kreiner, Kristian Larsen, and Jakob Egholt Søgaard, "Micro vs Macro Labor Supply Elasticities: The Role of Dynamic Returns to Effort," *Working Paper*, 2023.

— , **Martin Knudsen, Claus Kreiner, Søren Pedersen, and Emmanuel Saez**, "Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark," *Econometrica*, 2011, *79* (3), 651–692.

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

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

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

Kuziemko, Ilyana, Michael I Norton, Emmanuel Saez, and Stefanie Stantcheva, "How elastic are preferences for redistribution? Evidence from randomized survey experiments," *American Economic Review*, 2015, *105* (4), 1478–1508.

Lachowska, Marta, Alexandre Mas, Raffaele Saggio, and Stephen Woodbury, "Wage posting or wage bargaining? A test using dual jobholders," *Journal of Labor Economics*, 2022, *40* (S1), S469–S493.

Landais, Camille, Pascal Michaillat, and Emmanuel Saez, "A macroeconomic approach to optimal unemployment insurance: Applications," *American Economic Journal: Economic Policy*, 2018, *10* (2), 182–216.

— , — , **and** — , "A macroeconomic approach to optimal unemployment insurance: Theory," *American Economic Journal: Economic Policy*, 2018, *10* (2), 152–81.

Lavecchia, Adam, "Minimum wage policy with optimal taxes and unemployment," *Journal of Public Economics*, 2020, *190*, 104228.

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

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

Linos, Elizabeth, Allen Prohofsky, Aparna Ramesh, Jesse Rothstein, and Matt Unrath, "Can nudges increase take-up of the EITC?: Evidence from multiple field experiments," *American Economic Journal: Economic Policy*, 2022, *14* (4), 432–452.

Lobel, Felipe, Thiago Scot, and Pedro Zúñiga, "Corporate taxation and evasion responses: Evidence from a minimum tax in Honduras," *American Economic Journal: Economic Policy*, Forthcoming.

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

Mangin, Sephorah and Benoit Julien, "Efficiency in search and matching models: A generalized Hosios condition," *Journal of Economic Theory*, 2021, *193*, 105208.

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

_ , "Monopsony in labor markets: A review," *ILR Review*, 2021, *74* (1), 3–26.

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

Mayneris, Florian, Sandra Poncet, and Tao Zhang, "Improving or disappearing: Firm-level adjustments to minimum wages in China," *Journal of Development Economics*, 2018, *135*, 20–42.

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

Mill, John Stuart, *Principles of political economy*, D. Appleton, 1884.

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

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

Monras, Joan, "Minimum wages and spatial equilibrium: Theory and evidence," *Journal of Labor Economics*, 2019, *37* (3), 853–904.

Mousavi, Negin, "Optimal Labor Income Tax, Incomplete Markets, and Labor Market Power," *Working Paper*, 2022.

Naito, Hisahiro, "Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency," *Journal of Public Economics*, 1999, *71* (2), 165–188.

Onji, Kazuki, "The response of firms to eligibility thresholds: Evidence from the Japanese value-added tax," *Journal of Public Economics*, 2009, *93* (5-6), 766–775.

Pérez, Jorge, "City Minimum Wages and Spatial Equilibrium Effects," *Working Paper*, 2022.

Piketty, Thomas and Emmanuel Saez, "Optimal labor income taxation," in "Handbook of Public Economics," Vol. 5, Elsevier, 2013, pp. 391–474.

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

_ **, Emmanuel Saez, and Gabriel Zucman**, "Distributional national accounts: Methods and estimates for the United States," *The Quarterly Journal of Economics*, 2018, *133* (2), 553–609.

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

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

Riley, Rebecca and Chiara Rosazza Bondibene, "Raising the standard: Minimum wages and firm productivity," *Labour Economics*, 2017, *44*, 27–50.

Robinson, Joan, *The economics of imperfect competition*, Springer, 1933.

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

Ruffini, Krista, "Higher wages, service quality, and firm profitability: Evidence from nursing homes and minimum wage reforms," *The Review of Economics and Statistics*, Forthcoming.

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

_ , "Direct or indirect tax instruments for redistribution: short-run versus long-run," *Journal of Public Economics*, 2004, *88* (3-4), 503–518.

_ , "Public Economics and Inequality: Uncovering Our Social Nature," *AEA Papers and Proceedings*, 2021, *111*, 1–26.

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

Scheuer, Florian, "Entrepreneurial taxation with endogenous entry," *American Economic Journal: Economic Policy*, 2014, *6* (2), 126–163.

_ **and Iván Werning**, "The taxation of superstars," *The Quarterly Journal of Economics*, 2017, *132* (1), 211–270.

Simon, Andrew and Matthew Wilson, "Optimal minimum wage setting in a federal system," *Journal of Urban Economics*, 2021, *123*, 103336.

Sleet, Christopher and Hakki Yazici, "Taxation, redistribution and frictional labor supply," *Working Paper*, 2017.

Slemrod, Joel and Shlomo Yitzhaki, "Tax avoidance, evasion, and administration," in "Handbook of Public Economics," Vol. 3, Elsevier, 2002, pp. 1423–1470.

Stansbury, Anna, "Do US firms have an incentive to comply with the FLSA and the NLRA?," *Working Paper*, 2021.

_ **and Lawrence Summers**, "Declining worker power and American economic performance," *Brookings Papers on Economic Activity*, 2020.

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

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

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

Todd, Petra and Weilong Zhang, "Distributional Effects of Local Minimum Wages: A Spatial Job Search Approach," *Working Paper*, 2022.

Tørsløv, Thomas, Ludvig Wier, and Gabriel Zucman, "The missing profits of nations," *The Review of Economic Studies*, 2023, *90* (3), 1499–1534.

Vaghul, Kavya and Ben Zipperer, "Historical state and sub-state minimum wage data," *Washington Center for Equitable Growth*, 2016.

Vogel, Jonathan, "The race between education, technology, and the minimum wage," *Working Paper*, 2023.

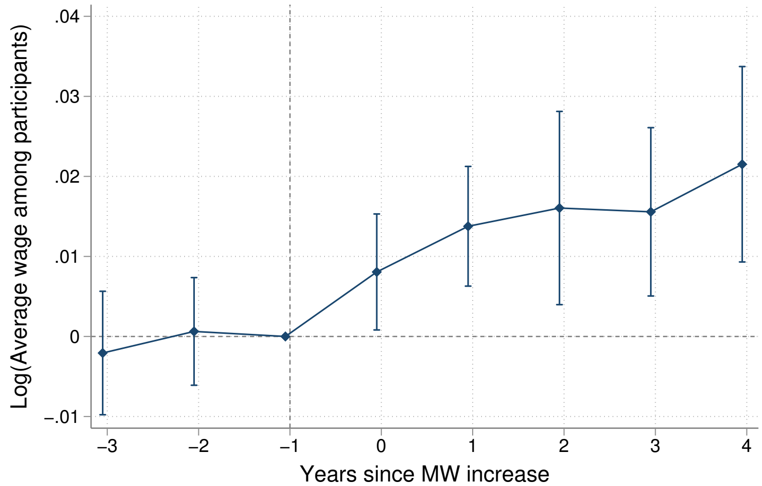
Wright, Randall, Philipp Kircher, Benoît Julien, and Veronica Guerrieri, "Directed Search and Competitive Search: A Guided Tour," *Journal of Economic Literature*, 2021, *59* (1), 90–148.

Wu, Liangjie, "Partially directed search in the labor market," *Working Paper*, 2021.

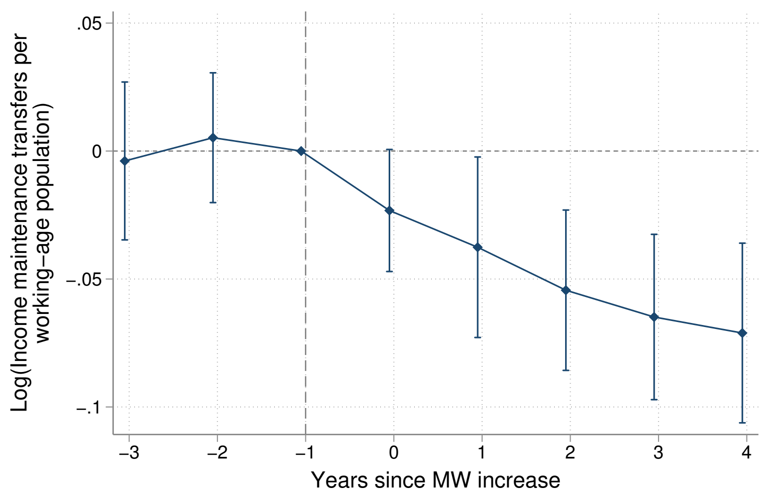
Zucman, Gabriel, "Globalisation, taxation and inequality," *Fiscal Studies*, 2023.

Zurla, Valeria, "Firm Responses to Earned Income Tax Credits: Evidence from Italy," *Working Paper*, 2021.

Figure 1: Effects of state-level minimum wage reforms on low-skill workers



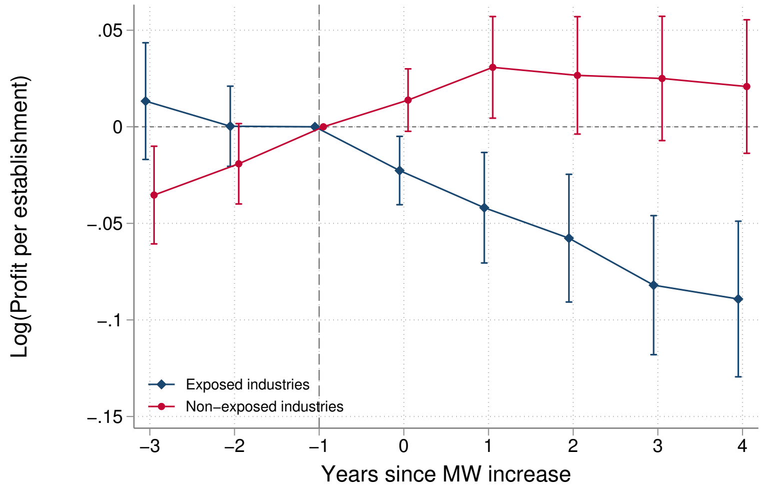
(a) Pre-tax wage (including 0s)



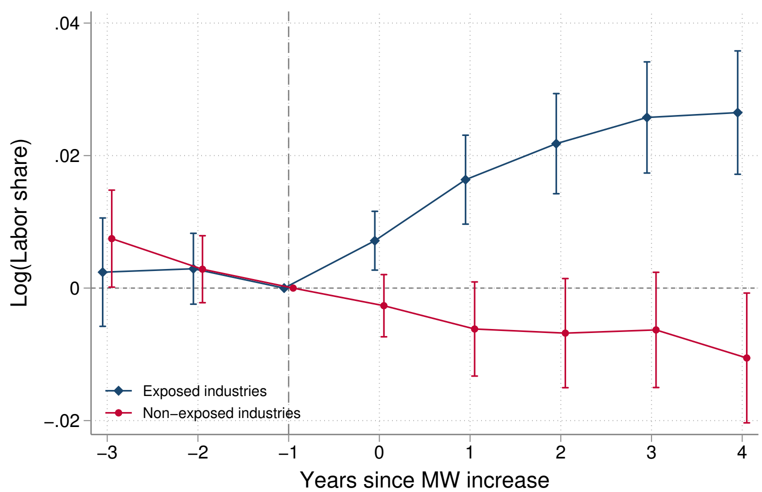
(b) Income maintenance benefits

Notes: These figures plot the estimated β_τ coefficients of equation (1) with their corresponding 95% confidence intervals. Panel (a) uses the log of the average pre-tax wage of active low-skill workers including the unemployed, equal to the average wage conditional on employment times the employment rate, as the dependent variable. Panel (b) uses the log of total income maintenance benefits per working-age population as the dependent variable. The analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. Outcome variables are computed using data from the CPS (MORG and Basic Monthly files) and the BEA regional accounts.

Figure 2: Effects of state-level minimum wage reforms on profits and labor share



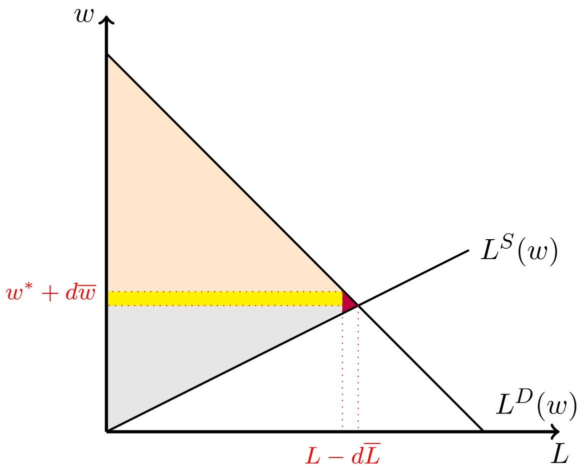
(a) Profits per establishment



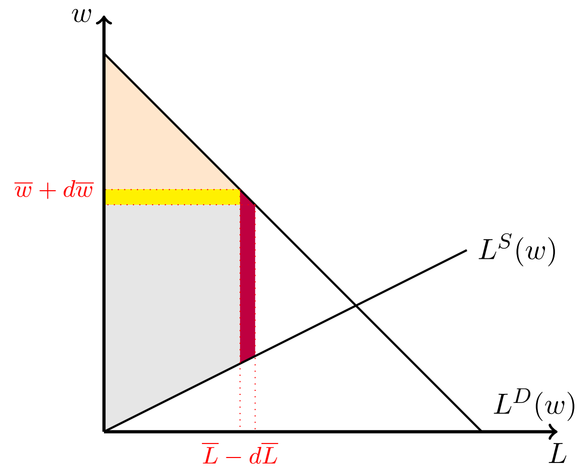
(b) Labor share

Notes: These figures plot the estimated β_τ coefficients of equation (1) with their corresponding 95% confidence intervals. Panel (a) uses the log of the profit per establishment as the dependent variable. Panel (b) uses the log of the labor share as the dependent variable. The analysis is at the state-by-industry-by-year level, standard errors are clustered at the state-by-industry level, and regressions are weighted by the average pre-period state-by-industry-by-year employment. Exposed industries include food and accommodation, retail trade, and low-skill health services. Outcome variables are computed using data from the BEA regional accounts and the QCEW.

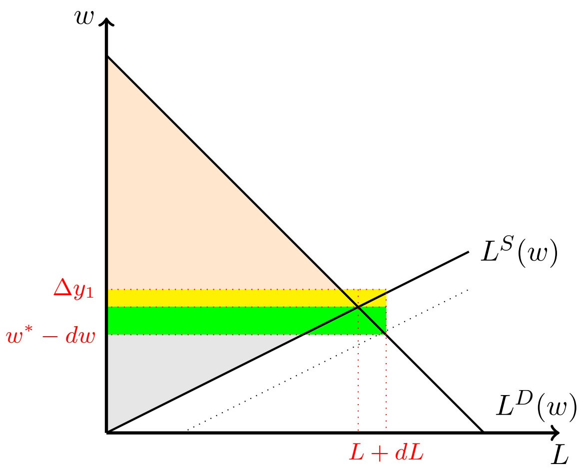
Figure 3: Welfare effects of different reforms



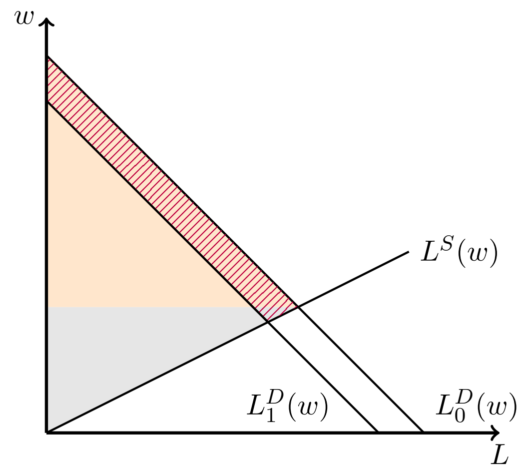
(a) Minimum wage increase just above w^*



(b) Minimum wage increase when $\overline{w}$ is large



(c) Effect of a wage subsidy



(d) Effect of corporate tax increase

Table 1: Descriptive statistics

	Obs.	Mean	Std. Dev.	Min	Max
Low-skill workers:					
Pre-tax wage including the unemployed (annualized)	1,173	19,397	1,226	16,176	24,002
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
Income maintenance benefits (per working-age individual)	1,173	1,057	329	402	2,194
Firms:					
Profit per establishment (Exposed industries)	1,173	170,217	50,459	95,477	539,061
Establishments (Exposed industries)	1,173	70,314	103,291	5,397	914,454
Labor share (Exposed industries)	1,173	0.67	0.04	0.57	0.79
Profit per establishment (Non-exposed industries)	1,173	1,014,998	269,346	423,976	1,826,289
Establishments (Non-exposed industries)	1,173	63,709	69,305	5,818	464,462
Labor share (Non-exposed industries)	1,173	0.45	0.04	0.29	0.62

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. The average pre-tax wage including the unemployed is annualized by computing Average Hourly Wage × Average Weekly Hours × Average Employment Rate × 52. Worker-level aggregates are computed using the CPS-MORG data and the Basic Monthly CPS files. Income maintenance benefits 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. The labor share corresponds to the compensation of employees over the compensation of employees plus taxes on production and imports net of subsidies plus gross operating surplus, all taken from the BEA regional accounts. The industry statistics consider 25 industries that have a relatively large coverage across states and years. Exposed industries 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 industries include manufacturing and non-exposed services. 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, and plastics and rubber products. 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.

Table 2: Difference-in-difference results

Panel (a): Low-skill workers and income maintenance benefits

Dependent variable:	Pre-tax wage (including 0s) (low-skill workers)			Employment rate (low-skill workers)			Inc. maint. benefits (per working-age ind.)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\hat{\beta}$	0.017	0.013	0.015	0.002	0.001	0.005	-0.040	-0.049	-0.050
	(0.006)	(0.006)	(0.005)	(0.003)	(0.003)	(0.003)	(0.015)	(0.012)	(0.015)
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
Elasticity estimate:									
First stage ($\Delta \log \text{MW}$)	0.114	0.117	0.109	0.114	0.117	0.109	0.114	0.117	0.109
	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)
F-test	80.039	83.904	88.700	80.039	83.904	88.700	80.039	83.904	88.700
Second stage (elasticity)	0.148	0.110	0.139	0.022	0.006	0.043	-0.352	-0.415	-0.453
	(0.045)	(0.045)	(0.037)	(0.030)	(0.029)	(0.024)	(0.128)	(0.116)	(0.148)

Panel (b): Profits, establishments, and labor share

Dependent variable:	Profits per establishment (exposed industries)			Number of establishments (exposed industries)			Labor share (exposed industries)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\hat{\beta}$	-0.057	-0.065	-0.063	-0.000	-0.003	-0.005	0.016	0.017	0.018
	(0.012)	(0.011)	(0.012)	(0.010)	(0.006)	(0.005)	(0.003)	(0.003)	(0.003)
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.	254,692	254,692	254,692	256,612	256,612	256,612	255,731	255,731	255,731
Events	50	50	50	50	50	50	50	50	50
Elasticity estimate:									
First stage ($\Delta \log \text{MW}$)	0.116	0.121	0.114	0.116	0.121	0.114	0.116	0.121	0.114
	(0.012)	(0.012)	(0.010)	(0.012)	(0.012)	(0.010)	(0.012)	(0.012)	(0.010)
F-test	97.718	108.492	120.718	97.718	108.492	120.718	97.718	108.492	120.718
Second stage (elasticity)	-0.488	-0.539	-0.554	-0.003	-0.024	-0.043	0.134	0.139	0.155
	(0.101)	(0.100)	(0.107)	(0.085)	(0.047)	(0.046)	(0.027)	(0.023)	(0.024)

Notes: This table shows the estimated β coefficient from equation (2) with corresponding standard errors reported in parentheses. All columns represent different regressions using different dependent variables (all in logarithms) and fixed effects. Columns (1) to (3) of Panel (a) use the average pre-tax wage of low-skill participants, equal to the average pre-tax wage of low-skill workers including the unemployed (alternatively, the average wage conditional on employment times the employment rate), as the dependent variable. Columns (4) to (6) of Panel (a) use the employment rate of low-skill workers as a dependent variable. Columns (7) to (9) of Panel (a) use income maintenance benefits per working-age individual as the dependent variable. Columns (1) to (3) of Panel (b) use the profit per establishment as the dependent variable. Columns (4) to (6) of Panel (b) use the number of establishments as the dependent variable. Columns (7) to (9) of Panel (b) use the labor share as a dependent variable. In Panel (b) I only report the coefficient on exposed industries, which include food and accommodation, retail trade, and low-skill health services. Year FE means that the regression includes year-by-event fixed effects. Year x CR FE means that the regression includes census region-by-year-by-event fixed effects. Year x CD FE means that the regression includes census division-by-year-by-event fixed effects. $\Delta \log \text{MW}$ is the average change in the log of the real state-level minimum wage across events as estimated from an analog of equation (2) that uses $\log \text{MW}_{ite}$ as the dependent variable. The implied elasticity is computed by dividing the point estimate by $\Delta \log \text{MW}$, which corresponds to the second stage of the instrumental variables estimation. In Panel (a), the analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. In Panel (b), the analysis is at the state-by-industry-by-year level, standard errors are clustered at the state-by-industry level, and regressions are weighted by the average pre-period state-by-industry-by-year employment. Outcome variables are computed using data from the CPS (MORG and Basic Monthly files), the BEA regional accounts, and the QCEW.

Minimum Wages and Optimal Redistribution: The Role of Firm Profits

Online Appendix

Damián Vergara - Princeton University[1]

[1]Email: damianvergara@princeton.edu. This version: October, 2023.

A Empirical appendix

This appendix provides more details on the empirical results presented in Section 2. I discuss the event definitions, the empirical strategy, the data, and descriptive statistics. I also present additional results not discussed in the main text.

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 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). This is done by computing employment counts by wage bins and checking whether, on average, the previous year's share of workers with wages below the new minimum wage is above 2% (Cengiz et al., 2019). 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 allows me 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 A.1. Table A.1 displays the list of the considered events with their corresponding treated states.

Estimating equation Estimating event studies in this setting is challenging for two reasons. First, states may increase their minimum wages several times over the period considered. Second, treatment effect heterogeneity may induce bias when treatment adoption is staggered (de Chaisemartin and D'Haultfœuille, 2023; Roth et al., 2023).

To deal with these issues, I implement stacked event studies (Cengiz et al., 2019, 2022; Gardner, 2021; Baker et al., 2022; Dube et al., 2023) 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:

$$\log Y_{ite} = \sum_{\tau=-3}^{4} \beta_\tau D_{i\tau e} + \alpha_{ie} + \gamma_{cd(i)te} + X'_{it}\rho_e + \epsilon_{ite}, \tag{A.1}$$

where i, t, and e index state, year, and event, respectively, Y_{ite} is an outcome of interest, $D_{i\tau e}$ are event indicators with τ the distance from the event (in years), α_{ie} are state-by-event fixed effects, $\gamma_{cd(i)te}$ are census division-by-year-by-event fixed effects, and X_{it} are time-varying controls that includes small state-level minimum wage increases and binding federal minimum wage increases, whose effect is allowed to vary by event e. The inclusion of flexible time fixed effects allows me to better account for time-varying confounders that differentially affect states and industries. I also report results with less conservative time fixed effects. 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 X_{it} includes all the interactions between $\{Early_t, Pre_t, Post_t\} \times \{Small_i, Fed_i\}$ for each event separately. 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 the state-by-year average total population. When the outcome varies at the state-by-industry level, I 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 pooled difference-in-difference regressions:

$$\log Y_{ite} = \beta T_{ie}\text{Post}_{te} + \alpha_{ie} + \gamma_{cd(i)te} + X'_{it}\rho_e + \epsilon_{ite}, \tag{A.2}$$

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 (A.1). The coefficient of interest is β, which captures the average treatment effect in the post-event years (from $\tau = 0$ to $\tau = 4$).

Data Outcome variables consist of state-level aggregates for 1997-2019 computed using publicly available data. The main outcomes include wages, employment, transfers, and profits.

I use the CPS-MORG data to compute average pre-tax hourly wages and the Basic CPS monthly files to compute employment rates at the state-by-year-by-skill level. I also use these sources to compute the average weekly hours worked and labor force participation rates. To estimate a summary effect on workers, I also compute the average pre-tax wage of active workers including the unemployed, which equals the average wage conditional on employment 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 worker-level net tax liabilities at the state-by-year level, I use data from the Bureau of Economic Analysis (BEA) regional accounts. I consider income maintenance benefits, medical benefits, and gross federal income tax liabilities. The BEA definition of income maintenance benefits is as follows: "Income maintenance benefits consist 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 assistance." Medical benefits consider both Medicaid and Medicare programs.

Absent firm-level microdata, I compute a measure of average profits per firm at the industry-by-state-by-year level using state-level aggregates. 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. 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)." 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 (when an industry has low representation in a given state-year cell, the BEA and QCEW do not report aggregates for privacy reasons). Noting that minimum wage workers are not evenly distributed across industries, I group industries into two large groups: exposed and non-exposed industries. Exposed industries mainly include food and accommodation, retail trade, and low-skill health services. I exclude agriculture and mining. I also exclude construction and finance since they experienced 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, and 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. While 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 Statistics of Income (SOI) tables to test for additional fiscal externalities.

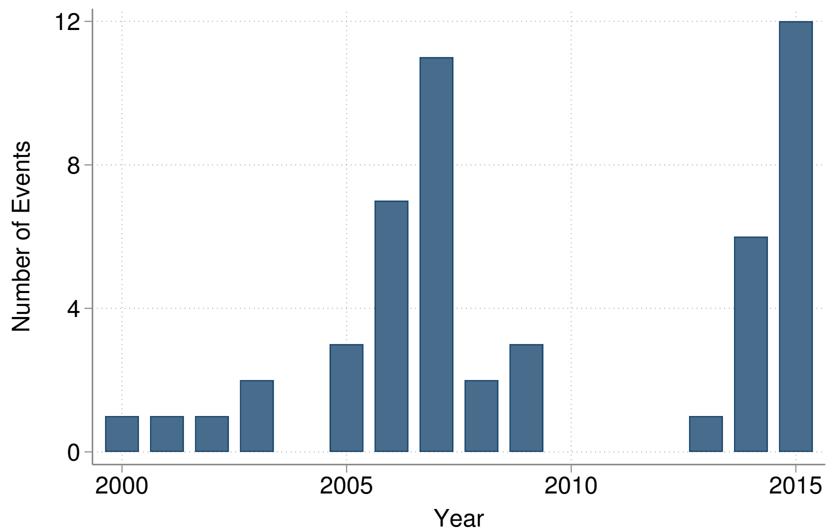
Within-industry labor shares are computed using state-by-industry BEA data on GOS, taxes on production and imports net of subsidies, and compensation of employees. The standard computation is Labor Share = Compensation of employees / (GOS + taxes on production and imports net of subsidies + compensation of employees).

Descriptive statistics Table A.2 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. While the empirical analysis of 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. 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.

Structure of figures and tables Figure A.1 shows the time distribution of the 50 events considered and Table A.1 presents a detailed list of the events considered. Figure A.2 shows the "first-stage", that is, the event study using the real hourly minimum wage as the dependent variable. Table A.2 presents descriptive statistics of the non-stacked estimation sample (state-by-year). Figure A.3 shows event studies for low- and high-skill workers for the average pre-tax wage including the unemployed (wage times employment). Figure A.4 shows results for wages and employment, and Figure A.5 shows results for hours and participation, for both low- and high-skill workers. Tables A.3, A.4, and A.5 show the pooled difference-in-difference estimates related to Figures A.3, A.4, and A.5, omitting the values presented in the main text. Figure A.6 presents a heterogeneity analysis for the low-skill workers' estimates. Figure A.7 tests the robustness of the low-skill workers' estimates to the choice of the wage percentile to truncate the sample. Figure A.8 shows the event studies for different worker-level fiscal externalities. Table A.6 presents related pooled difference-in-difference results, omitting the values presented in the main text. Figure A.9 shows event studies for profit per establishment and number of establishments for exposed and non-exposed industries. Figure A.10 shows event studies for additional firm-level fiscal externalities. Table A.7 shows the pooled difference-in-difference estimates related to Figures A.9 and A.10, omitting the values presented in the main text. Finally, Figure A.11 shoes events studies for the labor share for exposed and non-exposed industries.

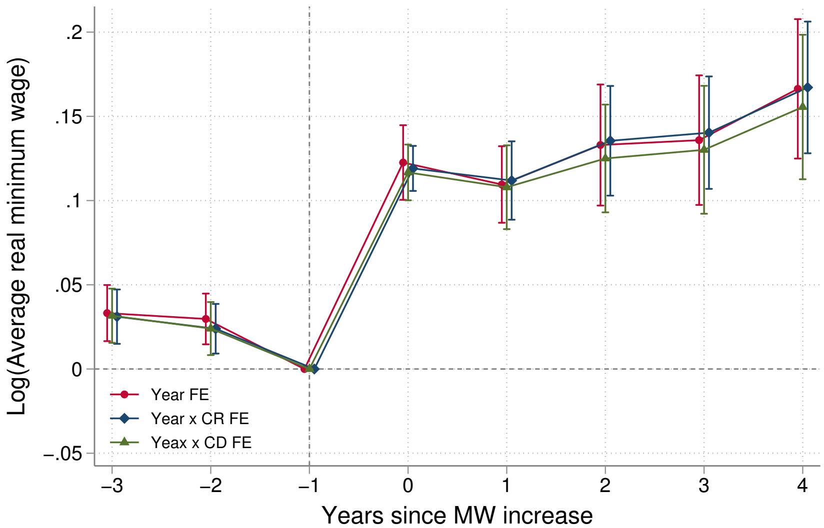
A.1 Additional figures and tables

Figure A.1: State-level events by year



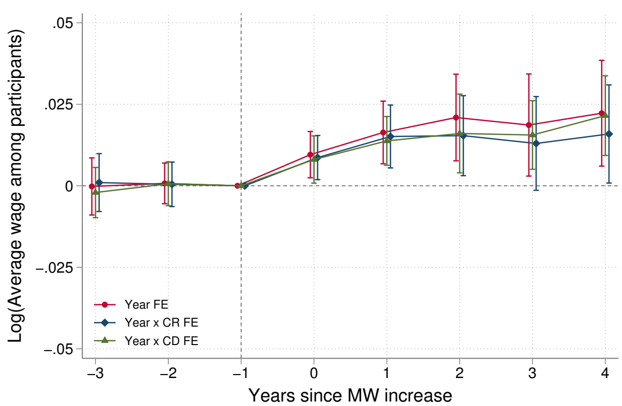
Notes: This figure plots the annual frequency of state-level minimum wage increases classified as events following Cengiz et al. (2019, 2022). Data on minimum wages is taken from Vaghul and Zipperer (2016). A state-level hourly minimum wage increase above the federal level is classified as an event if 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 A.2: First stage: Real minimum wage increase after the event

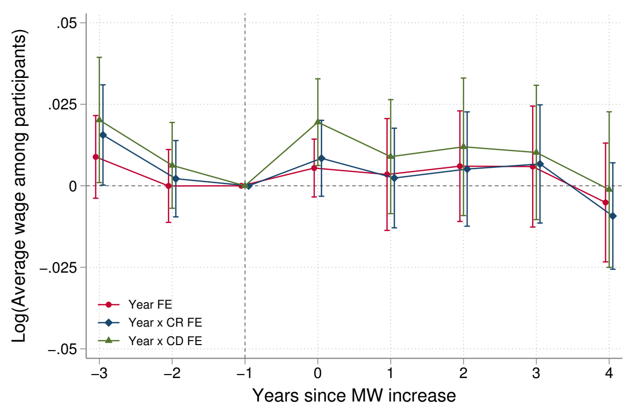


Notes: This figure plots the estimated β_τ coefficients of equation (A.1) with their corresponding 95% confidence intervals using the log real hourly minimum wage as the dependent variable. The analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. The different series correspond to different time fixed effects. Red series use year fixed effects. Blue series use census region-by-year fixed effects. Green series use census division-by-year fixed effects.

Figure A.3: Effects of state-level minimum wage reforms on low- and high-skill workers (average pre-tax wage including the unemployed)



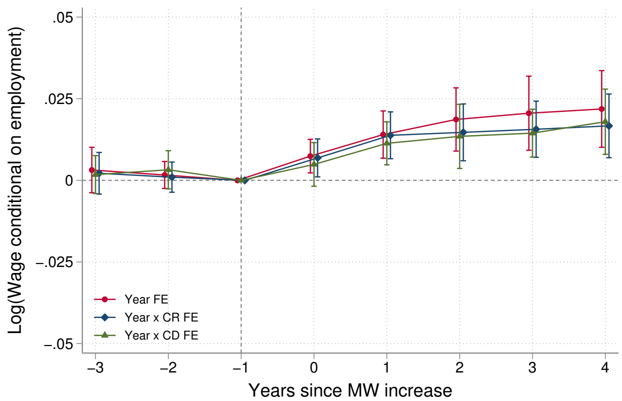
(a) Low-skill workers



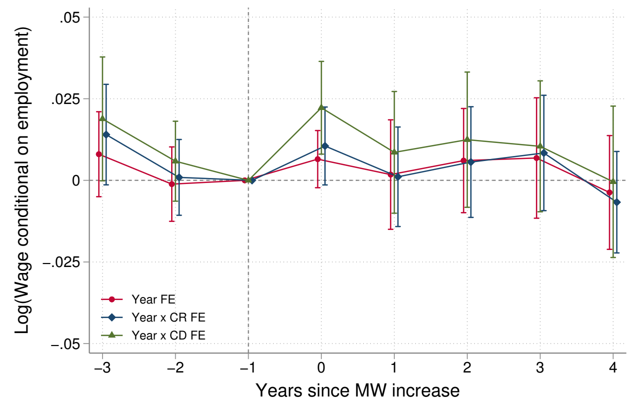
(b) High-skill workers

Notes: These figures plot the estimated β_τ coefficients of equation (A.1) with their corresponding 95% confidence intervals. Panel (a) uses the log of the average pre-tax wage of active low-skill workers including the unemployed as the dependent variable. Panel (b) uses the log of the average pre-tax wage of active high-skill workers including the unemployed as the dependent variable. The analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. In each figure, the different series correspond to different time fixed effects. Red series use year fixed effects. Blue series use census region-by-year fixed effects. Green series use census division-by-year fixed effects.

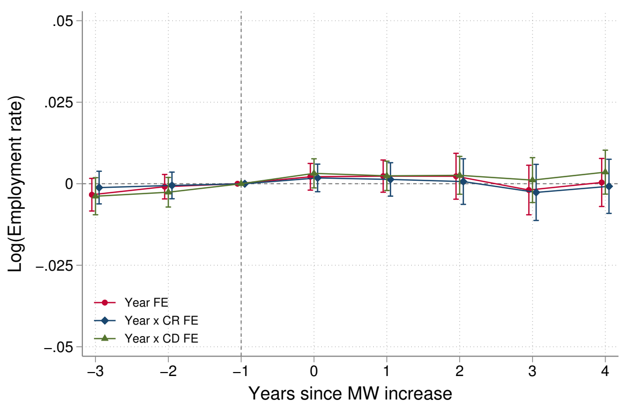
Figure A.4: Effects of state-level minimum wage reforms on low- and high-skill workers (wages and employment)



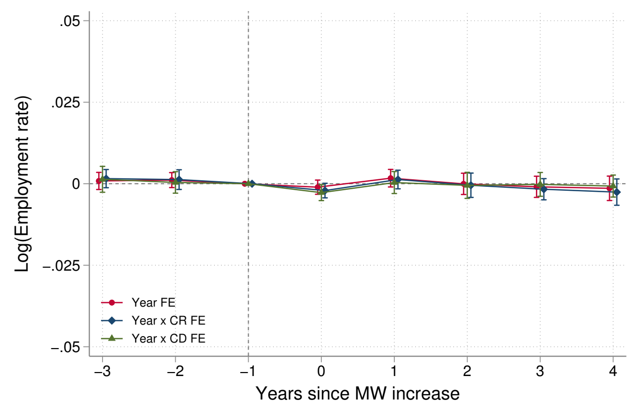
(a) Low-skill workers - Wage conditional on employment



(b) High-skill workers - Wage conditional on employment



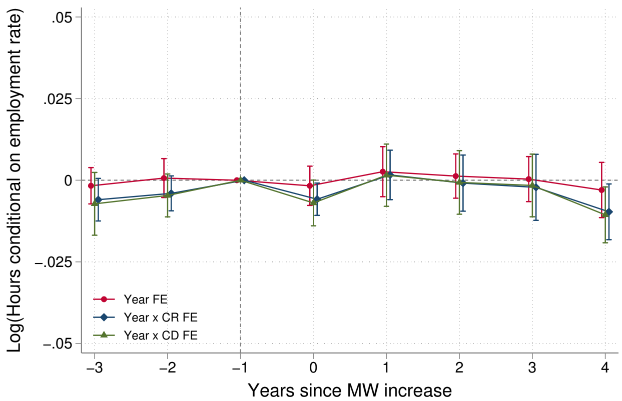
(c) Low-skill workers - Employment rate



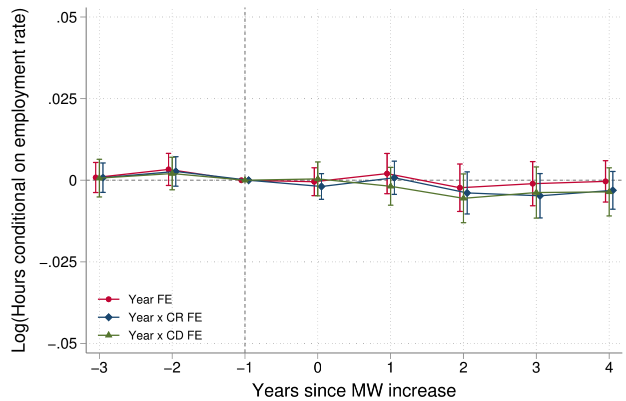
(d) High-skill workers - Employment rate

Notes: These figures plot the estimated β_τ coefficients of equation (A.1) with their corresponding 95% confidence intervals. Panel (a) uses the log of the average pre-tax wage of low-skill workers conditional on employment as the dependent variable. Panel (b) uses the log of the average pre-tax wage of high-skill workers conditional on employment as the dependent variable. Panel (c) uses the log of the employment rate of low-skill workers as the dependent variable. Panel (d) uses the log of the employment rate of high-skill workers. The analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. In each figure, the different series correspond to different time fixed effects. Red series use year fixed effects. Blue series use census region-by-year fixed effects. Green series use census division-by-year fixed effects.

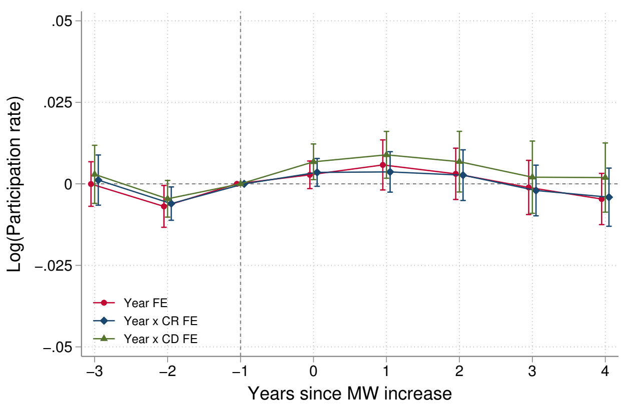
Figure A.5: Effects of state-level minimum wage reforms on low- and high-skill workers (hours and participation)



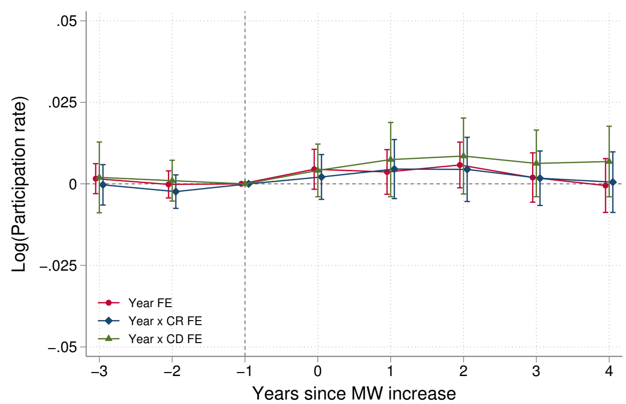
(a) Low-skill workers - Hours worked conditional on employment



(b) High-skill workers - Hours worked conditional on employment



(c) Low-skill workers - Participation rate

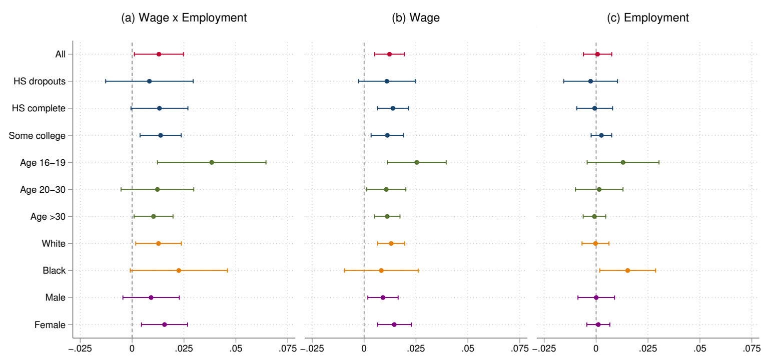


(d) High-skill workers - Participation rate

Notes: These figures plot the estimated β_τ coefficients of equation (A.1) with their corresponding 95% confidence intervals. Panel (a) uses the log of the average weekly hours worked by low-skill workers conditional on employment as the dependent variable. Panel (b) uses the log of the average weekly hours worked by high-skill workers conditional on employment as the dependent variable. Panel (c) uses the log of the participation rate of low-skill workers as the dependent variable. Panel (d) uses the log of the participation rate of high-skill workers. The analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. In each figure, the different series correspond to different time fixed effects. Red series use year fixed effects. Blue series use census region-by-year fixed effects. Green series use census division-by-year fixed effects.

Figure A.6: Minimum wage effects on low-skill workers: Heterogeneity

(a) Year fixed effects

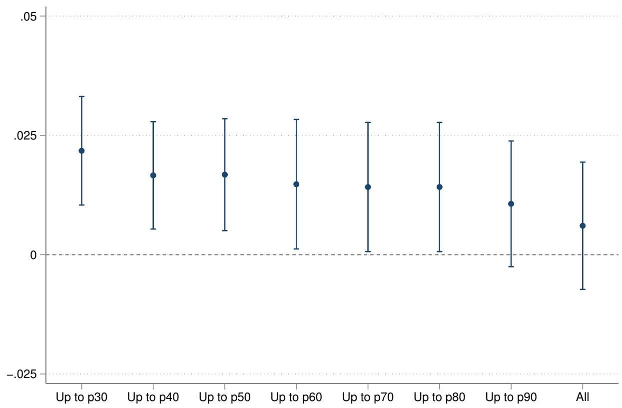


(b) Census region-by-year fixed effects

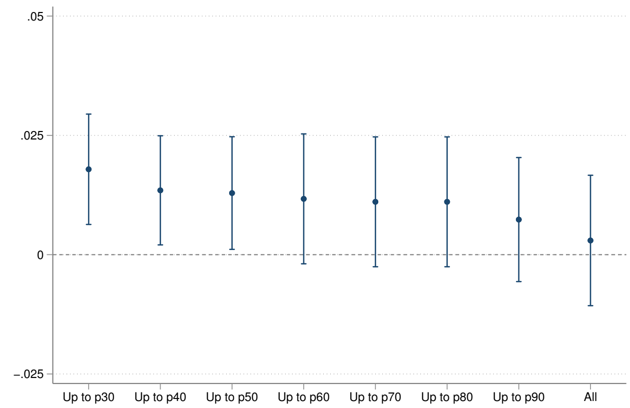
(c) Census division-by-year fixed effects

Notes: These figures plot the estimated β coefficient with its corresponding 95% confidence intervals from equation (A.2) 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 the dependent variable. Panel (b) uses the average pre-tax hourly wage of low-skill workers conditional on employment as the dependent variable. Panel (c) uses the average employment rate of low-skill workers as the dependent variable. 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). The analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. Each panel corresponds to different time fixed effects. Panel (a) year fixed effects. Panel (b) uses census region-by-year fixed effects. Panel (c) uses census division-by-year fixed effects.

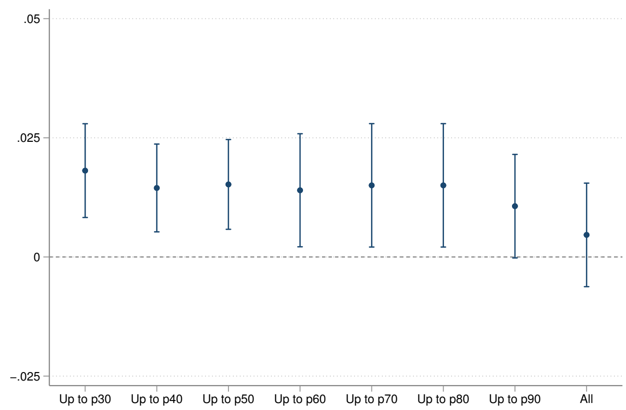
Figure A.7: Minimum wage effects on low-skill workers' welfare: change in percentile considered



(a) Year fixed effects



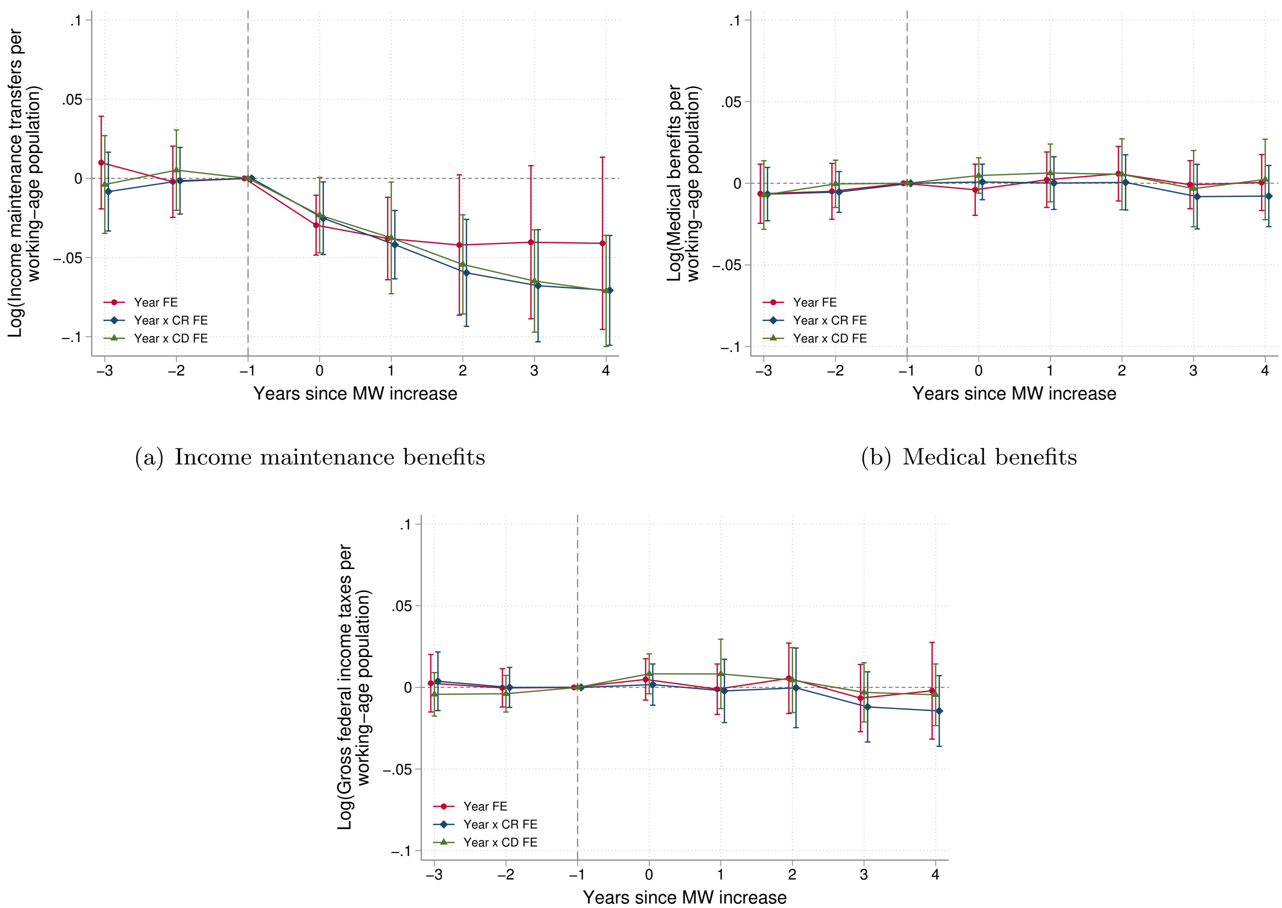
(b) Census region-by-year fixed effects



(c) Census division-by-year fixed effects

Notes: These figures plot the estimated β coefficient with its corresponding 95% confidence intervals from equation (A.2) using the average pre-tax wage of active low-skill workers including the unemployed as the dependent variable. Each coefficient comes from a different regression where the dependent variable is computed using different percentiles to truncate the sample of employed low-skill workers when computing the average wage. The analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. Each panel corresponds to different time fixed effects. Panel (a) year fixed effects. Panel (b) uses census region-by-year fixed effects. Panel (c) uses census division-by-year fixed effects.

Figure A.8: Worker-level fiscal externalities after minimum wage increases



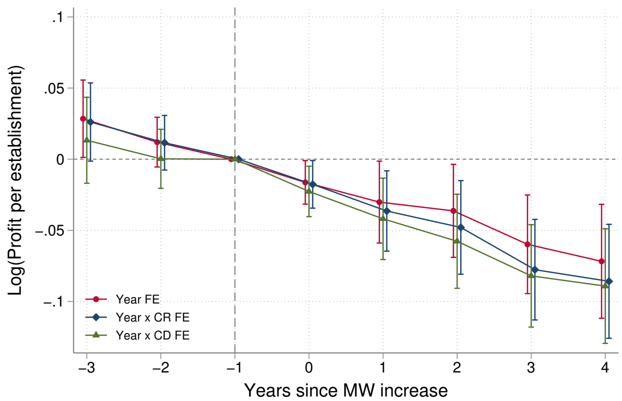
(a) Income maintenance benefits

(b) Medical benefits

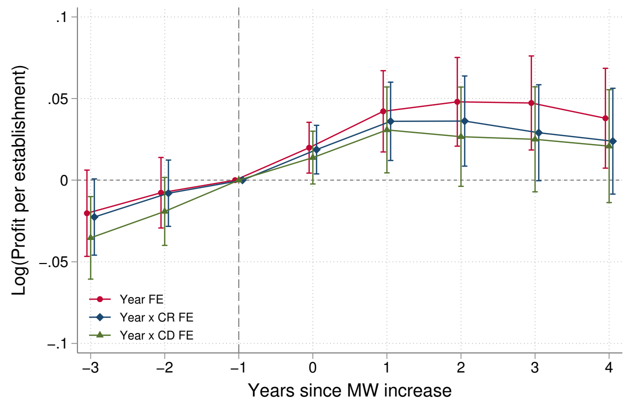
(c) Gross federal income taxes

Notes: These figures plot the estimated β_τ coefficients of equation (A.1) with their corresponding 95% confidence intervals. Panel (a) uses the log income maintenance benefits per working-age population as the dependent variable. Panel (b) uses the log medical benefits per working-age population as the dependent variable. Panel (c) uses the log gross federal income taxes per working-age population as the dependent variable. The analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. Exposed industries include food and accommodation, retail trade, and low-skill health services. In each figure, the different series correspond to different time fixed effects. Red series use year fixed effects. Blue series use census region-by-year fixed effects. Green series use census division-by-year fixed effects.

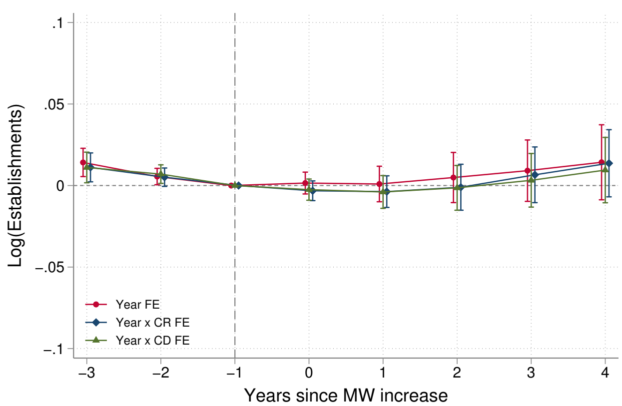
Figure A.9: Effects of state-level minimum wage reforms on profits and establishments



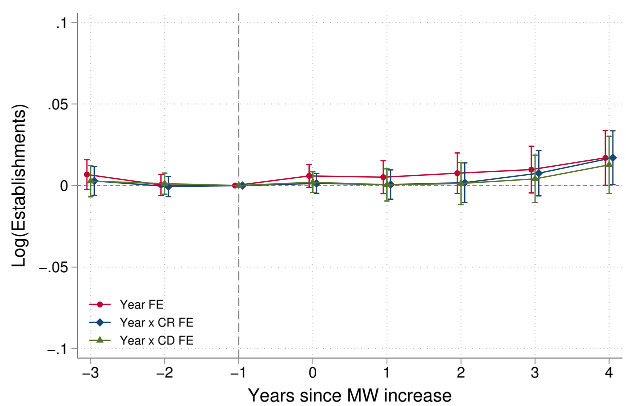
(a) Exposed industries - Profit per establishment



(b) Non-exposed industries - Profit per establishment



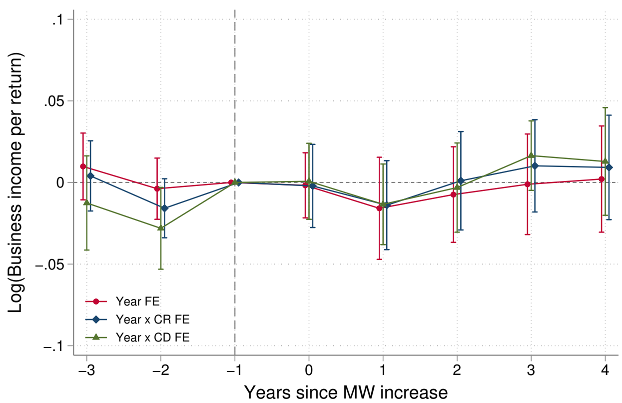
(c) Exposed industries - Establishments



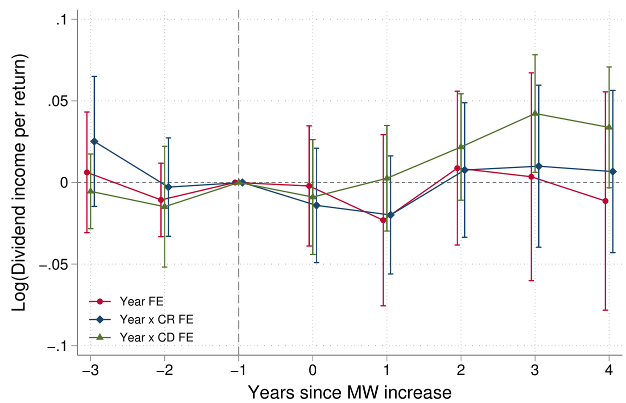
(d) Non-exposed industries - Establishments

Notes: These figures plot the estimated β_τ coefficients of equation (A.1) with their corresponding 95% confidence intervals. Panel (a) uses the log profit per establishment in exposed industries as the dependent variable. Panel (b) uses the log profit per establishment in non-exposed industries as the dependent variable. Panel (c) uses the log of the number of establishments in exposed industries as the dependent variable. Panel (d) uses the log of the number of establishments in non-exposed industries as the dependent variable. The analysis is at the state-by-industry-by-year level, standard errors are clustered at the state-by-industry level, and regressions are weighted by the average pre-event industry-by-state employment. Exposed industries include food and accommodation, retail trade, and low-skill health services. In each figure, the different series correspond to different time fixed effects. Red series use year fixed effects. Blue series use census region-by-year fixed effects. Green series use census division-by-year fixed effects.

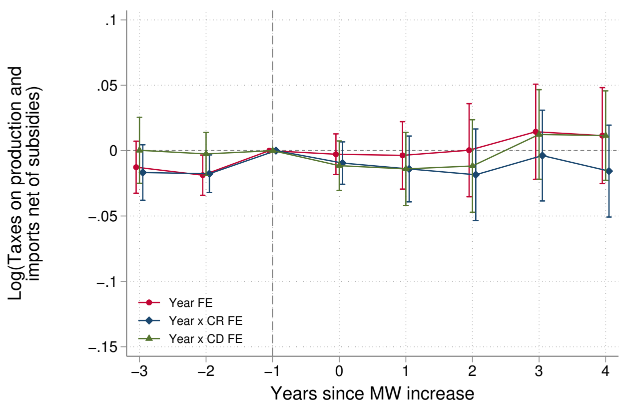
Figure A.10: Additional firm-level fiscal externalities



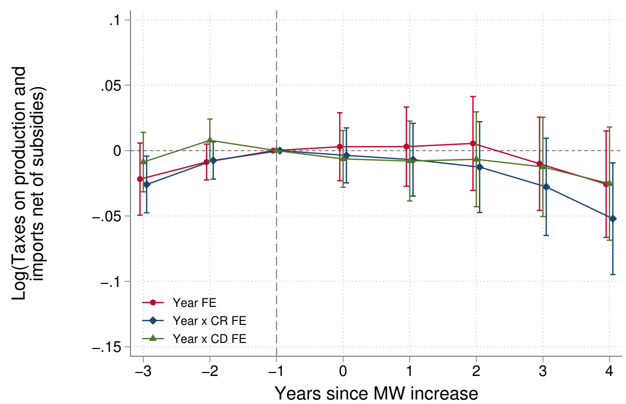
(a) Business income per tax return



(b) Dividend income per tax return



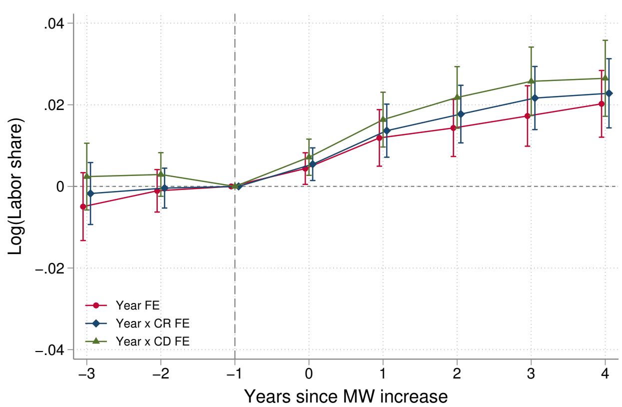
(c) Exposed industries - Taxes on production and imports net of subsidies



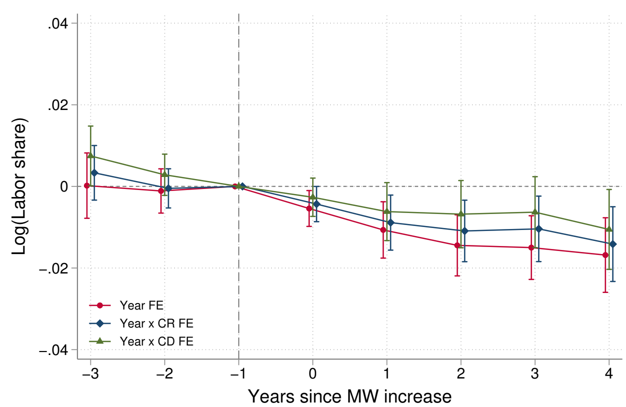
(d) Non-exposed industries - Taxes on production and imports net of subsidies

Notes: These figures plot the estimated β_τ coefficients of equation (A.1) with their corresponding 95% confidence intervals. Panel (a) uses the log business income per tax return as the dependent variable. Panel (b) uses the log dividend income per tax return as the dependent variable. Panel (c) uses the log taxes on production and imports net of subsidies in exposed industries as the dependent variable. Panel (d) uses the log taxes on production and imports net of subsidies in non-exposed industries as the dependent variable. In Panels (a) and (b), the analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. In Panels (c) and (d), the analysis is at the state-by-industry-by-year level, standard errors are clustered at the state-by-industry level, and regressions are weighted by the average pre-period state-by-industry-by-year employment. Exposed industries include food and accommodation, retail trade, and low-skill health services. In each figure, the different series correspond to different time fixed effects. Red series use year fixed effects. Blue series use census region-by-year fixed effects. Green series use census division-by-year fixed effects.

Figure A.11: Effects of state-level minimum wage reforms on the labor share



(a) Exposed industries - Labor share



(b) Non-exposed industries - Labor share

Notes: These figures plot the estimated β_τ coefficients of equation (A.1) with their corresponding 95% confidence intervals. Panel (a) uses the log labor share in exposed industries as the dependent variable. Panel (b) uses the log labor share in non-exposed industries as the dependent variable. The analysis is at the state-by-industry-by-year level, standard errors are clustered at the state-by-industry level, and regressions are weighted by the average pre-event industry-by-state employment. Exposed industries include food and accommodation, retail trade, and low-skill health services. In each figure, the different series correspond to different time fixed effects. Red series use year fixed effects. Blue series use census region-by-year fixed effects. Green series use census division-by-year fixed effects.

Table A.1: 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.

Table A.2: Descriptive statistics

	Obs.	Mean	Std. Dev.	Min	Max
Low-skill workers:					
Pre-tax wage including the unemployed (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:					
Pre-tax wage including the unemployed (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
Taxes and transfers (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
Firms:					
Profit per establishment (Exposed)	1,173	170,217.33	50,459.38	95,477.16	539,061.13
Establishments (Exposed)	1,173	70,313.94	103,291.48	5,397.00	914,454.00
Labor share (Exposed)	1,173	0.67	0.04	0.57	0.79
Profit per establishment (Non-exposed)	1,173	1,014,998.47	269,345.92	423,975.66	1,826,288.63
Establishments (Non-exposed)	1,173	63,709.17	69,305.42	5,818.00	464,462.00
Labor share (Non-exposed)	1,173	0.45	0.04	0.29	0.62

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. The average pre-tax wage including the unemployed is annualized by computing Hourly Wage × Weekly Hours × Employment Rate × 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. The labor share corresponds to the compensation of employees over the compensation of employees plus taxes on production and imports net of subsidies plus gross operating surplus, all taken from the BEA regional accounts.

Table A.3: Difference-in-difference results: Different margins for low-skill workers

Dependent variable:	Wages			Employment			Hours			Participation		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\hat{\beta}$	0.014	0.012	0.011	0.002	0.001	0.005	0.000	-0.000	-0.000	0.003	0.002	0.006
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)
Year FE	Y	N	N	Y	N	N	Y	N	N	Y	N	N
Year x CR FE	N	Y	N	N	Y	N	N	Y	N	N	Y	N
Year x CD FE	N	N	Y	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	10,300	10,300	9,653
Events	50	50	50	50	50	50	50	50	50	50	50	50
Elasticity estimate:												
First stage ($\Delta \log \text{MW}$)	0.114	0.117	0.109	0.114	0.117	0.109	0.114	0.117	0.109	0.114	0.117	0.109
	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)
F-test	80.039	83.904	88.700	80.039	83.904	88.700	80.039	83.904	88.700	80.039	83.904	88.700
Second stage (elasticity)	0.126	0.104	0.096	0.022	0.006	0.043	0.000	-0.004	-0.002	0.029	0.020	0.053
	(0.027)	(0.023)	(0.028)	(0.030)	(0.029)	(0.024)	(0.021)	(0.025)	(0.031)	(0.027)	(0.029)	(0.046)

Notes: This table shows the estimated β coefficient from equation (A.2) with corresponding standard errors reported in parentheses. All columns represent different regressions using different dependent variables (all in logarithms) and fixed effects. All variables are computed for low-skill workers, who are defined as workers without a college degree. Columns (1) to (3) use the average wage conditional on employment as the dependent variable. Columns (4) to (6) use the employment rate as the dependent variable. Columns (7) to (9) use the average weekly hours worked conditional on employment as the dependent variable. Columns (10) to (12) use the labor force participation rate as a dependent variable. Year FE means that the regression includes year-by-event fixed effects. Year x CR FE means that the regression includes census region-by-year-by-event fixed effects. Year x CD FE means that the regression includes census division-by-year-by-event fixed effects. $\Delta \log \text{MW}$ is the average change in the log of the real state-level minimum wage across events as estimated from an analog of equation (A.2) that uses $\log \text{MW}_{ite}$ as the dependent variable. The implied elasticity is computed by dividing the point estimate by $\Delta \log \text{MW}$, which corresponds to the second stage of the instrumental variables estimation. The analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. Outcome variables are computed using data from the CPS (MORG and Basic Monthly files).

Table A.4: Difference-in-difference results: Average pre-tax wage of high-skill workers (including the unemployed)

Dependent variable:	Pre-tax wage (including 0s) (high-skill workers)		
	(1)	(2)	(3)
$\hat{\beta}$	0.000	-0.003	0.002
	(0.007)	(0.006)	(0.008)
Year FE	Y	N	N
Year x CR FE	N	Y	N
Year x CD FE	N	N	Y
Obs.	10,300	10,300	9,653
Events	50	50	50
Elasticity estimate:			
First stage ($\Delta \log \mathrm{MW}$)	0.114	0.117	0.109
	(0.013)	(0.013)	(0.012)
F-test	80.039	83.904	88.700
Second stage (elasticity)	0.002	-0.026	0.015
	(0.062)	(0.050)	(0.077)

Notes: This table shows the estimated β coefficient from equation (2) with corresponding standard errors reported in parentheses. The dependent variable is the average pre-tax wage of high-skill workers including the unemployed, which equals the average wage conditional on employment times the employment rate. All columns represent different regressions. Year FE means that the regression includes year-by-event fixed effects. Year x CR FE means that the regression includes census region-by-year-by-event fixed effects. Year x CD FE means that the regression includes census division-by-year-by-event fixed effects. $\Delta \log \mathrm{MW}$ is the average change in the log of the real state-level minimum wage across events as estimated from an analog of equation (A.2) that uses $\log \mathrm{MW}_{ite}$ as the dependent variable. The implied elasticity is computed by dividing the point estimate by $\Delta \log \mathrm{MW}$, which corresponds to the second stage of the instrumental variables estimation. The analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. Outcome variables are computed using data from the CPS (MORG and Basic Monthly files).

Table A.5: Difference-in-difference results: Different margins for high-skill workers

Dependent variable:	Wages			Employment			Hours			Participation		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\hat{\beta}$	0.001	-0.001	0.003	-0.001	-0.002	-0.001	-0.002	-0.004	-0.004	0.003	0.003	0.006
	(0.007)	(0.006)	(0.008)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
Year FE	Y	N	N	Y	N	N	Y	N	N	Y	N	N
Year x CR FE	N	Y	N	N	Y	N	N	Y	N	N	Y	N
Year x CD FE	N	N	Y	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	10,300	10,300	9,653
Events	50	50	50	50	50	50	50	50	50	50	50	50
Elasticity estimate:												
First stage ($\Delta \log \mathrm{MW}$)	0.114	0.117	0.109	0.114	0.117	0.109	0.114	0.117	0.109	0.114	0.117	0.109
	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)
F-test	80.039	83.904	88.700	80.039	83.904	88.700	80.039	83.904	88.700	80.039	83.904	88.700
Second stage (elasticity)	0.012	-0.008	0.028	-0.009	-0.018	-0.013	-0.015	-0.031	-0.032	0.023	0.029	0.051
	(0.059)	(0.049)	(0.074)	(0.012)	(0.011)	(0.012)	(0.019)	(0.017)	(0.023)	(0.028)	(0.030)	(0.042)

Notes: This table shows the estimated β coefficient from equation (A.2) with corresponding standard errors reported in parentheses. All columns represent different regressions using different dependent variables (all in logarithms) and fixed effects. All variables are computed for high-skill workers, who are defined as workers with a college degree. Columns (1) to (3) use the average wage conditional on employment as the dependent variable. Columns (4) to (6) use the employment rate as the dependent variable. Columns (7) to (9) use the average weekly hours worked conditional on employment as the dependent variable. Columns (10) to (12) use the labor force participation rate as a dependent variable. Year FE means that the regression includes year-by-event fixed effects. Year x CR FE means that the regression includes census region-by-year-by-event fixed effects. Year x CD FE means that the regression includes census division-by-year-by-event fixed effects. $\Delta \log \mathrm{MW}$ is the average change in the log of the real state-level minimum wage across events as estimated from an analog of equation (A.2) that uses $\log \mathrm{MW}_{ite}$ as the dependent variable. The implied elasticity is computed by dividing the point estimate by $\Delta \log \mathrm{MW}$, which corresponds to the second stage of the instrumental variables estimation. The analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. Outcome variables are computed using data from the CPS (MORG and Basic Monthly files).

Table A.6: Difference-in-difference results: Additional fiscal externalities

Dependent variable:	Medical benefits (per working-age ind.)			Gross federal income taxes (per working-age ind.)		
	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\beta}$	0.004	0.001	0.006	-0.000	-0.006	0.005
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(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
Elasticity estimate:						
First stage ($\Delta \log \mathrm{MW}$)	0.114	0.117	0.109	0.114	0.117	0.109
	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)
F-test	80.039	83.904	88.700	80.039	83.904	88.700
Second stage (elasticity)	0.034	0.008	0.051	-0.003	-0.055	0.049
	(0.077)	(0.075)	(0.085)	(0.082)	(0.079)	(0.077)

Notes: This table shows the estimated β coefficient from equation (2) with corresponding standard errors reported in parentheses. All columns represent different regressions using different dependent variables (all in logarithms) and fixed effects. Columns (1) to (3) use the total medical benefits per working-age individual as the dependent variable. Columns (4) to (6) use the gross federal income taxes per working-wage individual as the dependent variable. Year FE means that the regression includes year-by-event fixed effects. Year x CR FE means that the regression includes census region-by-year-by-event fixed effects. Year x CD FE means that the regression includes census division-by-year-by-event fixed effects. $\Delta \log \mathrm{MW}$ is the average change in the log of the real state-level minimum wage across events as estimated from an analog of equation (A.2) that uses $\log \mathrm{MW}_{ite}$ as the dependent variable. The implied elasticity is computed by dividing the point estimate by $\Delta \log \mathrm{MW}$, which corresponds to the second stage of the instrumental variables estimation. The analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. Outcome variables are computed using data from the BEA regional accounts.

Table A.7: Difference-in-difference results: Additional effects on firms and capital income

Dependent variable:	Business income (per inc. tax return)			Dividend income per return (per inc. tax return)			Net taxes on prod. and imports (exposed industries)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\hat{\beta}$	-0.006	0.004	0.016	-0.004	-0.010	0.023	0.014	-0.001	-0.002
	(0.014)	(0.016)	(0.017)	(0.026)	(0.025)	(0.015)	(0.017)	(0.016)	(0.016)
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.	7,733	7,733	7,275	7,733	7,733	7,275	255,488	255,488	255,488
Events	38	38	38	38	38	38	50	50	50
Elasticity estimate:									
First stage ($\Delta \log \text{MW}$)	0.114	0.117	0.109	0.114	0.117	0.109	0.116	0.121	0.114
	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)	(0.010)
F-test	80.039	83.904	88.700	80.039	83.904	88.700	97.718	108.492	120.718
Second stage (elasticity)	-0.025	0.047	0.141	-0.019	-0.054	0.203	0.119	-0.008	-0.017
	(0.111)	(0.120)	(0.135)	(0.213)	(0.194)	(0.124)	(0.148)	(0.135)	(0.142)

Notes: This table shows the estimated β coefficient from equation (A.2) with corresponding standard errors reported in parentheses. All columns represent different regressions using different dependent variables (all in logarithms) and fixed effects. Columns (1) to (3) use total business income per income tax return as the dependent variable. Columns (4) to (6) use total dividend income per income tax return as the dependent variable. Columns (7) to (9) use total taxes on production and imports net of subsidies as the dependent variable. SOI data is only observed until 2018, so columns (1) to (6) omit events that happened in 2015. Year FE means that the regression includes year-by-event fixed effects. Year x CR FE means that the regression includes census region-by-year-by-event fixed effects. Year x CD FE means that the regression includes census division-by-year-by-event fixed effects. $\Delta \log \text{MW}$ is the average change in the log of the real state-level minimum wage across events as estimated from an analog of equation (A.2) that uses $\log \text{MW}_{ite}$ as the dependent variable. The implied elasticity is computed by dividing the point estimate by $\Delta \log \text{MW}$, which corresponds to the second stage of the instrumental variables estimation. In Columns (1) to (6), the analysis is at the state-by-year level, standard errors are clustered at the state level, and regressions are weighted by state-by-year average population. In Columns (7) to (9), the analysis is at the state-by-industry-by-year level, standard errors are clustered at the state-by-industry level, and regressions are weighted by the average pre-period state-by-industry-by-year employment. Outcome variables are computed using data from the SOI state-level tables and the BEA regional accounts.

B Theory appendix

B.1 Proofs of propositions

Proposition 1 Consider the case with no taxes and efficient rationing. Under efficient rationing, $\mathcal{C}_1 = [0, \underline{w}]$. Solving the social planner problem yields the following first-order condition:

$$\begin{aligned}\frac{dSWF}{d\overline{w}} &= -\frac{dL^D(\overline{w})}{d\overline{w}}\omega_L G(0) + \omega_L G(\overline{w} - \underline{w})\frac{d\underline{w}}{d\overline{w}} f(\underline{w}) \\ &\quad + \omega_L \int_0^{\underline{w}} G'(\overline{w} - c) dF(c) + N\omega_k G'(\Pi)\frac{d\Pi}{d\overline{w}}. \end{aligned} \tag{B.1}$$

Since $L^D(\overline{w}) = L^S(\underline{w})$, then $dL^D(\overline{w})/d\overline{w} = f(\underline{w})(d\underline{w}/d\overline{w})$. Note also that $-\left(dL^D(\overline{w})/d\overline{w}\right)\left(\overline{w}/L^D(\overline{w})\right) = \eta_w$. Then, using the SMWWs definitions, and defining by $g_u = \omega_L \left(G(\overline{w} - \underline{w}) - G(0)\right)/\gamma$ as the SMWW of workers that are involuntarily displaced, we can write:

$$\frac{1}{\gamma}\frac{dSWF}{d\overline{w}} = L\left(-\frac{\eta_w}{\overline{w}}g_u + g_1\right) + g_k N\frac{d\Pi}{d\overline{w}}. \tag{B.2}$$

If $\overline{w} = w^*$, then $\overline{w} = \underline{w}$ and $g_u = 0$. Then, increasing the minimum wage just above the market wage is welfare-improving if $Lg_1 + g_k N(d\Pi/d\overline{w}) > 0$. Further increases are welfare-improving if $L\left(-(\eta_w/\overline{w})g_u + g_1\right) + g_k N(d\Pi/d\overline{w}) > 0$. Noting that $d\Pi/d\overline{w} = -\Pi\epsilon_w/\overline{w}$ yields the result. □

Proposition 2 (Version 1 - Perturbation) Consider the allocation induced by the optimal tax system with no minimum wage and consider a small increase in the minimum wage, $d\overline{w}$, paired with a corresponding increase in T_1 of the same magnitude, $dT_1 = d\overline{w}$, and a decrease in the corporate tax rate, dt, such that $dL = 0$. This reform has no direct effect on workers' welfare since $dy_1 = d(w - T_1) = 0$, $dy_0 = 0$, and $dL = 0$. However, the reform generates a positive fiscal externality (driven by dT_1) and a negative fiscal externality (driven by dt and the possibly non-zero effect on profits) that are not necessarily offset by each other. Also, changes in after-tax profits have a welfare impact proportional to g_k.

The positive fiscal externality is given by $Ld\overline{w}$. Also, $dL = (\partial L/\partial\overline{w})d\overline{w} + (\partial L/\partial t)dt = 0$, which implies that $-(\eta_w/\overline{w})d\overline{w} - (\eta_t/t)dt = 0$. Then, the corporate tax cut that yields $dL = 0$ is given by:

$$dt = -\frac{\eta_w}{\eta_t}\frac{t}{\overline{w}}d\overline{w}. \tag{B.3}$$

Note that $d\Pi = \phi_w d\overline{w} + \phi_t dt - ld\overline{w} = -(\Pi/\overline{w})\epsilon_w d\overline{w} - (\Pi/t)\epsilon_t dt$. The negative fiscal externality is given by $d[tN\Pi] = dtN\Pi + tNd\Pi$, which equals:

$$
\begin{aligned}
d[tN\Pi] &= -\frac{\eta_w}{\eta_t}\frac{t}{\overline{w}}d\overline{w}N\Pi + tN\left(-\frac{\Pi}{\overline{w}}\epsilon_w + \frac{\Pi}{t}\epsilon_t\frac{\eta_w}{\eta_t}\frac{t}{\overline{w}}\right)d\overline{w}, \\
&= d\overline{w}\frac{N\Pi t}{\overline{w}}\left(-\frac{\eta_w}{\eta_t} - \epsilon_w + \epsilon_t\frac{\eta_w}{\eta_t}\right). \qquad \text{(B.4)}
\end{aligned}
$$

The welfare effect on capitalists is given by $Ng_k d[(1-t)\Pi] = Ng_k\left[-dt\Pi + (1-t)d\Pi\right]$, which equals:

$$
\begin{aligned}
Ng_k d[(1-t)\Pi] &= Ng_k\left(\frac{\eta_w}{\eta_t}\frac{t}{\overline{w}}\Pi d\overline{w} + (1-t)\left(-\frac{\Pi}{\overline{w}}\epsilon_w + \frac{\Pi}{t}\epsilon_t\frac{\eta_w}{\eta_t}\frac{t}{\overline{w}}\right)d\overline{w}\right), \\
&= d\overline{w}g_k\frac{N\Pi t}{\overline{w}}\left(\frac{\eta_w}{\eta_t} + \frac{1-t}{t}\left[-\epsilon_w + \epsilon_t\frac{\eta_w}{\eta_t}\right]\right). \qquad \text{(B.5)}
\end{aligned}
$$

Then, the reform is desirable if:

$$
Ld\overline{w} + d[tN\Pi] + g_k d[(1-t)N\Pi] > 0, \qquad \text{(B.6)}
$$

which can be written as:

$$
L + \frac{tN\Pi}{\overline{w}}\left(-(1-g_k)\frac{\eta_w}{\eta_t} - \left(1 + \frac{g_k(1-t)}{t}\right)\left(\epsilon_w - \epsilon_t\frac{\eta_w}{\eta_t}\right)\right) > 0, \qquad \text{(B.7)}
$$

or, alternatively:

$$
1 > \frac{tN\Pi}{L\overline{w}}\left((1-g_k)\frac{\eta_w}{\eta_t} + \left(1 + \frac{g_k(1-t)}{t}\right)\left(\epsilon_w - \epsilon_t\frac{\eta_w}{\eta_t}\right)\right). \quad \square \qquad \text{(B.8)}
$$

Proposition 2 (Version 2 - First principles) I first solve the optimal tax problem with no minimum wage and then introduce a minimum wage to derive conditions for its desirability.

The Lagrangian of the social planner is given by:

$$
\begin{aligned}
\mathcal{L}(T_0, T_1, t) &= (1-L)\omega_L G(-T_0) + \int_0^{w-T_1+T_0}\omega_L G(w - T_1 - c)dF(c) \\
&\quad + N\omega_k G((1-t)\Pi) + \gamma\left[(1-L)T_0 + LT_1 + tN\Pi\right]. \qquad \text{(B.9)}
\end{aligned}
$$

The first order condition with respect to T_0 is given by:

$$
\begin{aligned}
\frac{\partial\mathcal{L}}{\partial T_0} &= -\frac{\partial L}{\partial T_0}\omega_L G(-T_0) + \omega_L G(w - T_1 - w + T_1 - T_0)f(w - T_1 + T_0)\left(\frac{\partial w}{\partial T_0} + 1\right) \\
&\quad -(1-L)\omega_L G'(-T_0) + N\omega_k G'((1-t)\Pi)(1-t)\frac{\partial\Pi}{\partial w}\frac{\partial w}{\partial T_0} \\
&\quad +\gamma(1-L) + \gamma\frac{\partial L}{\partial T_0}(T_1 - T_0) + \gamma tN\frac{\partial\Pi}{\partial w}\frac{\partial w}{\partial T_0} = 0. \qquad \text{(B.10)}
\end{aligned}
$$

In the absence of a minimum wage, $L = F(w - T_1 + T_0)$ and, therefore, $\partial L/\partial T_0 = f(w-T_1+T_0)(\partial w/\partial T_0 + 1)$, so the first two terms cancel. Then, the expression reduces to:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial T_0} &= -(1-L)\omega_L G'(-T_0) + N\omega_k G'((1-t)\Pi)(1-t)\frac{\partial \Pi}{\partial w}\frac{\partial w}{\partial T_0} \\ &\quad + \gamma(1-L) + \gamma f(w - T_1 + T_0)\left(\frac{\partial w}{\partial T_0} + 1\right)(T_1 - T_0) + \gamma t N \frac{\partial \Pi}{\partial w}\frac{\partial w}{\partial T_0} = 0. \end{aligned} \tag{B.11}$$

The labor market clearing condition is given by $F(w - T_1 + T_0) = Nl^d$. Total differentiation yields:

$$f(w - T_1 + T_0)(dw - dT_1 + dT_0) = -\frac{L\eta_w}{w}dw - \frac{L\eta_t}{t}dt, \tag{B.12}$$

which, after imposing $dT_1 = dt = 0$, implies that:

$$\frac{\partial w}{\partial T_0} = \frac{-f(w - T_1 + T_0)}{f(w - T_1 + T_0) + \frac{L\eta_w}{w}} < 0, \tag{B.13}$$

so an increase in T_0 (a reduction in the transfer) decreases the equilibrium wage.

The FOC with respect to T_1 is given by:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial T_1} &= -\frac{\partial L}{\partial T_1}\omega_L G(-T_0) + \omega_L G(w - T_1 - w + T_1 - T_0) f(w - T_1 + T_0)\left(\frac{\partial w}{\partial T_1} - 1\right) \\ &\quad + \omega_L \int_0^{w - T_1 + T_0} G'(w - T_1 - c)\left(\frac{\partial w}{\partial T_1} - 1\right) dF(c) + N\omega_k G'((1-t)\Pi)(1-t)\frac{\partial \Pi}{\partial w}\frac{\partial w}{\partial T_1} \\ &\quad + \gamma L + \gamma \frac{\partial L}{\partial T_1}(T_1 - T_0) + \gamma t N \frac{\partial \Pi}{\partial w}\frac{\partial w}{\partial T_1} = 0. \end{aligned} \tag{B.14}$$

Per the same argument above, the first two terms cancel, so the expression reduces to:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial T_1} &= \omega_L \int_0^{w - T_1 + T_0} G'(w - T_1 - c)\left(\frac{\partial w}{\partial T_1} - 1\right) dF(c) + N\omega_k G'((1-t)\Pi)(1-t)\frac{\partial \Pi}{\partial w}\frac{\partial w}{\partial T_1} \\ &\quad + \gamma L + \gamma f(w - T_1 + T_0)\left(\frac{\partial w}{\partial T_1} - 1\right)(T_1 - T_0) + \gamma t N \frac{\partial \Pi}{\partial w}\frac{\partial w}{\partial T_1} = 0. \end{aligned} \tag{B.15}$$

Likewise, the wage effect can be written as:

$$\frac{\partial w}{\partial T_1} = \frac{f(w - T_1 + T_0)}{f(w - T_1 + T_0) + \frac{L\eta_w}{w}} > 0, \tag{B.16}$$

so an increase in T_1 (a reduction in the transfer) increases the equilibrium wage, in a symmetric fashion with respect to T_0, i.e., $\partial w/\partial T_0 = -\partial w/\partial T_1$.

The FOC with respect to t is given by:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial t} &= -\frac{\partial L}{\partial t}\omega_L G(-T_0) + \omega_L G(w - T_1 - w + T_1 - T_0) f(w - T_1 + T_0)\frac{\partial w}{\partial t} \\
&+\omega_L \int_0^{w-T_1+T_0} G'(w - T_1 - c)\frac{\partial w}{\partial t} dF(c) + N\omega_k G'((1-t)\Pi)(1-t)\left(\frac{\partial \Pi}{\partial w}\frac{\partial w}{\partial t} + \frac{\partial \Pi}{\partial t}\right) \\
&-N\omega_k G'((1-t)\Pi)\Pi + \gamma\frac{\partial L}{\partial t}(T_1 - T_0) + \gamma N\Pi + \gamma t N\left(\frac{\partial \Pi}{\partial w}\frac{\partial w}{\partial t} + \frac{\partial \Pi}{\partial t}\right) = 0. \qquad \text{(B.17)}
\end{aligned}$$

Per the same argument above, the first two terms cancel out so the expression reduces to:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial t} &= \omega_L \int_0^{w-T_1+T_0} G'(w - T_1 - c)\frac{\partial w}{\partial t} dF(c) + N\omega_k G'((1-t)\Pi)(1-t)\left(\frac{\partial \Pi}{\partial w}\frac{\partial w}{\partial t} + \frac{\partial \Pi}{\partial t}\right) \\
&-N\omega_k G'((1-t)\Pi)\Pi + \gamma f(w - T_1 + T_0)\frac{\partial w}{\partial t}(T_1 - T_0) \\
&+\gamma N\Pi + \gamma t N\left(\frac{\partial \Pi}{\partial w}\frac{\partial w}{\partial t} + \frac{\partial \Pi}{\partial t}\right) = 0. \qquad \text{(B.18)}
\end{aligned}$$

Likewise, the wage effect can be written as:

$$\frac{\partial w}{\partial t} = \frac{-\frac{L\eta_t}{t}}{f(w - T_1 + T_0) + \frac{L\eta_w}{w}} < 0, \qquad \text{(B.19)}$$

so an increase in t decreases the equilibrium wage.

Finally, the FOC with respect to γ yields:

$$\frac{\partial \mathcal{L}}{\partial \gamma} = (1 - L)T_0 + LT_1 + tN\Pi = 0, \qquad \text{(B.20)}$$

which closes the system of equations.

Now, assume that the optimal policy considers a binding minimum wage, $\overline{w} \geq w^*$. This affects the previous analysis in two dimensions. First, fixing the wage eliminates equilibrium responses on wages, so $\partial w/\partial T_0 = \partial w/\partial T_1 = \partial w/\partial t = 0$. Second, excess labor supply implies that employment is exclusively determined by labor demand, conditional on $\overline{w}$. This implies that $\partial L/\partial T_0 = \partial L/\partial T_1 = 0$. Then, the previous FOCs are reduced to:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial T_0} &= \omega_L G(\overline{w} - \underline{w} - T_0) f(\underline{w} - T_1 + T_0) - (1 - L)\omega_L G'(-T_0) + \gamma(1 - L) = 0. \qquad \text{(B.21)} \\
\frac{\partial \mathcal{L}}{\partial T_1} &= -\omega_L G(\overline{w} - \underline{w} - T_0) f(\underline{w} - T_1 + T_0) \\
&\quad -\omega_L \int_0^{\underline{w}-T_1+T_0} G'(\overline{w} - T_1 - c) dF(c) + \gamma L = 0. \qquad \text{(B.22)} \\
\frac{\partial \mathcal{L}}{\partial t} &= -\frac{L\eta_t}{\overline{w}}\left(G(\overline{w} - \underline{w} - T_0) - G(-T_0)\right) - N\omega_k G'((1-t)\Pi)(1-t)\frac{\Pi}{t}\epsilon_t \\
&\quad -N\omega_k G'((1-t)\Pi)\Pi - \lambda\frac{L}{t}\eta_t(T_1 - T_0) + \gamma N\Pi - \gamma t N\frac{\Pi}{t}\epsilon_t = 0. \qquad \text{(B.23)}
\end{aligned}$$

In terms of welfare weights, the previous expressions can be written as:

$$\frac{\partial \mathcal{L}}{\partial T_0} = \Omega - (1-L)g_0 + (1-L) = 0. \tag{B.24}$$

$$\frac{\partial \mathcal{L}}{\partial T_1} = -\Omega - Lg_1 + L = 0. \tag{B.25}$$

$$\frac{\partial \mathcal{L}}{\partial t} = -\frac{L\eta_t}{\overline{w}}g_u - Ng_k(1-t)\frac{\Pi}{t}\epsilon_t - Ng_k\Pi - \frac{L}{t}\eta_t(T_1 - T_0) + N\Pi - tN\frac{\Pi}{t}\epsilon_t = 0, \tag{B.26}$$

where $\Omega = \omega_L G(\overline{w} - \underline{w} - T_0)f(\underline{w} - T_1 + T_0)\gamma^{-1} > 0$, and $g_u = \omega_L \left(G(\overline{w} - \underline{w} - T_0) - G(-T_0)\right)/\gamma$. Note that equations (B.24) and (B.25) imply that $(1-L)g_0 + Lg_1 = 1$, and that $g_0 > 1 > g_1$.

Moreover, using similar arguments as above, if the optimum considers a binding minimum wage, the FOC with respect to $\overline{w}$ is given by:

$$\frac{\partial \mathcal{L}}{\partial \overline{w}} = -\frac{\eta_w L}{\overline{w}}g_u + \Omega + Lg_1 - Ng_k(1-t)\frac{\Pi}{\overline{w}}\epsilon_w - \frac{L}{\overline{w}}\eta_w(T_1 - T_0) - tN\frac{\Pi}{\overline{w}}\epsilon_w = 0. \tag{B.27}$$

From equation (B.25), we have that $\Omega + Lg_1 = L$. Also, from equation (B.26), we have that:

$$T_1 - T_0 = -g_u - \frac{tN\Pi}{L\eta_t}\left(\frac{1-t}{t}\epsilon_t + g_k - 1 + \epsilon_t\right). \tag{B.28}$$

Then, replacing in equation (B.27) yields:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \overline{w}} &= -\frac{\eta_w L}{\overline{w}}g_u + L - \frac{N\Pi}{\overline{w}}\left(g_k(1-t)\epsilon_w + t\epsilon_w\right) \\ &\quad + \frac{L}{\overline{w}}\eta_w\left(g_u + \frac{tN\Pi}{L\eta_t}\left(\frac{1-t}{t}\epsilon_t + g_k - 1 + \epsilon_t\right)\right) = 0,\end{aligned} \tag{B.29}$$

which can be written as:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \overline{w}} &= L - \frac{tN\Pi}{\overline{w}}\left(g_k\frac{(1-t)}{t}\epsilon_w + \epsilon_w\right) \\ &\quad + \frac{tN\Pi}{\overline{w}}\frac{\eta_w}{\eta_t}\left(\frac{1-t}{t}\epsilon_t + g_k - 1 + \epsilon_t\right) = 0.\end{aligned} \tag{B.30}$$

If, when evaluated at $\overline{w} = w^*$ the LHS of equation (B.30) is positive, then complementing the optimal tax system with a binding minimum wage is optimal. Grouping terms yields the exact same expression as in equation (B.8).

Finally note that, when $g_k = 0$, equation (B.26) is reduced to:

$$N\Pi(1-\epsilon_t) = \frac{L\eta_t}{\overline{w}}g_u + \frac{L}{t}\eta_t(T_1 - T_0). \tag{B.31}$$

At $\overline{w} = w^*$, $g_u = 0$. Then, if $T_1 > T_0$, $\epsilon_t < 1$. $\square$

Proposition 3 Consider the same reform used in the proof of Proposition 2 (Version 1). Starting from the allocation induced by the optimal tax system with no minimum wage, consider a small increase in the minimum wage, $d\overline{w}$, paired with a corresponding increase in T_1 of the same magnitude, $dT_1 = d\overline{w}$, and a decrease in the corporate tax rate, dt, such that $dL^l = 0$. This reform has no direct effect on low-skill workers' welfare since $dy_1^l = d(w^l - T_1) = 0$, $dy_0 = 0$, and $dL^l = 0$. However, the reform generates, in the low-skill sector, a positive fiscal externality (driven by dT_1) and a negative fiscal externality (driven by dt and the effect on profits) that are not necessarily offset by each other.

On top of the effect on the low-skill sector, the reform generates fiscal externalities in the high-skill sector. The decrease in the corporate tax rate generates an effect on the corporate tax revenue collected from firms in the high-skill sector and a corresponding increase in high-skill employment. The employment effect has no welfare effects since marginal high-skill workers are initially indifferent between working and not working, but the increase in employment generates a first-order fiscal externality if $T_0^* \neq T_2^*$.

Since labor markets are segmented, all the effects pertaining to the low-skill sector are equal to the ones derived in Proposition 2 after replacing $\omega_k = 0$:

$$L^l d\overline{w} - d\overline{w}\frac{tN^l\Pi^l}{\overline{w}}\left(\frac{\eta_w^l}{\eta_t^l}(1-\epsilon_t^l)+\epsilon_w^l\right). \tag{B.32}$$

On the other hand, the fiscal externality in the high-skill sector is given by $d[tN^h\Pi^h] + d[L^h\Delta T_2]$, where $\Delta T_2 = T_2 - T_0$. Then, $d[tN^h\Pi^h] = dtN^h\Pi^h + tN^h d\Pi^h = dtN^h\Pi^h(1-\epsilon_t^h)$. Likewise, $d[L^h\Delta T_2] = -dt\eta_t^h(L^h/t)\Delta T_2$. Together with equation (B.3), this implies that the fiscal externality on the high-skill sector is given by:

$$-\frac{\eta_w^l}{\eta_t^l}\frac{t}{\overline{w}}d\overline{w}N^h\Pi^h(1-\epsilon_t^h) + \frac{\eta_w^l}{\eta_t^l}\frac{L^h}{\overline{w}}d\overline{w}\eta_t^h\Delta T_2. \tag{B.33}$$

Then, the reform is desirable if:

$$L^l - \frac{tN^l\Pi^l}{\overline{w}}\left(\frac{\eta_w^l}{\eta_t^l}(1-\epsilon_t^l)+\epsilon_w^l\right) - \frac{\eta_w^l}{\eta_t^l}\frac{t}{\overline{w}}N^h\Pi^h(1-\epsilon_t^h) + \frac{\eta_w^l}{\eta_t^l}\frac{L^h}{\overline{w}}\eta_t^h\Delta T_2 > 0, \tag{B.34}$$

or, alternatively:

$$1 > \frac{tN^l\Pi^l}{L^l\overline{w}}\left(\frac{\eta_w^l}{\eta_t^l}(1-\epsilon_t^l)+\epsilon_w^l\right) + \frac{1}{L^l\overline{w}}\frac{\eta_w^l}{\eta_t^l}\left(tN^h\Pi^h(1-\epsilon_t^h) - L^h\eta_t^h\Delta T_2\right). \quad \square \tag{B.35}$$

Proposition 4 With no taxes, the SWF is given by:

$$\begin{aligned} SWF &= \left(L_I^l + L_I^h + K_I\right) G(0) + \alpha_l \int_0^{U^l} G(U^l - c) dF_l(c) \\ &+ \alpha_h \int_0^{U^h} G(U^h - c) dF_h(c) + K \sum_{j \in \mathcal{J}}^{J} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} G\left(\Pi^j(\psi) - \xi\right) dO_j(\psi). \end{aligned} \tag{B.36}$$

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:

$$\begin{aligned} \frac{dSWF}{d\overline{w}} &= \left(\frac{dK_I}{d\overline{w}} - \frac{dL_A^l}{d\overline{w}} - \frac{dL_A^h}{d\overline{w}}\right) G(0) \\ &+ \alpha_l G(0) f_l(U^l) \frac{dU^l}{d\overline{w}} + \alpha_l \frac{dU^l}{d\overline{w}} \int_0^{U^l} G'(U^l - c) dF_l(c) \\ &+ \alpha_h G(0) f_h(U^h) \frac{dU^h}{d\overline{w}} + \alpha_h \frac{dU^h}{d\overline{w}} \int_0^{U^h} G'(U^h - c) dF_h(c) \\ &+ K \sum_{j \in \mathcal{J}} \sigma_j \left(\int_{\psi_j^*}^{\overline{\psi}} G'(\Pi^j(\psi) - \xi) \frac{d\Pi^j(\psi)}{d\overline{w}} dO_j(\psi) - \frac{d\psi_j^*}{d\overline{w}} G(0) o_j(\psi_j^*) \right). \end{aligned} \tag{B.37}$$

Note that $dL_A^s/d\overline{w} = d(\alpha_s F_s(U^s))/d\overline{w} = \alpha_s f_s(U^s)(dU^s/d\overline{w})$. Also, $dK_I/d\overline{w} = d(K \sum_{j \in \mathcal{J}} \sigma_j O_j(\psi_j^*))/d\overline{w} = K \sum_{j \in \mathcal{J}} \sigma_j o_j(\psi_j^*)(d\psi_j^*/d\overline{w})$. Then, equation (B.37) is reduced to:

$$\begin{aligned} \frac{dSWF}{d\overline{w}} &= \alpha_s \frac{dU^l}{d\overline{w}} \int_0^{U^l} G'(U^l - c) dF_l(c) + \alpha_h \frac{dU^h}{d\overline{w}} \int_0^{U^h} G'(U^h - c) dF_h(c) \\ &+ K \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} G'(\Pi^j(\psi) - \xi) \frac{d\Pi^j(\psi)}{d\overline{w}} dO_j(\psi). \end{aligned} \tag{B.38}$$

Using the marginal welfare weights definitions, equation (B.38) can be written as:

$$\frac{1}{\gamma} \frac{dSWF}{d\overline{w}} = \frac{dU^l}{d\overline{w}} L_A^l g_1^l + \frac{dU^h}{d\overline{w}} L_A^h g_1^h + K \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} g_\psi^j \frac{d\Pi^j(\psi)}{d\overline{w}} dO_j(\psi). \quad \square \tag{B.39}$$

Proposition 5 With fixed taxes, the Lagrangian is given by:

$$\begin{aligned} \mathcal{L} &= \left(L_I^l + L_I^h + K_I\right) G(y_0) \\ &+ \alpha_l \int_0^{U^l - y_0} G(U^l - c) dF_l(c) + \alpha_h \int_0^{U^h - y_0} G(U^h - c) dF_h(c) \\ &+ K \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} G((1-t)\Pi^j(\psi, t) - \xi) dO_j(\psi) + \gamma \Bigg[\int \left(E_m^l T(w_m^l) + E_m^h T(w_m^h)\right) dm \\ &+ tK \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \Pi^j(\psi, t) dO_j(\psi) - y_0 \left(L_I^l + L_I^h + K_I + \rho^l \cdot L_A^l + \rho^h L_A^h\right) \Bigg], \end{aligned} \tag{B.40}$$

where γ is the budget constraint multiplier. Since $\rho^s L_A^s = L_A^s - \int E_m^s dm$, and using the fact that $L_I^l + L_I^h + L_A^l + L_A^h = 1$, equation (B.40) can be written as:

$$\begin{aligned}
\mathcal{L} \;=\; & \left(L_I^l + L_I^h + K_I\right) G(y_0) \\
& +\alpha_l \int_0^{U^l - y_0} G(U^l - c) dF_l(c) + \alpha_h \int_0^{U^h - y_0} G(U^h - c) dF_h(c) \\
& +K \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} G((1-t)\Pi^j(\psi, t) - \xi) dO_j(\psi) + \gamma \Bigg[\int \Big(E_m^l (T(w_m^l) + y_0) \\
& +E_m^h (T(w_m^h) + y_0) \Big) dm + tK \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \Pi^j(\psi, t) dO_j(\psi) - y_0 \left(1 + K_I\right) \Bigg] . \qquad \text{(B.41)}
\end{aligned}$$

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

$$\begin{aligned}
\frac{d\mathcal{L}}{d\overline{w}} \;=\; & \left(\frac{dK_I}{d\overline{w}} - \frac{dL_A^s}{d\overline{w}} - \frac{dL_A^h}{\overline{w}} \right) G(y_0) \\
& +G(y_0)\alpha_l f_l(U^l - y_0) \frac{dU^l}{d\overline{w}} + \alpha_l \frac{dU^l}{d\overline{w}} \int_0^{U^l - y_0} G'(U^l - c) dF_l(c) \\
& +G(y_0)\alpha_h f_h(U^h - y_0) \frac{dU^h}{d\overline{w}} + \alpha_h \frac{dU^h}{d\overline{w}} \int_0^{U^h - y_0} G'(U^h - c) dF_h(c) \\
& +K \sum_{j \in \mathcal{J}} \sigma_j \left[\int_{\psi_j^*}^{\overline{\psi}} G'((1-t)\Pi^j(\psi, t) - \xi)(1-t) \frac{d\Pi^j(\psi, t)}{d\overline{w}} dO_j(\psi) - G(y_0) o_j(\psi^*) \frac{d\psi_j^*}{d\overline{w}} \right] \\
& \gamma \left[\int \left(\frac{dE_m^l}{d\overline{w}} \left(T(w_m^l) + y_0 \right) + E_m^l T'(w_m^l) \frac{dw_m^l}{d\overline{w}} + \frac{dE_m^h}{d\overline{w}} \left(T(w_m^h) + y_0 \right) + E_m^h T'(w_m^h) \frac{dw_m^h}{d\overline{w}} \right) dm \right. \\
& \left. +tK \sum_{j \in \mathcal{J}} \sigma_j \left(\int_{\psi_j^*}^{\overline{\psi}} \frac{d\Pi^j(\psi, t)}{d\overline{w}} dO_j(\psi) - \Pi^j(\psi_j^*, t) o_j(\psi_j^*) \frac{d\psi_j^*}{d\overline{w}} \right) - y_0 \frac{dK_I}{d\overline{w}} \right] . \qquad \text{(B.42)}
\end{aligned}$$

We have that $dK_I/d\overline{w} = K \sum_{j \in \mathcal{J}} \sigma_j o_j(\psi_j^*)(d\psi_j^*/d\overline{w})$ and $dL_A^s/d\overline{w} = \alpha_s f_s(U^s - y_0)(dU^s/d\overline{w})$ for $s \in \{l, h\}$. Also, $\Pi^j(\psi_j^*, t) = (\xi + y_0)/(1-t) \equiv \Pi^{marg}$, for all j. Using the social marginal weights definitions, and grouping common terms, equation (B.42) can be written as:

$$\begin{aligned}
\frac{1}{\gamma} \frac{d\mathcal{L}}{d\overline{w}} \;=\; & \frac{dU^l}{d\overline{w}} L_A^l g_1^l + \frac{dU^h}{d\overline{w}} L_A^h g_1^h + K(1-t) \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} g_\psi^j \frac{d\Pi^j(\psi, t)}{d\overline{w}} dO_j(\psi) \\
& + \int \left(\frac{dE_m^l}{d\overline{w}} \left(T(w_m^l) + y_0 \right) + E_m^l T'(w_m^l) \frac{dw_m^l}{d\overline{w}} \right) dm \\
& + \int \left(\frac{dE_m^h}{d\overline{w}} \left(T(w_m^h) + y_0 \right) + E_m^h T'(w_m^h) \frac{dw_m^h}{d\overline{w}} \right) dm \\
& +tK \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \frac{d\Pi^j(\psi, t)}{d\overline{w}} dO_j(\psi) - \frac{dK_I}{d\overline{w}} \left(t\Pi^{marg} + y_0 \right) . \quad \square \qquad \text{(B.43)}
\end{aligned}$$

Proposition 6 Assuming either $\max_i w_i^l < \min_j w_j^h$ or the availability of skill-specific income tax schedules allows to solve the problem by pointwise maximization on final allocations. That is, the planner chooses y_0 and $\Delta y_m^s = y_m^s - y_0$, for all m and $s \in \{l, h\}$, and then recover taxes by $T(w_m^s) + y_0 = w_m^s - \Delta y_m^s$. The Lagrangian is given by:

$$\begin{aligned}
\mathcal{L} &= \left(L_I^l + L_I^h + K_I\right) G(y_0) + \alpha_l \int_0^{U^l - y_0} G(U^l - c) dF_l(c) + \alpha_h \int_0^{U^h - y_0} G(U^h - c) dF_h(c) \\
&\quad + K \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} G((1-t)\Pi^j(\psi, t) - \xi) dO_j(\psi) + \gamma \left[\int \left(E_m^l (w_m^l - \Delta y_m^l) + E_m^h (w_m^h - \Delta y_m^h) \right) dm \right. \\
&\quad \left. + tK \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \Pi^j(\psi, t) dO_j(\psi) - y_0 (1 + K_I) \right]. \qquad \text{(B.44)}
\end{aligned}$$

The main difference with respect to Proposition 5 is that the social planner leaves Δy_m^s constant, for all m and $s \in \{l, h\}$, when choosing $\overline{w}$. Then, the first-order condition with respect to $\overline{w}$ is given by:

$$\begin{aligned}
\frac{1}{\gamma} \frac{\partial \mathcal{L}}{\partial \overline{w}} &= \frac{\partial U^l}{\partial \overline{w}} L_A^l g_1^l + \frac{\partial U^h}{\partial \overline{w}} L_A^h g_1^h + K(1-t) \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} g_\psi^j \frac{\partial \Pi^j(\psi, t)}{\partial \overline{w}} dO_j(\psi) \\
&\quad + \int \left(\frac{\partial E_m^l}{\partial \overline{w}} \left(T(w_m^l) + y_0 \right) + E_m^l \frac{\partial w_m^l}{\partial \overline{w}} \right) dm \\
&\quad + \int \left(\frac{\partial E_m^h}{\partial \overline{w}} \left(T(w_m^h) + y_0 \right) + E_m^h \frac{\partial w_m^h}{\partial \overline{w}} \right) dm \\
&\quad + tK \sum_{j \in \mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \frac{\partial \Pi^j(\psi, t)}{\partial \overline{w}} dO_j(\psi) - \frac{\partial K_I}{\partial \overline{w}} \left(t \Pi^{marg} + y_0 \right), \qquad \text{(B.45)}
\end{aligned}$$

where, as in previous propositions, common terms are canceled, the definitions of the SMWWs are used, and $\Pi^j(\psi_j^*, t) = (\xi + y_0)/(1-t) \equiv \Pi^{marg}$, for all j. Since Δy_m^s is fixed, we have that $\partial U^s / \partial \overline{w} = p_\theta (\partial \theta_m^s / \partial \overline{w}) \Delta y_m^s$, $\partial E_m^s / \partial \overline{w} = p_\theta (\partial \theta_m^s / \partial \overline{w}) L_A^s + p \alpha_s f_s (U^s - y_0)(\partial U^s / \partial \overline{w})$, profit effects are given by equations (B.130) and (B.131), and wage spillovers are given by equation (B.127), except for firms for which the minimum wage is binding, case in which $\partial w_m^l / \partial \overline{w} = 1$. When Δy_m^s is fixed, changes in $\overline{w}$ do not affect L_A^s and L_m^s in partial equilibrium, so $\partial \theta_m^s / \partial \overline{w}$ is mediated by potential changes in vacancies, which in turn can generate a general equilibrium effect on applicants.

The first-order condition with respect to y_0 yields:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial y_0} = \left(L_I^l + L_I^h + K_I\right) G'(y_0) + \alpha_l \int_0^{U^l - y_0} G'(U^l - c) dF_l(c) \\
+ \alpha_h \int_0^{U^l - y_0} G'(U^l - c) dF_l(c) - \gamma(1 + K_I) &= 0, \qquad \text{(B.46)}
\end{aligned}$$

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

$\Gamma_0 g_0 + \Gamma_1^l g_1^l + \Gamma_1^h g_1^h = 1$, where $\Gamma_0 = \left(L_I^l + L_I^h + K_I\right)/(1+K_I)$, $\Gamma_1^s = L_A^s/(1+K_I)$, and $\Gamma_0 + \Gamma_1^l + \Gamma_1^h = 1$.

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

$$\begin{aligned}
\frac{1}{\gamma}\frac{\partial \mathcal{L}}{\partial t} = \; & L_A^l g_1^l \frac{\partial U^l}{\partial t} + L_A^h g_1^h \frac{\partial U^h}{\partial t} + K \sum_{j\in\mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} g_\psi^j \left[-\Pi^j(\psi,t) + (1-t)\frac{\partial \Pi^j(\psi,t)}{\partial t}\right] dO_j(\psi) \\
& + \int \left(\frac{\partial E_m^l}{\partial t}\left(w_m^l - \Delta y_m^l\right) + E_m^l \frac{\partial w_m^l}{\partial t}\right) dm + \int \left(\frac{\partial E_m^h}{\partial t}\left(w_m^h - \Delta y_m^h\right) + E_m^h \frac{\partial w_m^h}{\partial t}\right) dm \\
& + tK \sum_{j\in\mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \frac{\partial \Pi^j(\psi,t)}{\partial t} dO_j(\psi) + K \sum_{j\in\mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \Pi^j(\psi,t) dO_j(\psi) \\
& - tK \sum_{j\in\mathcal{J}} \sigma_j \Pi^j(\psi_j^*,t) o(\psi_j^*) \frac{\partial \psi_j^*}{\partial t} - y_0 \frac{\partial K_I}{\partial t} = 0. \qquad \text{(B.47)}
\end{aligned}$$

From equations (B.106) and (B.107), it follows that wages and vacancies (and, therefore, employment and profits) decrease with t, because $\phi_{tn} \leq 0$. Since Δy_m^s is fixed when choosing t, tightness decreases and, therefore, $\partial U^s/\partial t = p_\theta (\partial \theta_m^s/\partial t) \Delta y_m^s < 0$ for $s \in \{l, h\}$. Reordering terms yields:

$$\begin{aligned}
K \sum_{j\in\mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \left(1 - g_\psi^j\right) \Pi^j(\psi,t) dO_j(\psi) = \; & -\left(L_A^l g_1^l \frac{\partial U^l}{\partial t} + L_A^h g_1^h \frac{\partial U^h}{\partial t}\right) \\
& - K \sum_{j\in\mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \frac{\partial \Pi^j(\psi,t)}{\partial t} \left[g_\psi^j (1-t) + t\right] dO_j(\psi) \\
& - \int \left(\frac{\partial E_m^l}{\partial t}\left(w_m^l - \Delta y_m^l\right) + E_m^l \frac{\partial w_m^l}{\partial t}\right) dm \\
& - \int \left(\frac{\partial E_m^h}{\partial t}\left(w_m^h - \Delta y_m^h\right) + E_m^h \frac{\partial w_m^h}{\partial t}\right) dm \\
& + \frac{\partial K_I}{\partial t}\left(\Pi^{marg} + y_0\right). \qquad \text{(B.48)}
\end{aligned}$$

The right-hand side is positive,[2] which implies that:

$$\sum_{j\in\mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \Gamma_\psi^j (1 - g_\psi^j) dO_j(\psi) > 0, \qquad \text{(B.49)}$$

with $\Gamma_\psi^j = \Pi^j(\psi,t)/\sum_{j\in\mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \Pi^j(\psi,t) dO_j(\psi)$ and $\sum_{j\in\mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \Gamma_\psi^j dO_j(\psi) = 1$, so the profit-weighted average welfare weight on active capitalists is smaller than one. $\square$

[2] Provided the optimal income tax system is not giving employment subsidies that are large enough to encourage increases in the corporate tax rate to decrease employment, to a degree that more than compensates for all the negative welfare effects and fiscal externalities. This would imply that the income tax system is not optimal in the first place.

Proposition 7 The Lagrangian is given by:

$$\begin{aligned}\mathcal{L} &= (L_I^l + L_I^h)G(y_0) + \alpha_l \int_0^{U^l - y_0} G(U^l - c)dF_l(c) + \alpha_h \int_0^{U^h - y_0} G(U^h - c)dF_h(c) \\ &\quad + K_S G\left((1-t)\Pi^S\right) + K_M G\left((1-t)\Pi^M\right) \\ &\quad + \gamma \left[E^l(\overline{w} - \Delta y^l) + E^h(w^h - \Delta y^h) - y_0 + t\left(K_S \Pi^S + K_M \Pi^M\right)\right], \end{aligned} \tag{B.50}$$

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

$$\frac{1}{\gamma}\frac{\partial \mathcal{L}}{\partial \overline{w}} = \frac{\partial U^l}{\partial \overline{w}} L_A^l g_1^l + \frac{\partial E^l}{\partial \overline{w}}(\overline{w} - \Delta y^l) + E^l + K_S \frac{\partial \Pi^S}{\partial \overline{w}}\left((1-t)g_K^S + t\right). \tag{B.51}$$

Since Δy^l is fixed, we have that:

$$\frac{\partial U^l}{\partial \overline{w}} = p_\theta \frac{\partial \theta^l}{\partial \overline{w}} \Delta y^l, \tag{B.52}$$

$$\frac{\partial E^l}{\partial \overline{w}} = p_\theta \frac{\partial \theta^l}{\partial \overline{w}} L_A^l + p^l(\theta^l)\alpha_l f_l(U^l - y_0)\frac{\partial U^l}{\partial \overline{w}}, \tag{B.53}$$

$$\frac{\partial \Pi^S}{\partial \overline{w}} = \left(\phi_n^S - \overline{w}\right) q_\theta \frac{\partial \theta^l}{\partial \overline{w}} v^l - q^l v^l. \tag{B.54}$$

If follows that equation (B.51) can be written as:

$$\begin{aligned}\frac{1}{\gamma}\frac{\partial \mathcal{L}}{\partial \overline{w}} &= p_\theta \frac{\partial \theta^l}{\partial \overline{w}} \Delta y^l L_A^l g_1^l + \left[p_\theta \frac{\partial \theta^l}{\partial \overline{w}} L_A^l + p^l(\theta^l)\alpha_l f_l(U^l - y_0) p_\theta \frac{\partial \theta^l}{\partial \overline{w}} \Delta y^l\right](\overline{w} - \Delta y^l) \\ &\quad + E^l + K_S\left[\left(\phi_n^S - \overline{w}\right) q_\theta \frac{\partial \theta^l}{\partial \overline{w}} v^l - q^l v^l\right]\left((1-t)g_K^S + t\right), \\ &= \varepsilon_{\theta,\overline{w}}^l \frac{\theta^l}{\overline{w}}\left(p_\theta \Delta y^l L_A^l g_1^l + \left[p_\theta L_A^l + p^l(\theta^l)\alpha_l f_l(U^l - y_0) p_\theta \Delta y^l\right](\overline{w} - \Delta y^l)\right. \\ &\quad \left. + K_S\left(\phi_n^S - \overline{w}\right) q_\theta v^l \left((1-t)g_K^S + t\right)\right) + E^l\left(1 - \left((1-t)g_K^S + t\right)\right), \\ &= \varepsilon_{\theta,\overline{w}}^l \theta^l \left(p_\theta L_A^l \left((1-\tau_l)g_1^l + \tau_l\right) + K_S \frac{\left(\phi_n^S - \overline{w}\right)}{\overline{w}} q_\theta v^l \left((1-t)g_K^S + t\right)\right) \\ &\quad + E^l\left(1 + \varepsilon_{L,\overline{w}}^l \tau_l - \left((1-t)g_K^S + t\right)\right), \end{aligned} \tag{B.55}$$

$$\equiv \varepsilon_{\theta,\overline{w}}^l a + E^l\left(1 + \varepsilon_{L,\overline{w}}^l \tau_l - \left((1-t)g_K^S + t\right)\right), \tag{B.56}$$

with $\varepsilon_{\theta,\overline{w}}^l = (\partial\theta^l/\partial\overline{w})/(\overline{w}/\theta^l)$, $\varepsilon_{L,\overline{w}}^l = (\partial L_A^l/\partial\overline{w})/(\overline{w}/L_A^l)$, both holding Δy^l fixed, $\operatorname{sgn}\varepsilon_{\theta,\overline{w}}^l = \operatorname{sgn}\varepsilon_{L,\overline{w}}^l$, a is possibly positive, and $\Delta y^l = (1-\tau_l)\overline{w}$ and $\overline{w} - \Delta y^l = \tau_l \overline{w}$. Then, increasing $\overline{w}$ at the optimal tax allocation is desirable if equation (B.56) is positive, which means that the fiscal externality compensates the tightness distortions (given by potential negative vacancy distortions) and the reduction in profits. When $\varepsilon_{\theta,\overline{w}}^l \to 0$, $\varepsilon_{L,\overline{w}}^l \to 0$, and the condition for equation (B.56) being positive is reduced to $g_K^S < 1$.

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

$$\frac{1}{\gamma}\frac{\partial \mathcal{L}}{\partial \Delta y^l} = \frac{\partial U^l}{\partial \Delta y^l} L_A^l g_1^l + \frac{\partial E^l}{\partial \Delta y^l}(\overline{w} - \Delta y^l) - E^l + K_S \frac{\partial \Pi^S}{\partial \Delta y^l}\left((1-t)g_K^S + t\right). \tag{B.57}$$

When $\overline{w}$ is fixed, we have that:

$$\frac{\partial U^l}{\partial \Delta y^l} = p_\theta \frac{\partial \theta^l}{\partial \Delta y^l}\Delta y^l + p^l(\theta^l), \tag{B.58}$$

$$\begin{aligned}\frac{\partial E^l}{\partial \Delta y^l} &= p_\theta \frac{\partial \theta^l}{\partial \Delta y^l} L_A^l + p^l(\theta^l)\frac{\partial L_A^l}{\partial \Delta y^l} \\ &= p_\theta \frac{\partial \theta^l}{\partial \Delta y^l} L_A^l + \frac{p^l(\theta^l) L_A^l}{\Delta y^l}\varepsilon_{L,\Delta}^l,\end{aligned} \tag{B.59}$$

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

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \Delta y^l}\frac{1}{\gamma} &= \left(p_\theta \frac{\partial \theta^l}{\partial \Delta y^l}\Delta y^l + p^l(\theta^l)\right) L_A^l g_1^l + \left(p_\theta \frac{\partial \theta^l}{\partial \Delta y^l} L_A^l + \frac{p^l(\theta^l) L_A^l}{\Delta y^l}\varepsilon_{L,\Delta}^l\right)(\overline{w} - \Delta y^l) \\ &\quad -E^l + K_S q_\theta \frac{\partial \theta^l}{\partial \Delta y^l} v^l \left(\phi_n^S - \overline{w}\right)\left((1-t)g_K^S + t\right), \\ &= p_\theta \frac{\partial \theta^l}{\partial \Delta y^l} L_A^l \left(\Delta y^l g_1 + \overline{w} - \Delta y^l\right) + E^l \left(g_1 + \varepsilon_{L,\Delta}^l \frac{\overline{w} - \Delta y^l}{\Delta y^l} - 1\right) \\ &\quad + K_S q_\theta \frac{\partial \theta^l}{\partial \Delta y^l} v^l \left(\phi_n^S - \overline{w}\right)\left((1-t)g_K^S + t\right), \\ &= -\frac{p_\theta \theta^l L_A^l \varepsilon_{\theta,\Delta}^l}{\Delta y^l}\left(\Delta y^l g_1 + \overline{w} - \Delta y^l\right) + E^l \left(g_1 + \varepsilon_{L,\Delta}^l \frac{\overline{w} - \Delta y^l}{\Delta y^l} - 1\right) \\ &\quad - \frac{K_S v^l q_\theta \theta^l \left(\phi_n^S - \overline{w}\right)\varepsilon_{\theta,\Delta}^l}{\Delta y^l}\left((1-t)g_K^S + t\right),\end{aligned} \tag{B.60}$$

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

$$\begin{aligned}\frac{1}{\gamma}\frac{\partial \mathcal{L}}{\partial \Delta y^l} &= -p_\theta \theta^l L_A^l \varepsilon_{\theta,\Delta}^l \left(g_1 + \frac{\tau_l}{1-\tau_l}\right) + E^l \left(g_1 + \varepsilon_{L,\Delta}^l \frac{\tau_l}{1-\tau_l} - 1\right) \\ &\quad - \frac{K_S v^l q_\theta \theta^l \left(\phi_n^S - \overline{w}\right)\varepsilon_{\theta,\Delta}^l}{(1-\tau_l)\overline{w}}\left((1-t)g_K^S + t\right).\end{aligned} \tag{B.61}$$

To see whether a negative τ_l is optimal when $\overline{w}$ is fixed at its optimal value, equation (B.61) is evaluated at $\tau_l = 0$. At $\tau_l = 0$, we have that:

$$\frac{\partial \mathcal{L}}{\partial \Delta y^l} = E^l \gamma \left[-g_1 \varepsilon_{\theta,\Delta}^l B + g_1 - 1 + C\varepsilon_{\theta,\Delta}^l \left((1-t)g_K^S + t\right)\right], \tag{B.62}$$

where $B = p_\theta \theta^l / p(\theta^l) \in (0,1)$ since $p(\theta^l) > p_\theta \theta^l$ given Euler's theorem since the matching function has constant returns to scale, and $C = -\left(q_\theta \theta^l \left(\phi_n^S - \overline{w}\right)\right) / \left(\overline{w} q(\theta^l)\right) \in (0,1)$ since $q(\theta^l) > -q_\theta \theta$ 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 $\tau_l = 0$) when $\overline{w}$ is optimal is given by:

$$-g_1 \varepsilon_{\theta,\Delta}^l B + g_1 - 1 + C \varepsilon_{\theta,\Delta}^l \left((1-t) g_K^S + t\right) \quad > \quad 0, \tag{B.63}$$

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

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

$$\frac{\partial \theta^l}{\partial \Delta y^l} = \left(\frac{K_S}{L_A^l}\frac{\partial v^l}{\partial \Delta y^l} - \frac{K_S v^l}{L_A^{l2}}\frac{\partial L_A^l}{\partial \Delta y^l}\right) = \left(\frac{K_S}{L_A^l}\frac{\partial v^l}{\partial \Delta y^l} - \frac{K_S v^l}{L_A^l \Delta y^l}\varepsilon_{L,\Delta}^l\right). \tag{B.64}$$

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

$$\left(\phi_n^S - \overline{w}\right) q_\theta \left(\frac{\partial \theta^l}{\partial w}\frac{\partial \overline{w}}{\partial \Delta y^l} + \frac{\partial \theta^l}{\partial U^l}\frac{\partial U^l}{\partial \Delta y^l}\right) = \eta_{vv}^l \frac{\partial v^l}{\partial \Delta y^l}. \tag{B.65}$$

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

$$\frac{\left(\phi_n^S - \overline{w}\right) q_\theta}{\eta_{vv}^l}\left(\frac{\partial \theta^l}{\partial \Delta y^l}\Delta y^l + \frac{p^l(\theta^l)}{p_\theta}\right) = \frac{\partial v^l}{\partial \Delta y^l}. \tag{B.66}$$

Replacing in equation (B.64) yields:

$$\begin{aligned}
\frac{\partial \theta^l}{\partial \Delta y^l}\left(1 - \frac{\left(\phi_n^S - \overline{w}\right) q_\theta \Delta y^l K_S}{\eta_{vv}^l L_A^l}\right) \quad &= \quad \frac{(\phi_n^S - \overline{w}))q_\theta p^l(\theta^l) K_S}{\eta_{vv}^l L_A^l p_\theta} - \frac{\theta^l}{\Delta y^l}\varepsilon_{L,\Delta}^l, \\
\iff \frac{\partial \theta^l}{\partial \Delta y^l}\left(1 - \frac{\left(\phi_n^S - \overline{w}\right) q_\theta \Delta y^l \theta^l}{\eta_{vv}^l v^l}\right) \quad &= \quad \frac{\theta^l}{\Delta y^l}\left(\frac{\left(\phi_n^S - \overline{w}\right) q_\theta p^l(\theta^l) \Delta y^l}{\eta_{vv}^l v^l p_\theta} - \varepsilon_{L,\Delta}^l\right),
\end{aligned} \tag{B.67}$$

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

Proposition 8 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.45) implies that increasing the minimum wage when the corporate tax rate is optimal is welfare-improving if

$$\frac{\partial U^l}{\partial \overline{w}} L_A^l g_1^l + K_S \frac{\partial \Pi^S}{\partial \overline{w}}\left(g_K^S + t(1 - g_K^S)\right) > 0, \tag{B.68}$$

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, $U^l = p^l(\theta^l)\overline{w} + y_0$ and $U^h = p^h(\theta^h) w^h + y_0$. Then, it follows that $dU^l =$

$p_\theta d\theta^l \overline{w} + p^l(\theta^l)d\overline{w}$ and $dU^h = p_\theta d\theta^h w^h + p^h(\theta^h)dw^h$, so:

$$\begin{aligned}
\frac{\partial U^l}{\partial t} &= p_\theta \frac{\partial \theta^l}{\partial t}\overline{w}, && \text{(B.69)}\\
\frac{\partial U^h}{\partial t} &= p_\theta \frac{\partial \theta^h}{\partial t} w^h + p^h(\theta^h)\frac{\partial w^h}{\partial t}, && \text{(B.70)}\\
\frac{\partial U^l}{\partial \overline{w}} &= p_\theta \frac{\partial \theta^l}{\partial \overline{w}}\overline{w} + p^l(\theta^l), && \text{(B.71)}\\
\frac{\partial U^h}{\partial \overline{w}} &= 0. && \text{(B.72)}
\end{aligned}$$

Making explicit the capital allocation problem, differentiating the first-order conditions of the firms with respect to vacancies we have that

$$\begin{aligned}
\left(\widetilde{\phi}^S_{nk} dk_D - d\overline{w}\right) q^l + \left(\widetilde{\phi}^S_n - \overline{w}\right) q_\theta \left(\frac{\partial \theta^l}{\partial w}\cdot d\overline{w} + \frac{\partial \theta^l}{\partial U^l} dU^l\right) &= \eta^l_{vv} dv^l, && \text{(B.73)}\\
\left(\widetilde{\phi}^M_{nk} dk_D - dw^h\right) q^h + \left(\widetilde{\phi}^M_n - w^h\right) q_\theta \left(\frac{\partial \theta^h}{\partial w} dw^h + \frac{\partial \theta^h}{\partial U^h} dU^h\right) &= \eta^h_{vv} dv^h, && \text{(B.74)}
\end{aligned}$$

where I used that the differential effects driven by the curvature of $\widetilde{\phi}$ are second-order. Recall also from the first-order condition of manufacturing firms with respect to wages that $\left(\widetilde{\phi}^M_n - w^h\right) q_\theta \frac{\partial \theta^h}{\partial w^h} = q^h$. Then, equation (B.74) simplifies to:

$$\widetilde{\phi}^M_{nk} dk_D q^h + \left(\widetilde{\phi}^M_n - w^h\right) q_\theta \frac{\partial \theta^h}{\partial U^h} dU^h = \eta^h_{vv} dv^h. \tag{B.75}$$

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

$$\begin{aligned}
d\theta^l &= \frac{K_S}{L^l_A} dv^l - \frac{K_S v^l}{L^{l2}_A} f_l(U^l - y_0) dU^l, && \text{(B.76)}\\
d\theta^h &= \frac{K_M}{L^h_A} dv^h - \frac{K_M v^h}{L^{h2}_A} f_h(U^h - y_0) dU^h. && \text{(B.77)}
\end{aligned}$$

First, consider the comparative static with respect to t, i.e., equations (B.69) and (B.70). Setting $d\overline{w} = 0$ and combining equations (B.69), (B.73), and (B.76) yields:

$$\frac{\widetilde{\phi}^S_{nk} q^l}{\eta^l_{vv}} \frac{\partial k_D}{\partial t} + \left(\widetilde{\phi}^S_n - \overline{w}\right) \frac{q_\theta}{\eta^l_{vv}} \frac{\partial \theta^l}{\partial U^l} p_\theta \overline{w} \frac{\partial \theta^l}{\partial t} = \frac{\partial \theta^l}{\partial t} \frac{L^l_A}{K_S} \left(1 + \frac{K_S v^l}{L^{l2}_A} f_l(U_l - y_0) p_\theta \overline{w}\right), \tag{B.78}$$

which implies that $\partial \theta^l / \partial t = -a^l_t \varepsilon^S_{k,t}$, with $a^l_t > 0$ provided $\widetilde{\phi}^S_{nk} > 0$. Then, $\partial U^l / \partial t = -a^l_t p_\theta \overline{w} \varepsilon^S_{k,t}$. Then, $\left(\partial U^l / \partial t\right) / \partial \varepsilon^S_{k,t} < 0$.

Also, from equation (B.108), we know that $dw/dv > 0$ for firms that are not constrained by the minimum wage. Then, combining equations (B.70), (B.75), and (B.77) yields:

$$\begin{aligned}
&\frac{\widetilde{\phi}_{nk}^{M} q^h}{\eta_{vv}^h}\frac{\partial k_D}{\partial t} + \left(\widetilde{\phi}_{n}^{M} - w^h\right)\frac{q_\theta}{\eta_{vv}^h}\frac{\partial \theta^h}{\partial U^h} p_\theta w^h \frac{\partial \theta^h}{\partial t} \\
&= \frac{\partial \theta^h}{\partial t}\frac{L_A^h}{K_M}\left(1 + \frac{K_M v^h}{L_A^{h2}} f_h(U_h - y_0) p_\theta w^h\right)\frac{\left(1 - \left(\widetilde{\phi}_{n}^{M} - w^h\right)\frac{q_\theta}{\eta_{vv}^h}\frac{\partial \theta^h}{\partial U^h} p^h(\theta^h)\frac{\partial w^h}{\partial v^h}\right)}{\left(1 - \frac{v^h}{L_A^h} f_h(U^h - y_0)\frac{\partial w}{\partial v}\right)}. \qquad \text{(B.79)}
\end{aligned}$$

The only term that has an 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.[3] Then, this expression implies that $\partial \theta^h / \partial t = -a_t^h \varepsilon_{k,t}^M$, with $a_t^h > 0$ provided $\widetilde{\phi}_{nk}^{M} > 0$. Then:

$$\frac{\partial U^h}{\partial t} \quad = \quad -a_t^h \varepsilon_{k,t}^M \left(p_\theta w^h + p^h(\theta^h)\frac{\partial w^h}{\partial v^h} b\right), \qquad \text{(B.80)}$$

where $b = \frac{L_A^h}{K_M}\left(1 + \frac{K_M v^h}{L_A^{h2}} f_h(U_h - y_0) p_\theta w^h\right)\left(1 - \frac{v^h}{L_A^h} f_h(U^h - y_0)\frac{\partial w}{\partial v}\right) > 0$ 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.71), (B.73), and (B.76) yields:

$$\begin{aligned}
&\widetilde{\phi}_{nk}^{S}\frac{\partial k_D}{\partial \overline{w}} q^l - q^l + \left(\widetilde{\phi}_{n}^{S} - \overline{w}\right) q_\theta \frac{\partial \theta^l}{\partial w} + \left(\widetilde{\phi}_{n}^{S} - \overline{w}\right) q_\theta \frac{\partial \theta^l}{\partial U^l}\left(p_\theta \frac{\partial \theta^l}{\partial \overline{w}}\overline{w} + p^l(\theta^l)\right) \\
&= \eta_{vv}^l \left(\frac{\partial \theta^l}{\partial \overline{w}}\frac{L_A^l}{K_S}\left(1 + \frac{K_S v^l}{L_A^{l2}} f_l(U^l - y_0) p_\theta \overline{w}\right) + \frac{v^l}{L_A^l} f_l(U^l - y_0) p^l(\theta^l)\right). \qquad \text{(B.81)}
\end{aligned}$$

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

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

$$\begin{aligned}
&\frac{\partial U^l}{\partial t} L_A^l g_1^l + \frac{\partial U^h}{\partial t} L_A^h g_1^h + K_S g_K^S \left(-\Pi^S + (1-t)\frac{\partial \Pi^S}{\partial t}\right) \\
&+ K_M g_K^M \left(-\Pi^M + (1-t)\frac{\partial \Pi^M}{\partial t}\right) + K_S \Pi^S + K_M \Pi^M + K_S t \frac{\partial \Pi^S}{\partial t} + K_M t \frac{\partial \Pi^M}{\partial t} \quad = \quad 0. \qquad \text{(B.82)}
\end{aligned}$$

Grouping terms yields:

$$K_S\left(g_K^S + t(1 - g_K^S)\right) \quad = \quad \frac{A + B - \left(g_K^M + t(1 - g_K^M)\right)\gamma_t^M \varepsilon_{k,t}^M}{\gamma_t^S \varepsilon_{k,t}^S}, \qquad \text{(B.83)}$$

[3] 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 $U^h - y_0$.

where $A = K_S \Pi^S (1-g_K^S) + K_M \Pi^M (1-g_K^M)$ and $B = (\partial U^l/\partial t) L_A^l g_1^l + (\partial U^h/\partial t) L_A^h g_1^h$. Replacing equation (B.83) in equation (B.68) gives:

$$\frac{\partial U^l}{\partial \overline{w}} L_A^l g_1^l + \left(-\gamma_{\overline{w}}^S \varepsilon_{k,t}^S + \widetilde{\Pi}_{\overline{w}}^S\right) \frac{A + B - \left(g_K^M + t(1-g_K^M)\right) \gamma_t^M \varepsilon_{k,t}^M}{\gamma_t^S \varepsilon_{k,t}^S} > 0. \tag{B.84}$$

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 that the effect of capital mobility on welfare weights is of second-order, note that:

$$\frac{\partial \mathcal{F}(\varepsilon_{k,t}^S, \varepsilon_{k,t}^M)}{\partial \varepsilon_{k,t}^S} = \frac{\partial \left(\partial U^l/\partial \overline{w}\right)}{\partial \varepsilon_{k,t}^S} L_A^l g_1^l - \frac{\gamma_{\overline{w}}^S}{\gamma_t^S} \frac{\partial B}{\partial \varepsilon_{k,t}^S} + \widetilde{\Pi}_{\overline{w}}^S \frac{\left(\frac{\partial B}{\partial \varepsilon_{k,t}^S} - C\gamma_t^S\right)}{\gamma_t^S \varepsilon_{k,t}^S}, \tag{B.85}$$

where $C = \left(A + B - \left(g_K^M + t(1-g_K^M)\right) \gamma_t^M \varepsilon_{k,t}^M\right) / \gamma_t^S \varepsilon_{k,t}^S > 0$ following equation(B.83), 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:

$$\frac{\partial \mathcal{F}(\varepsilon_{k,t}^S, \varepsilon_{k,t}^M)}{\partial \varepsilon_{k,t}^M} = \frac{\left(-\gamma_{\overline{w}}^S \varepsilon_{k,t}^S + \widetilde{\Pi}_{\overline{w}}^S\right)}{\gamma_t^S \varepsilon_{k,t}^S} \left(\frac{\partial B}{\partial \varepsilon_{k,t}^M} - \left(g_K^M + t(1-g_K^M)\right) \gamma_t^M\right) > 0, \tag{B.86}$$

which is unambiguously positive. □

B.2 Additional results (Section 3)

Parametric example of Proposition 2 Assume the typical capitalist owns a fixed stock of capital, $\overline{k}$, that has to allocate between the domestic firm, k_D, and some foreign investment, k_F, with $k_D + k_F = \overline{k}$. The domestic revenue function is given by $\widetilde{\phi}(l, k_D) = \psi l^\alpha k_D^\beta$, with $\alpha + \beta < 1$, and the after-tax return of the foreign investment is $\widetilde{r}^*$. Then, the firm owner chooses (l, k) to maximize $(1-t)\left[\psi l^\alpha k_D^\beta - wl\right] + \left(\overline{k} - k_D\right)\widetilde{r}^*$. Assuming an interior solution, the first order conditions are given by:

$$l: \qquad \psi \alpha l^{\alpha-1} k_D^\beta = w, \tag{B.87}$$

$$k: \qquad (1-t)\psi \beta l^\alpha k_D^{\beta-1} = \widetilde{r}^*. \tag{B.88}$$

Combining equations (B.87) and (B.88) yields:

$$k_D = l \cdot \Omega(w, t), \tag{B.89}$$

where $\Omega(w,t) = (w(1-t)\beta)/\alpha\widetilde{r}^*$. Define $\phi(l,w,t) = \widetilde{\phi}(l, k_D(l,w,t)) = \psi l^{\alpha+\beta}\Omega(w,t)^\beta$. Solving for the optimal l yields:

$$l^* = \left[\frac{(\alpha+\beta)\psi\left[w(1-t)\beta\right]^\beta}{w\left(\alpha\widetilde{r}^*\right)^\beta}\right]^{\frac{1}{1-\alpha-\beta}} = \left[\frac{(\alpha+\beta)\psi\Omega(w,t)^\beta}{w}\right]^{\frac{1}{1-\alpha-\beta}}. \tag{B.90}$$

Then, noting that $\Omega_w = \Omega/w$ and $\Omega_t = -\Omega/(1-t)$, we have that:

$$\begin{aligned}
\eta_w &= -\frac{\partial l^*}{\partial w}\frac{w}{l^*} \\
&= -\frac{1}{1-\alpha-\beta}\left[\frac{(\alpha+\beta)\psi\Omega^\beta}{w}\right]^{\frac{1}{1-\alpha-\beta}-1}(\alpha+\beta)\psi\frac{\beta\Omega^{\beta-1}\Omega_w w - \Omega^\beta}{w^2}\frac{w}{l}, \\
&= -\frac{l(\alpha+\beta)\psi}{1-\alpha-\beta}\frac{w}{(\alpha+\beta)\psi\Omega^\beta}\frac{\Omega^\beta(\beta-1)}{w^2}\frac{w}{l}, \\
&= \frac{(1-\beta)}{1-\alpha-\beta} > 0, && \text{(B.91)} \\
\eta_t &= -\frac{\partial l^*}{\partial t}\frac{t}{l^*} \\
&= -\frac{1}{1-\alpha-\beta}\left[\frac{(\alpha+\beta)\psi\Omega^\beta}{w}\right]^{\frac{1}{1-\alpha-\beta}-1}\frac{(\alpha+\beta)\psi}{w}\beta\Omega^{\beta-1}\Omega_t\frac{t}{l}, \\
&= \frac{l(\alpha+\beta)\psi}{1-\alpha-\beta}\frac{w}{(\alpha+\beta)\psi\Omega^\beta}\frac{\Omega^\beta\beta}{w(1-t)}\frac{t}{l}, \\
&= \frac{\beta t}{(1-\alpha-\beta)(1-t)} > 0. && \text{(B.92)}
\end{aligned}$$

Then:

$$\frac{\eta_w}{\eta_t} = \frac{(1-\beta)}{1-\alpha-\beta}\frac{(1-\alpha-\beta)(1-t)}{\beta t} = \frac{(1-\beta)(1-t)}{\beta t}, \tag{B.93}$$

and:

$$\eta_t > \eta_w \quad\Leftrightarrow\quad \beta t > (1-\beta)(1-t) \quad\Leftrightarrow\quad \beta > 1-t. \tag{B.94}$$

Also:

$$
\begin{aligned}
\epsilon_w &= -\frac{w}{\Pi}\left(\frac{\partial \phi}{\partial w} - l\right), \\
&= -\frac{w}{\Pi}\left(\frac{\beta l}{\alpha+\beta} - l\right), \\
&= \frac{w}{\Pi}\frac{\alpha l}{\alpha+\beta} > 0, \qquad \text{(B.95)} \\
\epsilon_t &= -\frac{t}{\Pi}\frac{\partial \phi}{\partial t}, \\
&= \frac{t}{\Pi}\frac{\beta w l}{(\alpha+\beta)(1-t)} > 0, \qquad \text{(B.96)}
\end{aligned}
$$

which implies that:

$$
\epsilon_t > \epsilon_w \quad \Leftrightarrow \quad \beta t > \alpha(1-t) \quad \Leftrightarrow \quad \frac{\beta}{\alpha} > \frac{1-t}{t}. \tag{B.97}
$$

Then, consider equation (B.8) with $g_k = 0$:

$$
1 > \frac{tN\Pi}{L\overline{w}}\left(\frac{\eta_w}{\eta_t}(1-\epsilon_t) + \epsilon_w\right). \tag{B.98}
$$

Combining equations (B.93), (B.95), and (B.96) yields:

$$
\begin{aligned}
\epsilon_w - \epsilon_t\frac{\eta_w}{\eta_t} &= \frac{w}{\Pi}\frac{\alpha l}{\alpha+\beta} - \frac{t}{\Pi}\frac{\beta w l}{(\alpha+\beta)(1-t)}\frac{(1-\beta)(1-t)}{\beta t}, \\
&= \frac{wl}{\Pi}\frac{\alpha+\beta-1}{\alpha+\beta}. \qquad \text{(B.99)}
\end{aligned}
$$

Then:

$$
\frac{tN\Pi}{L\overline{w}}\left(\epsilon_w - \epsilon_t\frac{\eta_w}{\eta_t}\right) = \frac{tN\Pi}{L\overline{w}}\frac{wl}{\Pi}\frac{\alpha+\beta-1}{\alpha+\beta} = -\frac{t(1-\alpha-\beta)}{\alpha+\beta}. \tag{B.100}
$$

On the other hand, we have that:

$$
\frac{\Pi}{l\overline{w}} = \frac{l^{\alpha+\beta}\Omega^{\beta} - l\overline{w}}{l\overline{w}} = \frac{ll^{\alpha+\beta-1}\Omega^{\beta} - l\overline{w}}{l\overline{w}} = \frac{1-\alpha-\beta}{\alpha+\beta}, \tag{B.101}
$$

so:

$$
\frac{tN\Pi}{L\overline{w}}\frac{\eta_w}{\eta_t} = \frac{t(1-\alpha-\beta)}{\alpha+\beta}\frac{(1-\beta)(1-t)}{\beta t}. \tag{B.102}
$$

Then:

$$\frac{tN\Pi}{L\overline{w}}\left(\frac{\eta_w}{\eta_t}+\epsilon_w-\epsilon_t\frac{\eta_w}{\eta_t}\right) = \frac{t(1-\alpha-\beta)}{\alpha+\beta}\left(\frac{(1-\beta)(1-t)}{\beta t}-1\right), \\ = \frac{(1-\alpha-\beta)}{\alpha+\beta}\frac{(1-\beta-t)}{\beta}, \tag{B.103}$$

so the reform proposed in Proposition 2 is desirable if:

$$1 > \frac{(1-\alpha-\beta)}{\alpha+\beta}\frac{(1-\beta-t)}{\beta}, \tag{B.104}$$

which is equivalent to:

$$\beta > (1-t)(1-\alpha-\beta) = (1-t)(1-b), \tag{B.105}$$

with $b=\alpha+\beta$.

B.3 Additional results (Section 4)

Firm's problem To simplify exposition, I omit the superscript j from revenue and profit functions. The first-order conditions of firms are given by:

$$w^s : \qquad (\phi_s - w^s)\,\tilde{q}^s_w = \tilde{q}^s, \tag{B.106}$$

$$v^s : \qquad (\phi_s - w^s)\,\tilde{q}^s = \eta^s_v, \tag{B.107}$$

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 $\varepsilon^s = (\partial n^s/\partial w^s)\,(w^s/n^s) = \widetilde{q}^s_w w^s/\widetilde{q}^s$, we can rearrange equation (B.106) 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 $\tilde{q}^{s2} = \eta^s_v\tilde{q}^s_w$. Differentiating and rearranging terms yields:

$$\frac{dw^s}{dv^s} = \frac{\eta^s_{vv}\tilde{q}^s_w}{2\tilde{q}^s\tilde{q}^s_w - \eta^s_v\tilde{q}^s_{ww}} > 0, \tag{B.108}$$

provided $\tilde{q}^s_{ww} < 0$.[4] Moreover, differentiating equation (B.107) yields:

$$\left(d\phi_s - dw^s\right)\tilde{q}^s + \left(\phi_s - w^s\right)\tilde{q}^s_w dw^s \quad = \quad \eta^s_{vv} dv^s. \tag{B.109}$$

Note that:

$$d\phi_s \quad = \quad \phi_{ss}\left(\tilde{q}^s_w dw^s v^s + \tilde{q}^s dv^s\right) + \phi_{s,-s}\cdot\left(\tilde{q}^{-s}_w \cdot dw^{-s}\cdot v^{-s} + \tilde{q}^{-s}\cdot dv^{-s}\right), \tag{B.110}$$

where $-s$ is the other skill type. Replacing equations (B.106) and (B.110) in equation (B.109), yields:

$$\left(\phi_{ss}\cdot\left[\tilde{q}^s_w \cdot dw^s \cdot v^s + \tilde{q}^s \cdot dv^s\right] + \phi_{sj}\cdot\left[\tilde{q}^j_w \cdot dw^j \cdot v^j + \tilde{q}^j \cdot dv^j\right]\right)\cdot\tilde{q}^s = \eta^s_{vv}\cdot dv^s. \tag{B.111}$$

Rearranging terms gives:

$$\frac{dv^s}{dv^{-s}} \quad = \quad \left[\phi_{s,-s}\left(\tilde{q}^{-s}_w \cdot \frac{dw^{-s}}{dv^{-s}} v^{-s} + \tilde{q}^{-s}\right)\right]^{-1}\left[\frac{\eta^s_{vv}}{\psi\tilde{q}^s} - \phi_{ss}\left(\tilde{q}^s_w \frac{dw^s}{dv^s} v^s + \tilde{q}^s\right)\right], \tag{B.112}$$

which, given equation (B.108), implies that $\operatorname{sgn}\left(dv^s/dv^{-s}\right) = \operatorname{sgn}\phi_{s,-s}$.

Notion of equilibrium Define by $\Gamma : \left\{(\psi, j) \in \Psi \times \mathcal{J} : \psi \geq \psi^*_j\right\} \to [\underline{m}, \overline{m}]$ the function that maps active firm types to equilibrium wages that are indexed by sub-markets. The pre-image of Γ does not need to be single-valued (i.e., Γ is not necessarily a one-to-one function) since different (ψ, j) types may generate similar w^s values. Since matching functions and vacancy costs do not vary with j, then it follows from the FOCs of the firm that whenever two firms post the same wage, they also post the same vacancies. Then, given Γ, we can index wages and vacancies by m or (ψ, j) pairs. For each pair (s, m), let $I(s, m) \subset \left\{(\psi, j) \in \Psi \times \mathcal{J} : \psi \geq \psi^*_j\right\}$ be the (possibly singleton) set of (ψ, j) pairs that induce the same wage, w^s. Let ι be the index of elements within that set. An equilibrium of the model is given by:

$$\left\{U^l, U^h, \left\{\psi^*_j\right\}_{j=1}^J, \left\{v^s_m\right\}_{s\in\{l,h\}, m\in[\underline{m},\overline{m}]}, \left\{w^s_m\right\}_{s\in\{l,h\}, m\in[\underline{m},\overline{m}]}, \left\{L^s_m\right\}_{s\in\{l,h\}, m\in[\underline{m},\overline{m}]}\right\}, \tag{B.113}$$

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

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

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 $\mathcal{M}(L, V) = L^\delta V^{1-\delta}$, $\operatorname{sgn}\left[\tilde{q}_{ww}\right] = \operatorname{sgn}\left[\frac{-(1-T'(w))^2}{1-\delta} - T''(w)\right]$, 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 from the concavity and constant returns to scale of the matching function. To see why $\partial^2\theta/\partial w^2 > 0$, recall that $dU = p_\theta d\theta_m(w_m - T(w_m) - y_0) + p_m(1 - T'(w))dw_m$. Setting $dU = 0$ and differentiating again yields:

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

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

where $[\underline{m}, \overline{m}]$ is the mass of active sub-markets that are mapped from the distribution of active types (ψ, j) such that $\psi \in [\psi_j^*, \overline{\psi}]$. The equilibrium objects solve the following equations:

- Firm optimality: $\left(v_m^l, v_m^h, w_m^l, w_m^h\right)$ solve the FOCs of firms of type $(\psi, j) \in \Gamma^{-1}(m)$, taking ψ_j^*, U^l, and U^h as given, for all $m \in [\underline{m}, \overline{m}]$.

- Capitalists' participation constraint: ψ_j^* solves $\Pi^j(\psi_j^*, t) = \xi + y_0$, taking U^l and U^h as given, for all $j \in \mathcal{J}$.

- Across sub-markets equilibrium condition: L_m^s solve $U^s = p_m^s y_m^s + (1 - p_m^s) y_0$, taking ψ_j^*, U^l, U^h, v_m^s, and w_m^s as given, for $s \in \{l, h\}$ and for all $m \in [\underline{m}, \overline{m}]$, where $y_m^s = w_m^s - T(w_m^s)$ and $p_m^s = p^s\left(\theta_m^s\right) = p^s\left(\frac{K v_m^s \sum_{\iota \in \mathcal{I}(s,m)} h_{j_\iota} o_{j_\iota}(\psi_\iota)}{L_m^s}\right)$.

- Workers' participation constraint: U^s solves $\int L_m^s dm = F_s\left(U^s - y_0\right)$, taking L_m^s as given, for $s \in \{l, h\}$.

Efficiency properties of the decentralized equilibrium Without loss of generality, consider a case where there is a unique skill type and a unique j-type, the argument naturally extends given the segmented markets assumption and the fixed portions α_s and σ_j of types. A social planner who 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:

$$\mathcal{V} = K \int_{\psi^*}^{\overline{\psi}} \left[\phi\left(\psi, n\right) - \eta(v_\psi) - \xi\right] dO(\psi) - \alpha \int_0^{c^*} c dF(c), \tag{B.114}$$

and the restrictions are given by:

$$n = q\left(\frac{K v_\psi o(\psi)}{L_\psi}\right) v_\psi, \tag{B.115}$$

$$\int_{\psi^*}^{\overline{\psi}} L_\psi d\psi = \alpha F(c^*), \tag{B.116}$$

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 $\theta_\psi = \left(K v_\psi o(\psi)\right)/L_\psi$. The planner chooses $\{c^*, \psi^*\}$ and $\{v_\psi, L_\psi\}$ to maximize (B.114) subject to (B.115) (matches are endogenous to the number of applicants and vacancies) and (B.116) (the distribution of applicants across firms has to be consistent with the number of active workers). The Lagrangian is given by:

$$\begin{aligned}\mathcal{L} &= K\int_{\psi^*}^{\overline{\psi}}\left[\phi\left(\psi, q\left(\frac{Kv_\psi o(\psi)}{L_\psi}\right)v_\psi\right)-\eta(v_\psi)-\xi\right]dO(\psi) \\ &\quad -\alpha\int_0^{c^*}cdF(c)+\mu\left[\alpha F(c^*)-\int_{\psi^*}^{\overline{\psi}}L_\psi d\psi\right], \qquad \text{(B.117)}\end{aligned}$$

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

$$v_\psi : \qquad \phi_n\left(q_\theta\theta_\psi+q\right)=\eta_v, \qquad \text{(B.118)}$$

$$L_\psi : \qquad -\theta_\psi^2 q_\theta\phi_n=\mu, \qquad \text{(B.119)}$$

$$c^* : \qquad -\alpha c^* f(c^*)+\mu\alpha f(c^*)=0. \qquad \text{(B.120)}$$

Equation (B.120) implies that $\mu=c^*$. Using that and combining equations (B.118) and (B.119) yields:

$$q\phi_n-\frac{c^*}{\theta_\psi} = \eta_v. \qquad \text{(B.121)}$$

To assess the efficiency of vacancy posting decisions and application decisions, I check whether equation (B.121) is consistent with the decentralized equilibrium. In the absence of taxes, the threshold for workers' entry is given by $U=p(\theta_\psi)w_\psi$, which holds for any ψ. We also know, from the properties of the matching function, that $p(\theta_\psi)=\theta_\psi q(\theta_\psi)$. Replacing in equation (B.121) yields $q\left(\phi_n-w_\psi\right)=\eta_v$, which coincides with the decentralized first order condition of the firms for vacancies (see equation (B.107)). Then, the decentralized equilibrium is efficient in terms of vacancy postings and applications.

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

$$\psi^* : \qquad -K\left(\phi\left(\psi^*, q\left(\theta_{\psi^*}\right)v_{\psi^*}\right)-\eta\left(v_{\psi^*}\right)-\xi\right)o\left(\psi^*\right)-\mu L_{\psi^*}=0. \qquad \text{(B.122)}$$

Equation (B.122) can be written as:

$$\phi\left(\psi^*, q\left(\theta_{\psi^*}\right)v_{\psi^*}\right)-\eta\left(v_{\psi^*}\right)-\frac{\mu L_{\psi^*}}{Ko\left(\psi^*\right)} = \xi. \qquad \text{(B.123)}$$

Note that $\mu L_{\psi^*}/Ko\left(\psi^*\right)=c^*v_{\psi^*}/\theta_{\psi^*}$. Then, $c^*=p\left(\theta_{\psi^*}\right)w_{\psi^*}$ and $p(\theta_{\psi^*})=\theta_{\psi^*}q(\theta_{\psi^*})$ imply that $c^*v_{\psi^*}/\theta_{\psi^*}=w_{\psi^*}q(\theta_{\psi^*})v_{\psi^*}$, which implies that equation (B.122) is equivalent to $\Pi\left(\psi^*\right)=\xi$, which coincides with the decentralized equilibrium. Therefore, the decentralized equilibrium is efficient. □

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 (the extension of the problem discussed in Section 3) and an effort-allocation problem. To simplify notation, I omit superscripts j.

Regarding the capital-allocation problem, assume that the production function, $\widetilde{\phi}$, depends on capital, k, and labor, (n^l, n^h), 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 $(1-t)\widetilde{\Pi}(k_D, \psi)$, 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 $k_A \widetilde{r}^*$, where $\widetilde{r}^*$ is the after-tax return of capital abroad. Capitalists choose k_D to maximize $(1-t)\widetilde{\Pi}(k_D, \psi) + \left(\overline{k} - k_D\right) \widetilde{r}^*$. The first-order condition is given by $(1-t)\widetilde{\Pi}_k = \widetilde{r}^*$, which characterizes the optimal capital invested domestically, k^*_D, as a function of ψ and t.[5] Then, $k^*_D = k_D(\psi, t)$, so $\widetilde{\phi}(k_D(\psi, t), n^l, n^h, \psi) = \phi(\psi, n^l, n^h, t)$ and $\widetilde{\Pi}(k_D(\psi, t), \psi) = \Pi(\psi, t)$. If capitalists have no investment opportunities abroad or transportation costs are large (which would be analogous to $r^* \to 0$), then $k_D = \overline{k}$ and the revenue 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 $\overline{w}$, with ϕ_w ambiguous depending on the structural patterns of capital-labor substitution embedded in $\widetilde{\phi}$.

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 $\varepsilon_{k,t} = -(\partial k_D/\partial t)(t/k_D)$, and the elasticity of domestic capital to changes in the minimum wage (when it binds) by $\varepsilon_{k,\overline{w}} = -(\partial k_D/\partial \overline{w})(\overline{w}/k_D)$. $\varepsilon_{k,t}$ 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, $\varepsilon_{k,\overline{w}} = a\varepsilon_{k,t}$, with $a > 0$.[6]

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

$$\frac{\partial \Pi\left(\psi, t, \overline{w}\right)}{\partial t} = \frac{\widetilde{\Pi}\left(k_D(\psi, t, \overline{w}), \psi, \overline{w}\right)}{\partial t} = \widetilde{\Pi}_k \frac{\partial k_D}{\partial t} = -\gamma_t \varepsilon_{k,t} < 0, \tag{B.124}$$

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

$$\frac{\partial \Pi\left(\psi, t, \overline{w}\right)}{\partial \overline{w}} = \frac{\partial \widetilde{\Pi}\left(k_D(\psi, t, \overline{w}), \psi, \overline{w}\right)}{\partial \overline{w}} = \widetilde{\Pi}_k \frac{\partial k_D}{\partial \overline{w}} + \widetilde{\Pi}_{\overline{w}} = -\gamma_{\overline{w}} \varepsilon_{k,t} + \widetilde{\Pi}_{\overline{w}} < 0, \tag{B.125}$$

where $\gamma_{\overline{w}} = \widetilde{\Pi}_k k_D a/\overline{w} > 0$. 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

[5]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 $\widetilde{\Pi}_{kk} = \partial^2 \widetilde{\Pi}(k_D, \psi)/\partial k_D^2 < 0$.

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

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

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

minimum wages on pre-tax profits is driven by both the reduction in capital and the direct effect on labor costs. Expressions depicted in equations (B.124) and (B.125) can be easily transformed to profits elasticities, (ϵ_t, ϵ_w), as defined in Section 3.

As an alternative microfoundation, assume that the structural production function, $\widetilde{\phi}$, depends on the managerial effort of the capitalist, e, as well as labor inputs, (n^l, n^h), with $\widetilde{\phi}_e > 0$ and $\widetilde{\phi}_{ee} < 0$. If capitalists exert effort e, they get after-tax profits $(1-t)\widetilde{\Pi}(e, \psi)$, 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 $(1-t)\widetilde{\Pi}_e = c_e$, which characterizes the optimal effort, e^*, as a function of ψ and t. Then, $e^* = e(\psi, t)$, so $\widetilde{\phi}(e(\psi,t), n^l, n^h, \psi) = \phi(\psi, n^l, n^h, t)$ 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.

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 (B.106) and (B.107) for $s = \{l, h\}$) hold for firms for which the minimum wage is not binding, while (B.106) no longer holds for firms for which the minimum wage is binding. 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. As in previous appendix subsections, I omit superscript j to simplify notation.

In both cases, it is easier to work with equation (B.107) for $s = l$. When the minimum wage is not binding, totally differentiating the first-order condition yields:

$$\begin{aligned}\left(\left[\phi_{ll}\left(q_\theta^l d\theta^l v^l + q^l dv^l\right) + \phi_{lh}\left(q_\theta^h d\theta^h v^h + q^h dv^h\right)\right] - dw^l\right) q^l & \\ +(\phi_l - w^l) q_\theta^l d\theta^l &= \eta_{vv}^l dv^l, \end{aligned} \tag{B.126}$$

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

$$dw^l \left[\frac{dv^l}{dw^l}\left(\eta_{vv}^l - \phi_{ll} q^{l2} - \phi_{lh} q^h q^l \frac{dv^h}{dv^l}\right) + q^l\right] = d\theta^l q_\theta^l \left[(\phi_l - w^l) + \phi_{ll} v^l q^l\right] + d\theta^h q_\theta^h \phi_{lh} q^l. \tag{B.127}$$

Note that the sign and magnitude of $dw^l/d\overline{w}$ depend 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, when the minimum wage is binding, totally differentiating the first-order condition yields:

$$\left(\left[\phi_{ll}\left(q^l_\theta d\theta^l v^l + q^l dv^l\right) + \phi_{lh}\left(q^h_\theta d\theta^h v^h + q^h dv^h\right)\right] - d\overline{w}\right) q^l$$
$$+(\phi_l - \overline{w})q^l_\theta d\theta^l \quad = \quad \eta^l_{vv} dv^l, \tag{B.128}$$

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

$$\frac{dv^l}{d\overline{w}}\left(\eta^l_{vv} - \phi_{ll}q^{l2} - \phi_{lh}q^h q^l \frac{dv^h}{dv^l}\right) \quad = \quad \frac{d\theta^l}{d\overline{w}} q^l_\theta \left[(\phi_l - \overline{w}) + \phi_{ll} v^l q^l\right]$$
$$+\frac{d\theta^h}{d\overline{w}} q^h_\theta \phi_{lh} q^l - q^l. \tag{B.129}$$

The sign and magnitude depend 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 (conditional on j, omitted here for simplicity).

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 firms for which the minimum wage is binding:

$$\frac{d\Pi(\psi,t)}{d\overline{w}} = \frac{\partial\Pi(\psi,t)}{\partial\overline{w}} = q^l_\theta \frac{\partial\theta^l}{\partial\overline{w}} v^l(\phi_l - \overline{w}) - v^l q^l, \tag{B.130}$$

where $\partial\theta^l/\partial\overline{w} = \partial\theta^l/\partial w + (\partial\theta^l/\partial U^l)\cdot(\partial U^l/\partial\overline{w})$. This effect is possibly negative given that the first-order condition with respect to low-skill wages holds with inequality and is stronger for less productive firms. When $\overline{w} = w^l$, the envelope theorem cancels out part of the effect, although profits still can be affected by general equilibrium effects through U^l. That is why when the minimum wage is not binding:

$$\frac{d\Pi(\psi,t)}{d\overline{w}} = \frac{\partial\Pi(\psi,t)}{\partial\overline{w}} = q^l_\theta \frac{\partial\theta^l}{\partial U^l}\frac{\partial U^l}{\partial\overline{w}} v^l(\phi_l - w^l), \tag{B.131}$$

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

More intuition on the SWF The average social value of the expected utility of active workers of skill s is $\int_0^{U^s - y_0} \omega_L G(U^s - c) d\tilde{F}_s(c)$, where $\tilde{F}_s(c) = F_s(c)/F_s(U^s - y_0)$. Then, the total value is given by $L^s_A \int_0^{U^s - y_0} \omega_L G(U^s - c) d\tilde{F}_s(c)$, which yields the expressions of equation (16) noting that $L^s_A = \alpha_s F(U^s - y_0)$. The average social value of the utility of capitalists of type j is $\int_{\psi^*_j}^{\overline{\psi}} \omega_k G((1-t)\Pi^j(\psi,t) - \xi) d\tilde{O}_j(\psi)$, with $\tilde{O}_j(\psi) = O_j(\psi)/(1 - O_j(\psi^*))$. Then, their total value is $K^j_A \int_{\psi^*_j}^{\overline{\psi}} \omega_k G((1-t)\Pi^j(\psi,t) - \xi) d\tilde{O}_j(\psi) = K \int_{\psi^*_j}^{\overline{\psi}} \omega_k G((1-t)\Pi^j(\psi,t) - \xi) dO_j(\psi)$, since $K^j_A = 1 - O_j(\psi^*_j)$. Aggregating across types yields $K \sum_{j\in\mathcal{J}} \sigma_j \int_{\psi_j^*}^{\overline{\psi}} \omega_k G((1-t)\Pi^j(\psi,t)-\xi)dO_j(\psi)$.

Employment effects and workers' welfare For simplicity, assume away taxes. In equilibrium, $U^s = p_m^s w_m^s$. Multiplying by L_m^s at both sides and integrating over m yields $L_A^s U^s = \int E_m^s w_m^s dm$, where $E_m^s = L_m^s p_m^s$ is the mass of employed workers of skill s in sub-market m. Differentiating gives:

$$\frac{dU^s}{d\overline{w}}\left(L_A^s + U^s \alpha_s f_s(U^s)\right) = \int \left(\frac{dE_m^s}{d\overline{w}} w_m^s + E_m \frac{dw_m^s}{d\overline{w}}\right) dm, \tag{B.132}$$

where I used $L_A^s = \alpha_s F_s(U^s)$. 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, $\int (dE_m^s/d\overline{w}) w_m^s dm$, 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.

Equation (19) with taxes In the case with taxes, U^s is equal to the average pre-tax wage of active workers including the unemployed net of their average tax liabilities. To see why, recall that $U^s = p_m^s y_m^s + (1-p_m^s)y_0$. Multiplying both sides by the sub-market mass of applicants, L_m^s, and integrating over m, gives:

$$U^s = \frac{\int E_m^s(w_m^s - T(w_m^s) - y_0)dm}{L_A^s} + y_0 = \frac{\int E_m^s w_m^s dm}{L_A^s} - \frac{\int E_m^s(T(w_m^s)+y_0)dm}{L_A^s} + y_0, \tag{B.133}$$

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

$$\frac{dU^s}{d\overline{w}} = \frac{d}{d\overline{w}}\left(\frac{\int E_m^s w_m^s dm}{L_A^s}\right) - \frac{d}{d\overline{w}}\left(\frac{\int E_m^s(T(w_m^s)+y_0)dm}{L_A^s}\right). \tag{B.134}$$

The first term represents the change in the average pre-tax wage among active workers (see equation (19)). The second term represents the change in average tax liabilities net of transfers among active workers.

B.4 Additional discussions

I first discuss the omission of price and productivity effects. I then discuss the omission of dynamics, informality, and non-wage amenities.

- *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; Leung, 2021; Ashenfelter and Jurajda, 2022; Renkin et al., 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 if they 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, although a deeper exploration is needed.

- *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., 2022; Emanuel and Harrington, 2022; Ku, 2022; Ruffini, Forthcoming) 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.Importantly, the main policy results depend on reduced-form profit elasticities that are robust to productivity increases.

- *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. (2023) 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.

- *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).

- *Non-wage amenities*: The model requires workers to not have preferences over (ψ, j) beyond its effect on wages and vacancies. Hence, the model cannot accommodate heterogeneous non-wage amenities. For evidence on their relevance, see Bonhomme and Jolivet (2009), Mas and Pallais (2017), Sorkin (2018), Taber and Vejlin (2020), Le Barbanchon et al. (2021), Lamadon et al. (2022), Lindenlaub and Postel-Vinay (2022), Sockin (2022), Maestas et al. (2023), Roussille and Scuderi (2023), and Jäger et al. (Forthcoming). Amenities can affect the policy analysis for 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). Such effects could attenuate potential welfare gains to workers after minimum wage hikes.

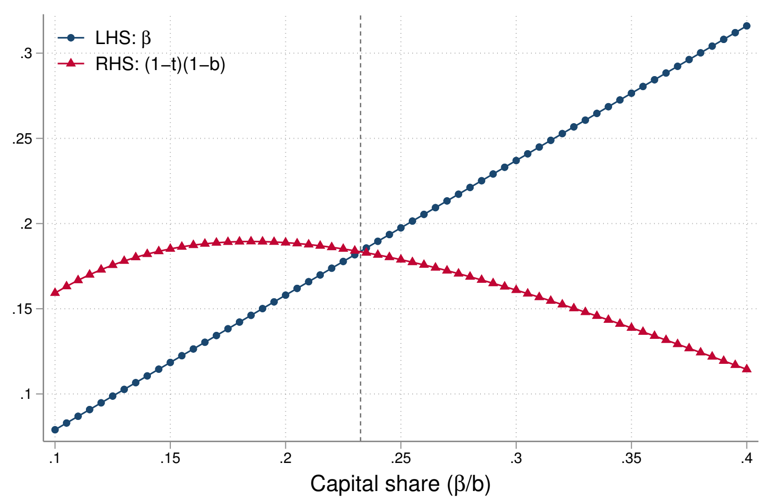
C Simulation appendix

C.1 Section 3

The following simulation is not meant to represent a real-world calibration but rather it is a proof of concept that the condition depicted in Proposition 2 is economically reasonable.

In terms of functional forms and parameters, I follow the parametric model described in Appendix B.2 and define $G(x) = x^{1-\zeta}/(1-\zeta)$, and $F(c,\lambda) = 1 - \exp(-\lambda c)$. I set $(\omega_L, \omega_K) = (1,0)$, $b = 0.79$ (Lamadon et al., 2022), $N = 1$, $\widetilde{r}^* = 0.042$ (see Appendix C.2), $\psi = 15$, $\zeta = 1$, and $\lambda = 0.04$. Then, I define a grid of β that defines a grid of capital shares β/b and, for each β/b, I solve for the optimal tax system by choosing (T_0, T_1, t, γ) to solve the non-linear system of equations determined by equations (B.24)-(B.25)-(B.26) and the budget constraint. Note that, given b, β also determines α. Then, for each simulation, I compute both the LHS and the RHS of equation (8) to determine whether an increase in the minimum wage above the market level is a desirable complement of the optimal tax system. Figure C.1 presents the results. In this particular example, the RHS displays a concave behavior which is explained by the convex relationship between t^* and β. In this example, whenever the capital share is above 23%, then using a binding minimum wage is desirable.

Figure C.1: Numerical example: Condition for minimum wage desirability



C.2 Section 4

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 $\widetilde{\phi}^j(k,n) = \psi^j k^{\beta^j_k} n^{\beta^j_n}$, for $j \in \{S,M\}$. The vacancy cost functions are given by $\eta^s(v) = \frac{\kappa_{0s} v^{1+\kappa_{1s}}}{1+\kappa_{1s}}$, for $s \in \{l,h\}$. The outside option is uniformly distributed with upper bound λ_s, for $s \in \{l,h\}$.

Calibration To simulate the model, I need to impute parameter values. For a subset of parameters, I take values from the related literature (*calibrated parameters*). The rest are chosen to match empirical moments (*estimated 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.

Table C.1: Calibrated 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. (2023)
$r^*(1-t^*)$ $(I = M)$	0.052	Piketty and Zucman (2014), Bachas et al. (2023)
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.

Calibrated parameters: Table C.1 summarizes the calibrated 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 $\beta_n^j = bLS$ and $\beta_k^I = b(1 - LS)$, for $j \in \{S, M\}$, 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). Since the global capital tax rate is around 30% (Bachas et al., 2023), this implies that the global after-tax return is around 4.2%. To accommodate differential capital mobility based on differential transportation costs paid in units of investment returns, I assume that the global after-tax return is 3.2% for $j = S$ and 5.2% for $j = M$. The estimation below is done assuming a fixed tax system, which I define as follows. I use $t = 20\%$, which is the statutory corporate tax rate. For the income tax system, 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.

Estimated parameters: Table C.2 summarizes the moments matched and the estimated parameters.

Table C.2: Estimated 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

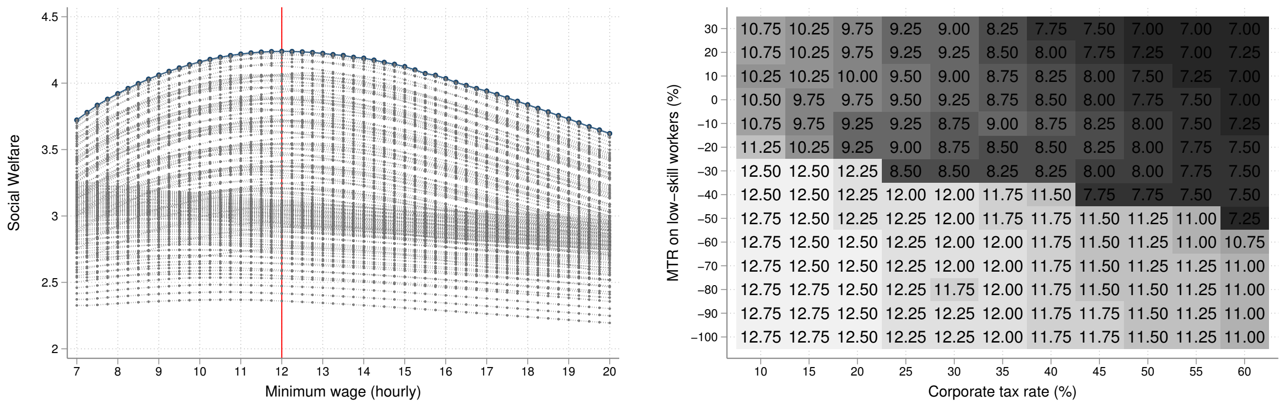
Parameter	Value	Parameter	Value
δ_{0l}	0.85	δ_{0h}	0.92
δ_{1l}	0.51	δ_{1h}	0.79
λ_l	15.62	λ_h	77.97
K_S	0.038	K_M	0.008
ψ^S	31.46	ψ^M	36.78
κ_{0l}	0.727	κ_{0h}	0.239
κ_{1l}	0.987	κ_{1h}	1.233

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

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, $\{\psi^S, \psi^M, \kappa_{0l}, \kappa_{0h}, \kappa_{1l}, \kappa_{1h}\}$.

Exercise and results I 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, 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 C.2 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 hourly minimum wage (i.e., the minimum wage that

Figure C.2: Simulation results



(a) Social welfare versus minimum wage

(b) Optimal hourly minimum wage given taxes

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

maximizes social welfare) for each of the 154 tax combinations.

There are three messages from Panel (a). First, given taxes, social welfare generally 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 ($12 dollars the hour) is far from the market wage ($7 dollars the hour). The exact number 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 an hour, the tax system consists of a substantial 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. 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, 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 in their values, the planner benefits from redistributing profits using both instruments, rather than just using larger corporate tax rates or large minimum wages.

D Sufficient statistics analysis appendix

A modified version of equation (20) suggests that increasing the minimum wage is welfare-improving if:

$$\frac{d\log U^l}{d\overline{w}}U^l L_A^l g_1^l + \frac{d\log \Pi^S}{d\overline{w}}\Pi^S K_A^S g_K^S + \text{Fiscal effects} \quad > \quad 0, \tag{D.1}$$

where I omit the high-skill workers component – because $dU^h/d\overline{w}$ is estimated to be zero (see Appendix A) – and denote as Π^S the average profit per establishment in exposed industries and g_K^S its corresponding marginal welfare weight. I work with average profits and omit firm-level heterogeneity (beyond exposed and non-exposed industries) since I have estimates based on aggregate data. The fiscal effects component considers both worker- and capitalist-level fiscal externalities.

$U^l L_A^l$ equals the sum of total pre-tax income of low-skill workers plus total income maintenance transfers,[7] so the first term of equation (D.1) can be written as $(\epsilon_{U_{PT}^l}\text{PTW} + \epsilon_{IT}\text{IT})g_1^l$, 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 (D.1) can be written as $\epsilon_{\Pi^S}\text{PTP}(1-t)g_K^S$, where ϵ_{Π^S} is the profit elasticity on exposed industries and PTP accounts for total annual pre-tax profits of exposed industries. Finally, fiscal effects can be written as $-\epsilon_{IT}\text{IT} + \epsilon_{\Pi^S}t\text{PTP}$, where I omit extensive margin responses on capitalists given the zero estimate on number of establishments presented in Appendix A. Collecting terms, I can write equation (D.1) as:

$$(\epsilon_{U_{PT}^l}\text{PTW} + \epsilon_{IT}\text{IT})g_1^l + \epsilon_{\Pi}\text{PTP}(1-t)g_K^S - \epsilon_{IT}\text{IT} + \epsilon_{\Pi^S}t\text{PTP} \quad > \quad 0. \tag{D.3}$$

Values for $\{\epsilon_{U_{PT}^l}, \epsilon_{IT}, \epsilon_{\Pi^S}\}$ can be taken from Tables 2. I focus on the estimates using the stricter set of time fixed effects which yield $\{\epsilon_{U_{PT}^l}, \epsilon_{IT}, \epsilon_{\Pi^S}\} = \{0.015, -0.050, -0.063\}$. Likewise, values for $\{\text{PTW}, \text{IT}, \text{PTP}\}$ are 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 predict the effects of small minimum wages today.[8] Consequently, I impute values for t using two assumptions.

[7]Note from equation (B.133) that

$$U^l L_A^l \quad = \quad \int E_m^l w_m^l dm - \int E_m^l T(w_m^l)dm + y_0 L_A^l \rho^l, \tag{D.2}$$

that is, $U^l L_A^l$ equals total pre-tax income plus the net tax liabilities which are composed of 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 given the estimated zero effects on medical transfers and gross federal income tax payments, as shown in Appendix A.

[8]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.

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).[9] 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 (D.3): the SMWWs, $\{g_1^l, g_K^S\}$. I calibrate g_K^S and then back up the welfare weight on low-skill workers that makes equation (D.3) 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 a 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, $g_1^l/g_K^S = (U^l/(1-t)\Pi^S)^{-\zeta}$, so g_K^S can be jointly determined with g_1^{l*}.[10]

Results Table D.1 summarizes the results. Each cell reports g_1^{l*} for a different permutation of the 16 calibration choices discussed above. Table D.1 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 requires a welfare weight of workers of 1.95 or more to justify the policy. That is, if the planner does not care about inequality between low-skill workers and exposed 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.

This exercise highlights the importance of including redistributive preferences in the analysis. Even

[9] 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.

[10] 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:

$$g_1^{l*} = \frac{-\left(\epsilon_\Pi \mathrm{PTP}(1-t) - \epsilon_{IT}\mathrm{IT} + \epsilon_{\Pi^S} t\mathrm{PTP}\right)}{\epsilon_{U^l_{PT}}\mathrm{PTW} + \epsilon_{IT}\mathrm{IT}}. \tag{D.4}$$

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

$$g_1^{l*} = \frac{-\left(-\epsilon_{IT}\mathrm{IT} + \epsilon_{\Pi^S} t\mathrm{PTP}\right)}{\epsilon_{U^l_{PT}}\mathrm{PTW} + \epsilon_{IT}\mathrm{IT} + \epsilon_\Pi \mathrm{PTP}(1-t)\omega(\zeta)^{-1}}. \tag{D.5}$$

if total output falls, the incorporation of distributional concerns makes the case for the minimum wage unambiguously favorable.[11] Intuitively, the empirical analysis shows that minimum wages benefit low-skill workers, hurt firm owners in the exposed industries, and generate fiscal savings in transfers and 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. Also, the net fiscal effect is very small relative to baseline incomes: the net fiscal effect never represents more than 0.5% of total pre-tax incomes of low-skill workers. Then, in the absence of preferences for redistribution, the policy is not far from 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 fact makes a positive case for the minimum wage because the distinction between winners and losers is aligned with the social planner's preferences.

Table D.1: Welfare effects of minimum wage reforms under fixed taxes (g_1^{l*})

	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	1.95	0.19	0.14	0.12	1.97	0.00	0.00	0.00
Effective t	1.95	0.00	0.00	0.00	1.97	0.00	0.00	0.00

[11]The degree of concavity of the social welfare function (ζ) does not affect the analysis because 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 equation (D.3) manifest even when the concavity of the social welfare function is moderate.

Appendix Bibliography

Aaronson, Daniel, Eric French, Isaac Sorkin, and Ted To, "Industry dynamics and the minimum wage: A putty-clay approach," *International Economic Review*, 2018, *59* (1), 51–84.

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

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

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

Bachas, Pierre, Matthew Fisher-Post, Anders Jensen, and Gabriel Zucman, "Globalization and Factor Income Taxation," *Working Paper*, 2023.

Baker, Andrew, David Larcker, and Charles Wang, "How Much Should We Trust Staggered Difference-In-Differences Estimates?," *Journal of Financial Economics*, 2022, *144* (2), 370–395.

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

Berger, David, Kyle Herkenhoff, and Simon Mongey, "Minimum wages and welfare," *Working Paper*, 2022.

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

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

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

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The effect of minimum wages on low-wage jobs," *The Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

_ , _ , _ , and David Zentler-Munro, "Seeing Beyond the Trees: Using Machine Learning to Estimate the Impact of Minimum Wages on Labor Market Outcomes," *Journal of Labor Economics*, 2022, *40* (S1), S203–S247.

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

_ **, Lisa Kahn, and Jonathan Meer**, "The minimum wage, fringe benefits, and worker welfare," *Working Paper*, 2018.

Coviello, Decio, Erika Deserranno, and Nicola Persico, "Minimum wage and individual worker productivity: Evidence from a large US retailer," *Journal of Political Economy*, 2022, *130* (9), 2227–2500.

de Chaisemartin, Clément and Xavier D'Haultfœuille, "Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey," *Econometrics Journal*, 2023, *26* (3), C1–C30.

Dube, Arindrajit, Daniele Girardi, Oscar Jorda, and Alan Taylor, "A local projections approach to difference-in-differences event studies," *Working Paper*, 2023.

_ **, William Lester, and Michael Reich**, "Minimum wage shocks, employment flows, and labor market frictions," *Journal of Labor Economics*, 2016, *34* (3), 663–704.

Emanuel, Natalia and Emma Harrington, "Firm frictions and the payoffs of higher pay: Labor supply and productivity responses to a voluntary firm minimum wage," *Working Paper*, 2022.

Gardner, John, "Two-stage differences in differences," *Working Paper*, 2021.

Gittings, Kaj and Ian Schmutte, "Getting handcuffs on an octopus: Minimum wages, employment, and turnover," *ILR Review*, 2016, *69* (5), 1133–1170.

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

Harasztosi, Péter and Attila Lindner, "Who Pays for the minimum Wage?," *The American Economic Review*, 2019, *109* (8), 2693–2727.

Hurst, Erik, Patrick Kehoe, Elena Pastorino, and Thomas Winberry, "The Macroeconomic Dynamics of Labor Market Policies," *Working Paper*, 2023.

Jäger, Simon, Christopher Roth, Nina Roussille, and Benjamin Schoefer, "Worker Beliefs About Outside Options," *Quarterly Journal of Economics*, Forthcoming.

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

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

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

Lindenlaub, Ilse and Fabien Postel-Vinay, "The Worker-Job Surplus," *Working Paper*, 2022.

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

Maestas, Nicole, Kathleen J Mullen, David Powell, Till Von Wachter, and Jeffrey B Wenger, "The value of working conditions in the United States and implications for the structure of wages," *American Economic Review*, 2023, *113* (7), 2007–2047.

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

Mayneris, Florian, Sandra Poncet, and Tao Zhang, "Improving or disappearing: Firm-level adjustments to minimum wages in China," *Journal of Development Economics*, 2018, *135*, 20–42.

Meghir, Costas, Renata Narita, and Jean-Marc Robin, "Wages and informality in developing countries," *American Economic Review*, 2015, *105* (4), 1509–46.

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

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

_ , Emmanuel Saez, and Gabriel Zucman, "Distributional national accounts: Methods and estimates for the United States," *The Quarterly Journal of Economics*, 2018, *133* (2), 553–609.

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

Riley, Rebecca and Chiara Rosazza Bondibene, "Raising the standard: Minimum wages and firm productivity," *Labour Economics*, 2017, *44*, 27–50.

Robinson, Joan, *The economics of imperfect competition*, Springer, 1933.

Roth, Jonathan, Pedro Sant'Anna, Alyssa Bilinski, and John Poe, "What's trending in difference-in-differences? A synthesis of the recent econometrics literature," *Journal of Econometrics*, 2023, *235* (2), 2218–2244.

Roussille, Nina and Benjamin Scuderi, "Bidding for talent: A test of conduct in a high-wage labor market," *Working Paper*, 2023.

Ruffini, Krista, "Higher wages, service quality, and firm profitability: Evidence from nursing homes and minimum wage reforms," *The Review of Economics and Statistics*, Forthcoming.

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

Sockin, Jason, "Show me the amenity: Are higher-paying firms better all around?," *Working Paper*, 2022.

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

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

Taber, Christopher and Rune Vejlin, "Estimation of a Roy/search/compensating differential model of the labor market," *Econometrica*, 2020, *88* (3), 1031–1069.

Vaghul, Kavya and Ben Zipperer, "Historical state and sub-state minimum wage data," *Washington Center for Equitable Growth*, 2016.

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