Setting Minimum Wages as Regulating a Monopsonist with Unknown Productivity

James Stratton*

July 1, 2026

Abstract

I view minimum wage-setting as regulating a monopsonist with privately known productivity. This approach gives four insights into optimal minimum wages. First, the optimal minimum wage falls as productivity becomes more dispersed. Second, conditioning the minimum wage on industry or location raises the optimal minimum wage on average. Third, smaller firms should face lower minimum wages. Fourth, unlike under perfect competition, whether minimum wages should be accompanied by employment subsidies or taxes depends on the regulator's redistributive preferences: a regulator who values only worker surplus taxes employment, even in the presence of the optimal minimum wage.

*Email: jstratton@g.harvard.edu. I am grateful for discussions with Josh Angrist, Raj Chetty, Cameron Deal, Ed Glaeser, Oliver Hart, Larry Katz, Eric Maskin, Nicolaj Thor, and workshop participants at Harvard and Brown. I am also grateful to have received financial support from the Kenneth C. Griffin Fund. The code and data to reproduce the analysis in Online Appendix C are available at https://github.com/JamesStratton/regulating-a-monopsonist-with-unknown-productivity.

1. Introduction

How should regulators set minimum wages? In competitive labor markets, minimum wages prohibit some positive-surplus matches (an efficiency loss), but transfer surplus from firms to workers (an equity gain for a planner with redistributive preferences) (Stigler 1946; Lee and Saez 2012). The modern case for minimum wages offers an alternative rationale: if low-wage labor markets are monopsonistic, minimum wages generate efficiency gains by reducing deadweight loss from monopsony wage-setting (Manning 2011; Dube and Lindner 2024; Kline 2025).[1]

On this latter view, there is a sense in which the challenge in setting minimum wages is informational. Consider a "company town" with a single employer whose only input is labor, with constant productivity θ. The competitive wage is θ, leaving the monopsonist no profit. Absent regulation, the monopsonist chooses a wage strictly below θ, generating positive profit but deadweight loss. If the regulator knew θ, she could restore first-best employment through a minimum wage equal to θ.[2] The difficulty in practice is that productivity θ is *private information*, unobserved by the regulator; equivalently, the regulator must set a *single* minimum wage, which applies *uniformly* across monopsonists with different productivity.

Through this lens, minimum wage-setting is a problem of regulating a monopsonist of unknown productivity, closely related to the monopoly regulation problem of Baron and Myerson (1982). I show that applying the Baron-Myerson approach to minimum wage-setting yields new insights into optimal minimum wages. A CRS monopsonist's labor productivity is drawn from a known distribution, but its realization is private information. The regulator chooses a minimum wage policy to maximize a weighted sum of worker and firm surplus; the monopsonist chooses a wage to maximize profit, subject to the minimum wage policy. I consider three policy regimes: a uniform minimum wage; a minimum wage conditional on exogenous characteristics (*e.g.,* industry or location); and a minimum wage that is differentiated by firm size (*e.g.,* exemptions for small firms).

Main results. Consider first the uniform minimum wage. As in the "company town", if the regulator knew the monopsonist's productivity, she would set a minimum wage equal to it. Under uncertainty, the regulator faces a trade-off: setting the minimum wage too high closes low-productivity

[1] Though I term this the "modern case," the theoretical possibility dates back at least to Robinson (1933) and Stigler (1946); see Boal and Ransom (1997) for a review. This possibility received renewed attention following the early empirical results of Katz and Krueger (1992) and Card and Krueger (1994) (for a discussion, see Manning 2003, 2021; Kline 2025).

[2] As well as maximizing joint surplus, this leaves the firm with zero profit, and hence maximizes worker surplus. Of course, if labor productivity varied with employment, the regulator would also need information about labor supply.

firms; setting it too low leaves monopsony distortions in place. I show that, under natural assumptions on the productivity distribution, this uncertainty tends to push optimal minimum wages down. Intuitively, the cost of setting the minimum wage too high exceeds the cost of setting it too low; as a result, an increase in the dispersion of firm productivity creates a precautionary motive for a lower minimum wage.

Now suppose the regulator can condition minimum wages on an observable (*e.g.*, industry or state), as in the second regime. If the observable is informative about productivity, it is straightforward that the regulator should condition on it. The more interesting question is how conditioning affects the *distribution* of surplus. I show that, when minimum wages are set optimally, the average *conditional* minimum wage exceeds the *unconditional* minimum wage. Intuitively, since uncertainty about productivity drags minimum wages down, conditioning on an informative signal increases the minimum wage on average. As a result, differentiated minimum wages shift expected surplus from firms to workers, relative to a uniform minimum wage: unions should prefer local minimum wage-setting, whereas firms should prefer national minimum wage-setting.

The third regime, in which minimum wages vary with employment, is most closely related to Baron and Myerson (1982). I show that the optimal minimum wage rises with employment: smaller firms face lower minimum wages. Intuitively, the regulator uses employment as a screening tool, exploiting the fact that more productive firms have a stronger incentive to hire more workers. This is consistent with practice in some jurisdictions, where minimum wages vary by firm size or small firms are exempted (ILO 2015; 2016b). On this view, firm-size variation in minimum wages reflects screening concerns, rather than political concessions to small-business lobbying.

Finally, I consider employment subsidies. In a competitive labor market, Lee and Saez (2012) argue that minimum wages should be accompanied by employment subsidies. This logic survives only partially under monopsony. Because monopsonistic firms earn rents, a subsidy risks transferring surplus from workers to firms—a concern absent under perfect competition. A regulator with strong redistributive preferences will prefer to *tax* rather than subsidize employment, so the sign of the optimal subsidy is no longer unambiguous. Nonetheless, the two instruments remain complements in the weaker sense that, under the optimal minimum wage, the regulator is more inclined to subsidize employment than under no minimum wage.

Related literature. Methodologically, the most relevant literature studies regulating monopolists with unknown cost (*e.g.*, Baron and Myerson

1982; Laffont and Tirole 1986). Substantively, the most relevant literature studies minimum wages, with seminal contributions including Stigler (1946); Guesnerie and Roberts (1987); Lee and Saez (2012); Cahuc and Laroque (2014).[3] I am not aware of any existing work viewing minimum wage-setting as a problem of regulation under private information about firm productivity. As a result, this prior literature has not studied my four main results— that uncertainty about productivity tends to drag down minimum wages; that local minimum wages tend to exceed national minimum wages; that smaller firms should face lower minimum wages; and that, under monopsony, a minimum wage may optimally be accompanied by an employment tax rather than a subsidy, if the regulator places sufficiently little weight on firm profit.

I take a mechanism design approach to minimum wages. A recent literature takes this approach to price controls (Dworczak, Kominers, and Akbarpour 2021; Kang and Watt 2024; Doligalski et al. 2025), but focuses on competitive markets and planners who are fully informed about production technologies. Loertscher and Muir (2023) analyze labor procurement under a given minimum wage; I study the design of the minimum wage itself. Loertscher and Muir (2024) note that a regulator can restore first-best employment by equating the minimum wage and the competitive wage. This requires the regulator to observe productivity; I examine the consequences of private information about productivity (or, equivalently, of setting a single minimum wage for heterogeneous firms).

The analysis applies the *spirit*, not the *letter*, of Baron and Myerson (1982). In both settings, a regulator designs policy for a privately informed firm. Nonetheless, the four main results discussed above are not implied by Baron and Myerson (1982) or subsequent work. Even the analysis of size-specific minimum wages, which comes closest to that framework, departs from it in two ways: (i) I consider a regulator who can set minimum wages, but cannot freely tax or subsidize employment; and (ii) I allow for the possibility that employment is rationed. The motivation for (i) is that the institutions that set minimum wages typically have authority over the wage floor but not over taxes or subsidies (International Labour Organization 2016a; Dickens 2023).[4]

[3] For reviews of the empirical literature, see Brown (1999); Neumark and Wascher (2008); Dube and Lindner (2024); for a review of the normative literature, see Rocheteau and Tasci (2008).

[4] Examples of statutory bodies with authority over minimum wages but not other fiscal instruments include Australia's Fair Work Commission, Germany's Minimum Wage Commission ("Mindestlohnkommission"), Mexico's CONASAMI (Comisión Nacional de los Salarios Mínimos), and South Korea's Minimum Wage Commission. In the UK, an independent statutory body, the Low Pay Commission, makes recommendations on the minimum wage which must then be approved by the central government. Examples in

In the context of regulating a monopolist, Amador and Bagwell (2022) and Kolotilin and Zapechelnyuk (2025) consider a setting without transfers (as in departure (i)). In their settings, a regulator without access to taxes or subsidies imposes a quantity regulation on a monopolist with private information about its costs. However, both of these papers rule out rationing by assumption (and for that reason do not consider departure (ii)).[5] The possibility of rationing is especially salient in the context of regulating a monopsonist, since rationed employment is a commonly-discussed consequence of minimum wages (Lee and Saez 2012); indeed, I show that the regulator's optimal mechanism does involve rationing. In the last part of the paper, I allow the regulator to choose both minimum wages and taxes/subsidies.

Outline. Section 2 outlines the model. Section 3 discusses uniform minimum wages. Section 4 turns to local minimum wages. Section 5 allows minimum wages to depend on firm size. Section 6 adds employment subsidies. Section 7 concludes. Proofs are in the appendix. The online appendices provide additional results, further empirical details, and some intermediate steps for the main proofs.

2. Setting

I begin by describing the basic setting absent regulation. I then turn to the regulator's preferences, and then the instruments available to her.

2.1. Setting absent regulation. A constant-returns monopsonist has labor productivity θ drawn from the CDF F, with full support on some interval $\Theta \subseteq \mathbb{R}$ and corresponding continuous PDF f. Revenue from employing L workers is thus θL.

The monopsonist faces a continuum of prospective workers of mass one. Workers differ only in outside options, drawn from a distribution G with full support on $[\underline{o}, \bar{o}] \subseteq \mathbb{R}_+$ and corresponding continuous PDF g, with $g(\underline{o}) > 0$. The CDF G can equivalently be interpreted as the distribution of the value of home production, the disutility of work, differences in preferences for non-wage aspects of work at the monopsonist (Card et al. 2018), or the wage available in informal markets, unaffected by the minimum wage. I assume that employment is at least sometimes jointly efficient ($F(\underline{o}) < 1$),

the U.S. include California's Fast Food Council, Minnesota's Nursing Home Workforce Standards Board, and New York's wage boards. In some European countries, sectoral wage floors are set by collective bargaining agreements between trade unions and employers' associations rather than a statutory wage floor. The parties to these agreements have the authority to set minimum wages, but not to impose taxes/subsidies (Boeri 2012; Garnero, Kampelmann, and Rycx 2015; International Labour Organization 2016a).

[5] Formally, these papers assume that, once the monopolist has chosen its quantity (subject to the regulation), the price is pinned down as the market-clearing price.

and that there are some firms whose productivity sits at the very bottom of the distribution of workers' outside options ($f(\underline{o}) > 0$).

The monopsonist chooses a wage w, and workers accept employment whenever the wage exceeds their outside option, so labor supply at wage w is $G(w)$, with elasticity $\epsilon(w) := g(w)w/G(w)$. Given wage w and employment L, the monopsonist's profit is $\pi_\theta(w, L) = L(\theta - w)$. Absent regulation, the monopsonist's problem is:

$$\max_{w,L} \pi_\theta(w, L), \text{ subject to labor supply: } L \leq G(w). \tag{1}$$

I explicitly include the constraint $L \leq G(w)$ to emphasize the possibility of excess labor supply. Absent regulation, the constraint binds and the problem is $\max_w \pi_\theta(w, G(w))$, yielding the familiar markdown rule for the monopsony wage $w^m(\theta)$:

$$w^m(\theta) = \theta - \frac{G(w^m(\theta))}{g(w^m(\theta))} = \theta \frac{\varepsilon(w^m(\theta))}{1 + \varepsilon(w^m(\theta))}. \tag{2}$$

I make the standard assumption that G is log-concave (Kline 2025), which ensures the first-order condition in (2) is sufficient for a solution to (1), at least among firms which do in fact employ workers.

2.2. Role of regulator. The monopsonist is subject to minimum wages set by a regulator who knows the productivity distribution F but not the realization θ. The regulator's payoff depends on firm profits, $\pi_\theta(w, L)$, and worker surplus. The surplus generated by employing a worker with outside option o at wage w is $w - o$. Given employment L and wage w, total worker surplus is

$$W(w, L) = wL - \int_{\tilde{o}=\underline{o}}^{G^{-1}(L)} \tilde{o} \, \mathrm{d}G(\tilde{o}). \tag{3}$$

Equation (3) imposes efficient rationing: if $L < G(w)$, workers with worse outside options are employed. This follows an extensive literature on minimum wages and price controls.[6] In Online Appendix B, I show that the main conclusions hold under alternative rationing assumptions.

The regulator's payoff is social welfare, defined as

$$S_\theta(w, L) = W(w, L) + \lambda \pi_\theta(w, L), \tag{4}$$

[6] For minimum wages, see Allen (1987); Lee and Saez (2012); generally, see Glaeser and Luttmer (2003); Davis and Kilian (2011). In an analysis of the 1990–1991 federal minimum wage increase, Luttmer (2007) finds evidence consistent with efficient rationing. Gerritsen and Jacobs (2020) discuss the effects of rationing on the optimal minimum wage.

where $\lambda \in [0, 1]$ parameterizes the regulator's weight on firm profits relative to worker surplus. When $\lambda = 1$, the regulator maximizes joint surplus; when $\lambda = 0$, she maximizes worker surplus.

2.3. Instruments available to the regulator. I consider three possible regimes for the regulator: a uniform minimum wage; a minimum wage that conditions on an exogenous firm characteristic, such as industry or geographic location; and a minimum wage that depends on endogenously-chosen firm size. Table 1 gives examples of each case.

TABLE 1. Illustrative examples of minimum wage-setting regimes

Regime (1) Uniform national minimum wage	Regime (2) Differentiated minimum wage by observable characteristics	Regime (3) Differentiated minimum wage by employment
Belgium	Australia	Parts of the U.S.
France	Costa Rica	Dominican Republic
Germany	Indonesia	Honduras
Ireland	Japan	Indonesia
Netherlands	Mexico	Malaysia
Poland	Panama	Pakistan
Portugal	Philippines	Panama
Spain	Vietnam	Philippines

Note: *This table presents examples of jurisdictions in which minimum wages are either (1) set nationally at a uniform level; (2) differentiated by observable firm characteristics; or (3) differentiated by employment. Jurisdictions appear in both columns (2) and (3) if their minimum wages are differentiated by both observable firm characteristics other than employment, and by employment itself. See Online Appendix C for details.*

In the first regime, the regulator chooses a uniform minimum wage. The first column of Table 1 gives examples, mostly from Europe: 22 out of 27 EU members have uniform minimum wages (Eurostat 2026). The regulator's problem is:

$$\text{(5)} \qquad \max_{\underline{w}} \int_{\theta \in \Theta} S_\theta\left(\hat{w}(\theta, \underline{w}), \hat{L}(\theta, \underline{w})\right) \mathrm{d}F(\theta), \text{ where}$$
$$\left(\hat{w}(\theta, \underline{w}), \hat{L}(\theta, \underline{w})\right) = \underset{w \geq \underline{w},\ L \leq G(w)}{\arg\max} \left\{\pi_\theta(w, L)\right\}, \text{ breaking ties in the regulator's favor.}$$

The assumption that ties are broken in the regulator's favor is a technical assumption which ensures that a solution to the regulator's problem exists. Call (5) the *problem of setting a uniform minimum wage*. I assume there is a unique solution to this problem, as will generically be the case.[7]

In the second regime, the regulator sets minimum wages conditional on an observable $c \in \mathcal{C}$ with CDF H. Conditional on c, productivity has CDF $F(\theta \mid c)$ and PDF $f(\theta \mid c)$. The second column of Table 1 gives examples: c either represents industry or geographic location.[8] I assume c cannot be changed in response to the minimum wage. In principle, firms could move state or industry in response to differentiated minimum wages; in practice, these effects are usually small (Rohlin 2011; Allegretto and Reich 2018; Rao and Risch 2026). Call this the *problem of setting local minimum wages*, and continue to assume there is a unique solution for each c. I will refer to the optimal $w(c)$ as the *optimal minimum wage in region c*.

In the third regime, the minimum wage depends on the monopsonist's chosen employment L. The third column of Table 1 gives examples. The regulator's problem is:

$$\max_{\underline{w}:\, \mathbb{R}_+\to\mathbb{R}_+} \int_{\theta\in\Theta} S_\theta\big(\hat{w}(\theta,\underline{w}), \hat{L}(\theta,\underline{w})\big)\, \mathrm{d}F(\theta), \text{ where}$$

$$\big(\hat{w}(\theta,\underline{w}), \hat{L}(\theta,\underline{w})\big) = \operatorname*{arg\,max}_{w\geq \underline{w}(L),\, L\leq G(w)} \Big\{\pi_\theta(w,L)\Big\}, \text{ breaking ties in the regulator's favor.}$$

As usual, we can apply the revelation principle to restate the regulator's problem in terms of a direct mechanism. Appendix Lemma A1 shows formally that the regulator's problem can be rewritten as:

$$\text{(6)} \qquad \max_{\hat{w}:\,\Theta\to\mathbb{R},\, \hat{L}:\,\Theta\to\mathbb{R}} \int_{\theta\in\Theta} S_\theta\big(\hat{w}(\theta), \hat{L}(\theta)\big)\, \mathrm{d}F(\theta),$$

$$\text{subject to (IC): } \theta \in \operatorname*{arg\,max}_{\hat{\theta}} \pi_\theta(\hat{w}(\hat{\theta}), \hat{L}(\hat{\theta})).$$

$$\text{(PC): } \pi_\theta(\hat{w}(\theta), \hat{L}(\theta)) \geq 0.$$

$$\text{(labor supply): } \hat{L}(\theta) \leq G(\hat{w}(\theta)), \text{ for all } \theta \in \Theta.$$

Call (6) the *problem of setting minimum wages by firm size*.

2.4. Discussion of assumptions. Throughout, I assume the productivity PDF f is log-concave, ensuring regularity in the sense of Myerson (1981).[9]

[7] Without the assumption of uniqueness, certain results comparing optimal minimum wages across settings are true in the strong set order commonly used in the literature on monotone comparative statics (Milgrom and Shannon 1994).

[8] The U.S. does not fall neatly in either regime: state and city minimum wages operate on top of a federal floor. Some European countries with a uniform statutory minimum wage in practice have sectoral bargaining that may effectively set industry-specific minimum wages; see Bijnens et al. (2024) for a discussion in Belgium.

[9] For background on log-concave distributions, see Bagnoli and Bergstrom (2005). In the discussion of setting local minimum wages, I assume that $f(\cdot \mid c)$ is log-concave for each c, and that the unconditional distribution f is log-concave.

Log-concavity is satisfied by the distributions most commonly used to model firm productivity—normal, exponential, type-1 extreme value (Gumbel), Gamma, or Weibull (with shape parameter at least 1)—which are themselves motivated by the empirical productivity distribution (Lentz and Mortensen 2008; Combes et al. 2012; Nigai 2017).[10] Under log-concavity, the regulator's problem has a solution, which does indeed involve a minimum wage.

Lemma 1. Consider the problems of setting a uniform wage, setting local minimum wages, or setting minimum wages by firm size. There is a solution to each problem. Moreover, the regulator strictly prefers the solution to setting no minimum wage.

In addition, log-concavity implies f is single-peaked at its mode q_f; I assume the optimal uniform minimum wage binds only firms with productivity below q_f. This assumption is empirically innocuous: for example, the modal labor productivity in U.S. manufacturing plants in Bernard et al. (2003) is around \$42 per hour in 2026 dollars, well above even the most ambitious minimum wage proposals.[11]

I conclude by discussing several assumptions. First, I assume the minimum wage does not affect workers' outside options. This excludes, for instance, spillovers to the rest of the wage distribution; there is some evidence of spillovers, though they are usually quantitatively small (Autor, Manning, and Smith 2016; Gopalan et al. 2021; Fortin, Lemieux, and Lloyd 2021). Second, I assume CRS production, following much of the existing literature on optimal minimum wages (Lee and Saez 2012; Cahuc and Laroque 2014). This is partly for tractability: the Baron-Myerson approach typically requires strong functional-form restrictions. Under CRS production, an increase in the uniform minimum wage reduces employment primarily through firm exit, while employment at continuing firms is unaffected or increases. Several empirical studies find evidence of this pattern (Aaronson et al. 2018; Luca and Luca 2019; Bossavie, Erdoğan, and Makovec 2019; Jardim and van Inwegen 2023; Rao and Risch 2026); indeed, Berger, Herkenhoff, and Mongey (2022, 2025) estimate almost-constant returns in quantitative models of U.S. labor market power. For this reason,

[10] Of course, caution is needed in comparing these estimates to F: most empirical studies estimate TFP rather than labor productivity; the relevant object for the model is productivity among firms plausibly affected by minimum wages; and productivity is measured with error (White, Reiter, and Petrin 2018; Bils, Klenow, and Ruane 2021). An exception is that there is some evidence that, at the very right tail, productivity follows a Pareto distribution (Nigai 2017); in any case, these highly-productive firms are less likely to be bound by the minimum wage.

[11] I detail this calculation in Online Appendix C.

I consider the CRS case a natural starting point, though I later return to some alternative assumptions.

Third, I assumed the monopsonist has private information about productivity θ, but not about labor supply G. Presumably monopsonists have much more private information about their own production function than about labor supply. The approach in Lewis and Sappington (1988a,b) could be used to extend the analysis to private information about labor supply.

Fourth, I treat labor as the firm's only input. One interpretation is that the firm also uses other inputs, but its production technology is constant returns to scale, and the firm is a price-taker in the market for those other inputs; I show in Online Appendix B that the problem described here can be viewed as a reduced-form representation of that general problem.

Finally, the regulator only has access to a single instrument: the minimum wage. As I discussed in the introduction, this is a leading case of interest in practice. Section 6 relaxes this assumption by introducing employment subsidies and taxes.

3. Uniform minimum wages

I begin with uniform minimum wages. I first show a general condition for an optimal uniform minimum wage, and then turn to the effects of uncertainty about productivity on the optimal uniform minimum wage.

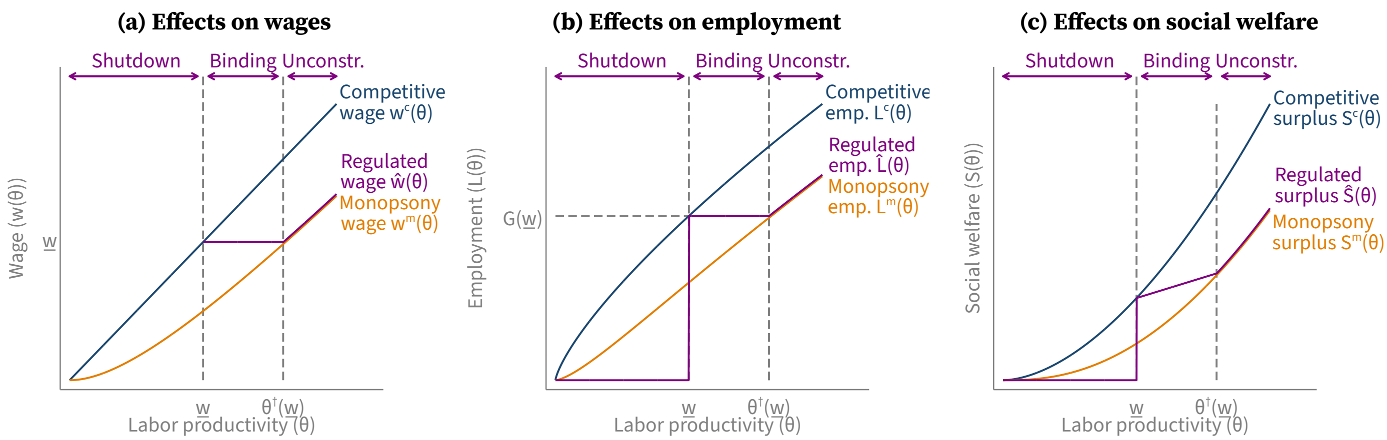
3.1. Effects of uniform minimum wages. Figure 1 depicts the effects of a minimum wage $\underline{w}$ on wages, employment, and surplus. As benchmarks, the navy curves show competitive outcomes: wage $w^c(\theta) = \theta$, employment $L^c(\theta) = G(\theta)$, and welfare $S^c(\theta) = S_\theta(\theta, L^c(\theta))$. The orange curves show unregulated monopsony outcomes: wage $w^m(\theta)$, employment $L^m(\theta) = G(w^m(\theta))$, and welfare $S^m(\theta) = S_\theta(w^m(\theta), L^m(\theta))$. For each θ, monopsony wages, employment, and social welfare fall below the competitive level.

The purple curves show wages, employment, and welfare under uniform minimum wage $\underline{w}$, denoted by $(\hat{w}(\theta), \hat{L}(\theta), \hat{S}(\theta))$. Define $\theta^\dagger(\underline{w})$ as the productivity level at which the minimum wage just binds, implicitly defined by $w^m(\theta^\dagger(\underline{w})) = \underline{w}$.[12] A minimum wage $\underline{w}$ partitions productivity into three intervals: (i) a *shutdown interval* $\{\theta \in \Theta\colon \theta < \underline{w}\}$, in which firms do not operate, since the minimum wage exceeds output from employing workers; (ii) a *binding interval* $\{\theta \in \Theta\colon \theta \geq \underline{w}, \theta < \theta^\dagger(\underline{w})\}$, in which

[12] We can always find a unique such productivity level for any minimum wage $\underline{w}$ which lies within the range of workers' outside options $(\underline{o}, \bar{o})$. To see this, note that, under our assumption that G is log-concave, the mapping $\theta \to w^m(\theta)$ is strictly increasing on the set of types for which the monopsony optimum is within $(\underline{o}, \bar{o})$. Hence, for any $\underline{w} \in (\underline{o}, \bar{o})$, there is a unique θ such that $w^m(\theta) = \underline{w}$. Of course, this productivity level may not actually arise: there is no guarantee that $\theta^\dagger(\underline{w}) \in \Theta$.

firms are constrained by the minimum wage, with $\hat{w}(\theta) = \underline{w}$; and (iii) an *unconstrained interval* $\{\theta \in \Theta\colon \theta \geq \theta^{\dagger}(\underline{w})\}$, in which firms continue to set monopsony wages.[13] Figure 1 depicts wages, employment, and welfare in each of these intervals.

FIGURE 1. The effects of a uniform minimum wage on wages, employment, and welfare



Note: *Panels (a), (b) and (c) depict wages, employment, and social welfare, respectively, under competition (navy curves), monopsony (orange curves), and a fixed uniform minimum wage* $\underline{w}$ *(purple curves). The annotations above each graph indicate the shutdown, binding, and unconstrained regions. I do not depict any wage in the shutdown region: these firms have no workers, and so are indifferent to any wage.*

As Figure 1(c) shows, the benefit of a higher minimum wage is higher welfare in the binding interval; the cost is a larger shutdown interval. The optimal minimum wage balances these costs and benefits at the margin.[14]

Lemma 2. The optimal uniform minimum wage $\underline{w}$ satisfies:

$$(7)\quad \underbrace{[1-\lambda]G(\underline{w})\left[F\left(\theta^{\dagger}(\underline{w})\right) - F(\underline{w})\right]}_{\text{Marginal redistributive gain}} + \underbrace{\lambda g(\underline{w}) \int_{\tilde{\theta}=\underline{w}}^{\theta^{\dagger}(\underline{w})} \left[\tilde{\theta} - \underline{w}\right] \, \mathrm{d}F(\tilde{\theta})}_{\text{Marginal efficiency gain}} = \underbrace{f(\underline{w}) W\left(\underline{w}, G(\underline{w})\right)}_{\text{MC of shutdowns}}.$$

The LHS of (7) is the marginal benefit of a higher minimum wage, arising from firms in the binding interval. The first term is the redistributive gain from transferring surplus from firms to workers. The expected transfer is the probability a firm lies in the binding interval, $F\left(\theta^{\dagger}(\underline{w})\right) - F(\underline{w})$, times employment at such a firm, $G(\underline{w})$. The transfer reduces firm surplus, which the regulator values at λ, but increases worker surplus, which she

[13] Of course, depending on the value of $\underline{w}$, some of these intervals could be empty.

[14] Since there is no guarantee that the regulator's objective function is concave, Lemma 2 provides a necessary condition for an optimal uniform minimum wage, but I do not claim that this is a sufficient condition.

values one-for-one; the resulting redistributive gain is weighted by $(1 - \lambda)$. The second term is the efficiency gain from higher employment. By the envelope theorem, newly-employed workers are indifferent, so the first-order efficiency gain accrues to the firm: a marginal increase in $\underline{w}$ raises employment by $g(\underline{w})$, increasing revenue by $g(\underline{w})(\theta - \underline{w})$. The RHS of (7) is the marginal cost: firms at the shutdown threshold cease operating. By the envelope theorem, this loss is borne by workers, since the marginal firm earns no profit.

3.2. Effects of uncertainty about productivity on uniform minimum wages. This basic analysis sets the stage for the question of interest: how does uncertainty about productivity affect the minimum wage? If productivity were known to be $\hat{\theta}$, the optimal minimum wage would be $\hat{\theta}$, restoring the competitive outcome. Suppose the regulator faces a productivity distribution centered at $\hat{\theta}$. Is the optimal minimum wage still $\hat{\theta}$, as would be the case under a quadratic-loss objective? Or does uncertainty systematically push the minimum wage in one direction?

The intuition for why uncertainty should tend to push minimum wages down is that the welfare cost of errors is asymmetric. Setting the minimum wage "too low" relative to productivity leaves wages inefficiently low, but still preserves some surplus from employment at the firm; setting it "too high" risks shutting down the firm entirely. Since the losses from overshooting are typically much larger than the losses from undershooting, uncertainty makes a cautious policy attractive.

Making this intuition precise requires some further structure on how dispersion increases: dispersion concentrated entirely above $\theta^{\dagger}(\underline{w})$ or below $\underline{w}$ leaves equation (7) unchanged.[15] I therefore focus on global changes in dispersion. Recall that a *location-scale family* with shape H_0, indexed by location m and scale s, is a collection of distributions with CDFs $H_{m,s}(x) = H_0\left(\frac{x-m}{s}\right)$. For example, the family of normal distributions is a location-scale family, indexed by their mean and variance. Given a density f with unique mode q_f, I will say that f' is a *shape-preserving spread* of f if f and f' belong to the same location-scale family, have the same mode, and f' has strictly larger variance.[16]

I consider the notion of a shape-preserving spread a reasonable way of comparing the dispersion between two distributions, while holding fixed their basic shape. Figure 2(a) provides a graphical example: the distribution in orange is a shape-preserving spread of the distribution in navy. For

[15] Indeed, even in the two-type case, it is insufficient to assume f' is a mean-preserving spread of f, as Online Appendix B shows.

[16] Equivalently, $f'(\theta) = \frac{1}{s} f\left(q_f + \frac{\theta - q_f}{s}\right)$, for some $s > 1$. The equivalence between these two statements is presented in Lemma A2.

instance, if f is a normal PDF with mean μ and variance σ^2, and f' is a normal PDF with mean μ and variance ${\sigma'}^2$, then f' is a shape-preserving spread of f iff ${\sigma'}^2 > \sigma^2$. Similarly, if f is the PDF of a logistic distribution with mode μ and scale s, and f' is also a logistic PDF with mode μ and scale s', then f' is a shape-preserving spread of f iff $s' > s$, and likewise for a type-1 extreme value (Gumbel) or Laplace distribution. The notion of a shape-preserving spread is connected to other intuitive measures of dispersion in economics and statistics; working within a single location-scale family allows me to change the dispersion of productivity, holding fixed its basic shape (Meyer 1987).[17]

Proposition 1. *If f' is a shape-preserving spread of f, the optimal minimum wage is strictly lower under f' than under f.*

The proof of Proposition 1 clarifies that, if we had not assumed that there is a unique optimal minimum wage, then we could have demonstrated that the *set* of optimal minimum wages under f dominates the *set* of optimal minimum wages under f', in the sense of the strong set order.[18]

Figure 2(a) gives some graphical intuition for Proposition 1. Suppose $\underline{w}$ was the optimal minimum wage under distribution f, drawn in navy. By Lemma 2, $\underline{w}$ satisfies $\int_{\tilde{\theta}=\underline{w}}^{\theta^\dagger(\underline{w})} \mathrm{MB}(\underline{w}, \tilde{\theta}) f(\tilde{\theta}) d\tilde{\theta} = f(\underline{w})\mathrm{MC}(\underline{w})$, where $\mathrm{MB}(\underline{w}, \tilde{\theta})$ is the marginal benefit to social welfare of an increase in the minimum wage $\underline{w}$ given that the realized firm type is θ, and $\mathrm{MC}(\underline{w})$ is the corresponding marginal cost.[19] The marginal cost term depends on the density at $\underline{w}$, as indicated by the dashed red line; the marginal benefit terms depend on the densities between $\underline{w}$ and $\theta^\dagger(\underline{w})$, with an example such density indicated by the dashed green line.

[17]For examples of economics papers analyzing increased dispersion as an increase in scale within a location-scale family, see Sandmo (1971); Meyer (1987); Sinn (1983); Maskin and Riley (2000); Johnson and Myatt (2006); Wong and Ma (2008). For background on location-scale families, see Lehmann and Casella (1998, section 1.4) or Casella and Berger (2024, section 3.5). The notion of a shape-preserving spread is also related to the canonical definition of "peakedness" (Birnbaum 1948); indeed, f' is a shape-preserving spread of f if they belong to the same location-scale family, and f is strictly more peaked around their common mode, in the sense of Birnbaum (1948).

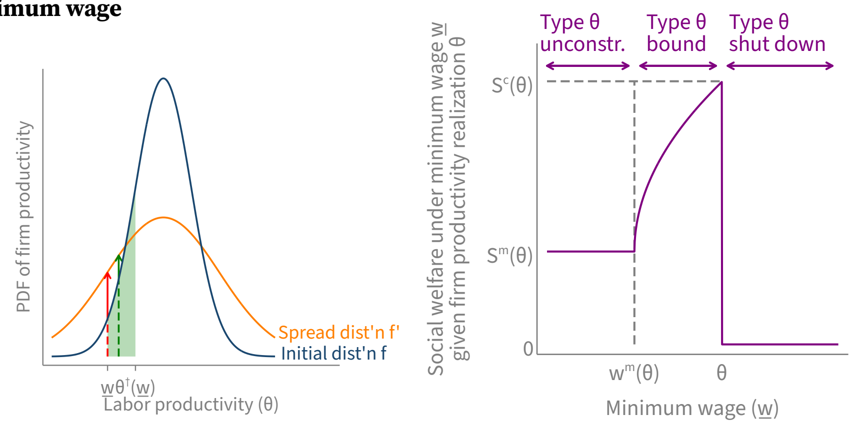
[18]Recall that, for two subsets of $\mathbb{R}$, S and S', the set S dominates S' if, for any $x \in S$ and $x' \in S'$, we have $\max\{x, x'\} \in S$ and $\min\{x, x'\} \in S'$.

[19]That is, $\mathrm{MB}(\underline{w}, \tilde{\theta}) := \mathbb{1}\{\theta \in [\underline{w}, \theta^\dagger(\underline{w}))\}\{[1 - \lambda]G(\underline{w}) + \lambda g(\underline{w})[\tilde{\theta} - \underline{w}]\}$ and $\mathrm{MC}(\underline{w}) := W(\underline{w}, G(\underline{w}))$.

FIGURE 2. Graphical intuition for Proposition 1

(a) Effects of a shape-preserving spread on marginal cost and benefit of a higher minimum wage

(b) Effects of varying the minimum wage on social welfare, for a fixed firm type



Note: *Panel (a) illustrates the effects of moving from an initial distribution (in navy) to a shape-preserving spread (in orange) on the marginal costs (in red) and benefits (in green) of the minimum wage. Panel (b) depicts social welfare for a type-θ monopsonist, as a function of the uniform minimum wage* $\underline{w}$.

Now consider replacing f with the shape-preserving spread f'. At the old minimum wage $\underline{w}$, the marginal cost is multiplied by the density ratio $f'(\underline{w})/f(\underline{w})$, whereas the marginal benefit to each type $\tilde{\theta} \in [\underline{w}, \theta^{\dagger}(\underline{w})]$ is multiplied by $f'(\tilde{\theta})/f(\tilde{\theta})$. This density ratio is decreasing on $[\underline{w}, \theta^{\dagger}(\underline{w})]$: graphically, the red arrow is proportionally much longer than the red dashed line, but the green arrow is only marginally longer than the green dashed line. As a result, whereas the marginal benefit and cost of $\underline{w}$ given distribution f were equal, the marginal cost is higher under distribution f', and hence the optimal minimum wage is lower.

As another piece of graphical intuition, fix a productivity type θ and consider the shape of social welfare $S(\theta)$ as the minimum wage $\underline{w}$ varies, as in Figure 2(b). When the minimum wage is below the monopsony wage $w^m(\theta)$, it does not bind: social welfare is at its monopsony level, $S^m(\theta)$. As the minimum wage increases above $w^m(\theta)$, welfare rises until reaching the competitive level $S^c(\theta)$ at $w = \theta$. When the minimum wage is raised marginally above θ, the firm shuts down and welfare falls to zero. That is, the optimal minimum wage is $\underline{w} = \theta$: a marginally lower minimum wage generates a second-order welfare loss, whereas a marginally higher minimum wage generates a first-order welfare loss.

This discussion highlights the role of the CRS assumption: under CRS, the marginal cost of the minimum wage at $\underline{w}$ falls entirely on the type-$\underline{w}$ firm, which exits. Do the conclusions of Proposition 1 hold only in the CRS case? In Online Appendix B, I consider decreasing returns to scale production technologies. I show that Proposition 1 continues to hold allowing for a range of other technologies, such as the possibility that revenue is θL^{α}, for some $\alpha \in (0, 1]$. The intuition is unchanged: the costs of the minimum wage are borne on the left tail of productivity, and a more dispersed distribution increases the mass in that tail.

3.3. Policy implications. Proposition 1 has several implications. First, productivity dispersion varies substantially between industries (Syverson 2004; Cunningham et al. 2023), regions (Combes et al. 2012), and countries (Hsieh and Klenow 2009; Bartelsman, Haltiwanger, and Scarpetta 2013; Kehrig and Vincent 2025; Restuccia 2025). Proposition 1 suggests these differences imply differences in optimal minimum wages. For instance, Hsieh and Klenow (2009) find much more productivity dispersion in China than the U.S.; Proposition 1 suggests this dispersion may contribute to substantially lower minimum wages in China (even adjusted for differences in median wages).[20] Similarly, rising within-country productivity dispersion in the OECD (Andrews, Criscuolo, and Gal 2016) and the U.S. (Decker et al. 2020) depresses optimal minimum wages.

Second, there is some evidence that productivity dispersion is countercyclical: during periods of recession, the variance of firm-level productivity increases (Kehrig et al. 2011; Bloom et al. 2018). As a result, Proposition 1 implies that, during recessions, minimum wages should fall not *just* because *average* productivity falls, but because dispersion *rises*, which makes regulating the labor market using minimum wages more challenging.

Third, if the regulator can differentiate by industry or location, Proposition 1 suggests that sectors with more dispersion should face lower minimum wages. Conversely, some policies aim to compress the productivity distribution by raising productivity at low-productivity firms—for instance, by providing low-performers with managerial training (Adhvaryu, Murathanoglu, and Nyshadham 2023) or consulting services (Bijnens, Jäger, and Schoefer 2025). Proposition 1 shows that these policies have an unintended secondary benefit: by reducing the mass of firms vulnerable to a given minimum wage, they relax the employment cost that otherwise constrains the minimum wage, and thus facilitate higher minimum wages.

[20] Dautović, Hau, and Huang (2024) find that the ratio of the minimum wage to the median wage—sometimes called the Kaitz ratio—was around 18% in China in 2009, but around 30% in the U.S.

4. Local minimum wages

4.1. The value of local minimum wages. The preceding discussion suggests there are benefits to differentiating minimum wages. Suppose the regulator observes a characteristic c—such as industry or geographic location—that may be informative about productivity. How does conditioning on c affect minimum wages?

Clearly, the regulator should generally condition on c: the optimal minimum wage given c satisfies Lemma 2's first-order condition, replacing the unconditional distribution $F(\cdot)$ with the conditional $F(\cdot \mid c)$. Thus, if c is informative about productivity, the optimal local minimum wage depends on c, following the intuition of Holmström's (1979) "informativeness theorem". Moreover, as is intuitive, regions or sectors with "higher" productivity are optimally assigned higher minimum wages, in the following sense. I previously defined a shape-preserving *spread*; I will say f' is a *shape-preserving shift* of f if they belong to the same location-scale family and have the same scale, but differ in location, and that f' is a *shape-preserving rightward shift* of f if f' is a shape-preserving shift of f and its location is strictly higher than f's location. For example, two normal distributions with the same variance are shape-preserving shifts of one another; the distribution with the higher mean is a shape-preserving rightward shift of the other distribution.

Lemma 3. Say $f(\cdot \mid c')$ is a shape-preserving rightward shift of $f(\cdot \mid c)$. Then the optimal minimum wage in region c' is higher than the optimal minimum wage in region c.

4.2. The effects of differentiation on the average level of minimum wages. Lemma 3 is intuitive: higher-productivity regions or industries are assigned higher minimum wages. The more interesting question is how conditioning on c affects the *average* level of minimum wages, and hence the distribution of surplus. I analyze this question under two additional assumptions. First, I assume the productivity distribution f is *symmetric* around its mode: $f(q_f - a) = f(q_f + a)$—as would be the case, for instance, if productivity is normally-distributed. Second, I previously defined a shape-preserving *spread*; I will say f' is a *shape-preserving shift* of f if they belong to the same location-scale family and have the same scale, but differ in location.

Proposition 2. *Say the conditional productivity distributions are symmetric and shape-preserving shifts of one another, and the unconditional distribution is in the same location-scale family. Then the optimal uniform minimum wage is weakly lower than the average optimal local minimum wage.*

Thus, local wage-setting shifts surplus from firms to workers, relative to a national minimum wage. The proof is in two steps. Fix a location-scale family and write the optimal minimum wage under mode m and variance σ^2 as $\underline{w}(m, \sigma^2)$.[21] The optimal local minimum wage given c is $\underline{w}(m_c, \sigma^2)$, where m_c is the mode given c and σ^2 is the variance, while the optimal uniform minimum wage is $\underline{w}(\bar{m}, \bar{\sigma}^2)$, for $(\bar{m}, \bar{\sigma}^2) := (\mathbb{E}[m_c], \sigma^2 + \mathrm{Var}(m_c))$ the national mode and variance.[22] The proposition asserts $\mathbb{E}[\underline{w}(m_c, s)] \geq \underline{w}(\bar{m}, \bar{s})$. The first step is to note that the mapping $m \mapsto \underline{w}(m, s)$ is convex, and hence $\mathbb{E}[\underline{w}(m_c, s)] \geq \underline{w}(\bar{m}, s)$. Second, since $\bar{s} \geq s$, the distribution defined by $(\bar{m}, \bar{s})$ is a shape-preserving spread of $(\bar{m}, s)$; Proposition 1 implies $\underline{w}(\bar{m}, s) \geq \underline{w}(\bar{m}, \bar{s})$. Combining the inequalities gives the result.

Intuitively, conditioning on an informative signal reduces uncertainty about productivity. Since Proposition 1 implies that greater uncertainty tends to lower optimal minimum wages, allowing the regulator to condition on c raises the minimum wage on average.

Proposition 2 speaks to debates over the level of wage-setting. In the U.S., there is an active debate over local *vs.* state *vs.* federal minimum wages (Simon and Wilson 2021; Dube and Lindner 2021, 2024). In the EU, the Adequate Minimum Wages Directive has prompted related debates over national *vs.* supra-national wage-setting (Schulten and Müller 2021; European Parliament and Council of the European Union 2022; Eurofound 2023; Anxo 2024). Proposition 2 shows that the level of wage-setting affects both the level of wages, and the distribution of surplus. Of course, different layers of government may also place different weights on firm *vs.* worker surplus, corresponding to different values of λ. The proposition shows that the level of wage-setting matters even abstracting from these political economy differences.

This result echoes practitioners' views of national *vs.* local minimum wages. Consistent with the proposition, business groups have tended to favor "preemption laws" in which state governments regulate minimum wages to prevent local governments from doing so (Treskon et al. 2021). Conversely, labor unions have tended to favor local and industry-specific wage-setting (UK Trades Union Congress 2013; NZCTU Te Kauae Kaimahi 2025). Proposition 2 suggests one way to interpret these stances: workers' representatives prefer local wage-setting, since tailoring minimum wages to local conditions tends to increase them, on average. On the other hand,

[21] Under our maintained assumption that the productivity distributions are unimodal, once we have fixed a location-scale family, each distribution is fully characterized by (m, σ^2).

[22] The unconditional mode coincides with the expected value of the conditional modes only because of our assumption that the distributions are symmetric. Appendix Lemma A4 derives these expressions.

firms tend to prefer centralized wage-setting because reducing the information available to the regulator tends to lower the optimal minimum wage.

5. Minimum wages by firm size

Now consider the problem of setting minimum wages by firm size. Should smaller firms face relaxed minimum wages, as in some countries? How does conditioning on firm size affect wages and employment? Despite the use of size-contingent minimum wages, I am not aware of any previous analysis of these questions.

This section provides such an analysis. The benefit of size-contingent minimum wages is screening: in the first-best case, the regulator would set a minimum wage exactly equal to the firm's productivity; when the regulator cannot observe the firm's productivity directly, she can nonetheless use firm size as a signal of productivity, since larger firms will tend to be more productive. The cost of size-contingent minimum wages is that they create an additional margin for distortion, by encouraging some firms to reduce employment in order to qualify for a lower minimum wage. Indeed, an existing empirical literature finds that size-contingent regulations around labor regulations encourage firms to reduce employment (Schivardi and Torrini 2008; Garicano, Lelarge, and Van Reenen 2016; Amirapu and Gechter 2020). I show that, despite this distortion, the regulator nonetheless prefers to employ size-contingent minimum wages: intuitively, the benefit of avoiding firm exit exceeds the losses from incentivizing some firms to reduce employment in order to pay lower minimum wages.[23]

For this section only, I assume that the type space is a compact interval $\Theta = [\underline{\theta}, \bar{\theta}]$, which avoids some technical complications relating to the existence and implementation of the optimal mechanism. The assumption that the type space is bounded below is without loss, since clearly firms with negative productivity would choose not to operate. The assumption that the type space is bounded above is with some loss, but the upper bound on the type space $\bar{\theta}$ can be taken to be arbitrarily large.

5.1. Incentive-compatible mechanism. I begin with the implications of incentive-compatibility. For types $\theta_H > \theta_L$ assigned positive employment, incentive-compatibility requires:

$$\hat{L}_H[\theta_H - \hat{w}_H] \geq \hat{L}_L[\theta_H - \hat{w}_L]. \tag{IC_H}$$

[23] On the other hand, I abstract from evasion responses to regulation by firm size. These forms of evasion include misreporting of firm size (Almunia and Lopez-Rodriguez 2018), and firm "splitting" (Onji 2009), both of which have been documented in response to size-specific taxes.

$$(\mathrm{IC}_L) \qquad \hat{L}_L[\theta_L - \hat{w}_L] \geq \hat{L}_H[\theta_L - \hat{w}_H].$$

Adding (IC_H) and (IC_L) gives $\hat{L}_H \geq \hat{L}_L$, so *employment* is increasing in firm type. For *wages*, suppose instead $\hat{w}_H < \hat{w}_L$. Since the L-type firm employs workers, its participation constraint requires $\theta_L - \hat{w}_L \geq 0$, and so $\theta_L - \hat{w}_H > \theta_L - \hat{w}_L \geq 0$. Combining this with $\hat{L}_H \geq \hat{L}_L$ yields $\hat{L}_H[\theta_L - \hat{w}_H] > \hat{L}_L[\theta_L - \hat{w}_L]$, contradicting ($\mathrm{IC}_L$), and hence wages are also increasing in type.

Lemma 4. Any mechanism that satisfies the incentive, participation, and labor supply constraints can be implemented by a minimum-wage schedule that is weakly increasing in employment.

5.2. The optimal mechanism. To determine when this monotonicity is strict, I now turn to the optimal mechanism.

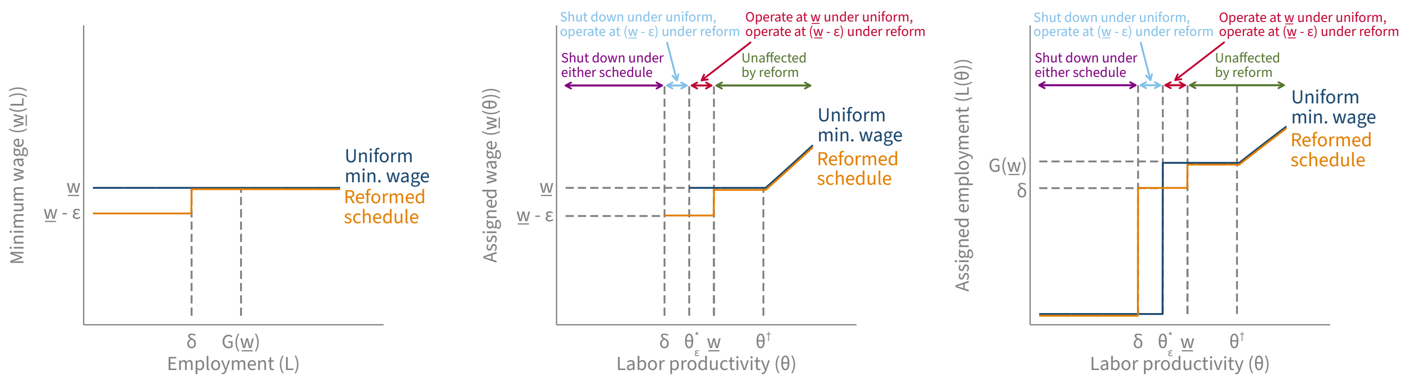
Proposition 3. *The optimal size-contingent minimum wage schedule is strictly increasing in employment.*

Thus, the regulator strictly benefits from allowing smaller firms to pay lower minimum wages. To see the intuition, start from a uniform minimum wage $\underline{w}$, shown in navy in Figure 3(a). Fix $\delta \in (0, G(\underline{w}))$, and consider a reform permitting firms employing at most δ workers to instead pay $(\underline{w} - \varepsilon)$ (shown in orange). For $\varepsilon > 0$ sufficiently small, this reform raises welfare, as I now show.

Let $\theta^\dagger$ denote the highest type bound by the uniform minimum wage. Under the original policy, firms with $\theta \in [\underline{\theta}, \underline{w})$ do not operate; firms with $\theta \in [\underline{w}, \theta^\dagger)$ choose wage-employment pair $(\underline{w}, G(\underline{w}))$; and firms with $\theta \geq \theta^\dagger$ pay monopsony wages (shown in navy in panels (b) and (c)). The reform changes behavior in two ways (shown in orange). First, firms with $\theta \in [\underline{w} - \varepsilon, \underline{w})$, previously shut down, now operate at wage-employment pair $(\underline{w} - \varepsilon, \delta)$. Second, some firms that previously chose $(\underline{w}, G(\underline{w}))$ instead downsize to $(\underline{w} - \varepsilon, \delta)$ to qualify for the lower minimum wage. It is simple to show that this is true for the interval $[\underline{w}, \theta^*_\varepsilon)$, for $\theta^*_\varepsilon := \underline{w} + \delta\varepsilon/[G(\underline{w}) - \delta]$.

FIGURE 3. The benefit of size-contingent minimum wages over uniform minimum wages

(a) Reform to minimum wage schedule **(b) Wage by firm productivity, before and after reform** **(c) Employment by firm productivity, before and after reform**



Note: *Panel (a) depicts the original uniform minimum wage (in navy) and the reformed schedule (in orange). Panel (b) depicts wages under the original and reformed schedules. Panel (c) depicts employment under the original and reformed schedules.*

The first effect raises welfare by bringing into operation firms that generate positive surplus; the second lowers it, as some incumbent firms cut wages and employment. The net welfare effect is

$$\text{(8)} \qquad \Delta_\varepsilon := \underbrace{\int_{\tilde{\theta}=\underline{w}-\varepsilon}^{\underline{w}} S_{\tilde{\theta}}(\underline{w}-\varepsilon, \delta)\, \mathrm{d}F(\tilde{\theta})}_{\uparrow \text{ in welfare from more firms operating}} - \underbrace{\int_{\tilde{\theta}=\underline{w}}^{\theta^*_\varepsilon} \left[S_{\tilde{\theta}}(\underline{w}, G(\underline{w})) - S_{\tilde{\theta}}(\underline{w}-\varepsilon, \delta)\right] \mathrm{d}F(\tilde{\theta})}_{\downarrow \text{ in welfare from some firms downsizing}}.$$

The first term in (8) reflects the fact that the reform allows some low-productivity firms that were shut down by the uniform minimum wage $\underline{w}$ to operate under minimum wage $\underline{w} - \varepsilon$, albeit at the reduced scale δ. The second term is the cost of the reform: some firms that previously operated at minimum wage $\underline{w}$ will reduce employment in order to qualify for the new, lower, minimum wage $\underline{w} - \varepsilon$.

As $\varepsilon \to 0$, taking a first-order expansion to (8) around $\varepsilon = 0$ gives

$$\text{(9)} \qquad \Delta_\varepsilon = \varepsilon f(\underline{w}) \frac{G(\underline{w})\delta}{G(\underline{w}) - \delta} \left[\frac{W(\underline{w}, \delta)}{\delta} - \frac{W(\underline{w}, G(\underline{w}))}{G(\underline{w})}\right] + o(\varepsilon),$$

where worker surplus is $W(w, L) = wL - c(L)$, for $c(L) = \int_{\tilde{o}=\underline{o}}^{G^{-1}(L)} \tilde{o}\, \mathrm{d}G(\tilde{o})$. Since G^{-1} is increasing, c is convex, and so $c(L)/L$ is increasing. Hence per-worker surplus $W(\underline{w}, L)/L$ is strictly decreasing in L: intuitively, since workers with worse outside options are employed first, per-worker surplus falls as we employ more workers. As a result, since $\delta < G(\underline{w})$, the bracketed term is strictly positive, and hence $\Delta_\epsilon > 0$. Thus, for small $\varepsilon > 0$, the reform raises welfare. Hence, the regulator prefers a schedule in which the minimum wage is *sometimes* strictly increasing; Proposition 3 further shows that the optimal schedule is *generally* strictly increasing.

Proposition 3 thus provides a way to rationalize policies which either set lower minimum wages for smaller firms, or entirely exempt them from minimum wages. These policies have traditionally been seen as political concessions to small businesses (Seltzer 1995; Levin-Waldman 2001). Proposition 3 instead suggests that these lower minimum wages can instead be viewed as part of an optimal screening device.[24]

It is worth distinguishing this rationale for size-contingent minimum wages from existing explanations of size-dependent regulation. For instance, Keen and Mintz (2004) presents a model to explain why small firms are, in some jurisdictions, exempted from value-added tax (VAT). In the model, the administrative and compliance costs of VAT are constant by firm size, but the benefits to the regulator from imposing VAT are increasing in firm size, since large firms have a greater quantity of taxable sales; as a result, the optimal policy exempts small firms from VAT, but imposes VAT on large firms. Similar arguments are found in Dharmapala, Slemrod, and Wilson (2011); Kaplow (2019). A similar justification could be provided for small firms being exempted from the minimum wage, to the extent that the costs of enforcement are large relative to the benefits of compliance (Ashenfelter and Smith 1979; Basu, Chau, and Kanbur 2015). The mechanism here is different: enforcement is costless and the firm's choices are perfectly observed, but productivity is not, and the regulator uses employment as a screen for productivity.

5.3. Other features of the optimal mechanism. A full characterization of the optimal mechanism is more involved and is deferred to Online Appendix D. In addition to our conclusion that minimum wages are smaller for smaller firms, two other qualitative features of the optimal mechanism are of interest. First, employment and wages are distorted *even for the highest-productivity type* $\bar{\theta}$, contrary to the usual intuition of no distortion at the top. The distortion arises because the minimum wage serves two roles simultaneously: screening firms by productivity and determining employment. Efficient employment for the highest-type firm is $G(\bar{\theta})$; to

[24]Alternatively, if the regulator is constrained to use a single minimum wage, but can choose where to use resources to enforce minimum wage regulations, the logic of Proposition 3 suggests that the regulator may prefer to expose small firms to a lower probability of enforcement. For a discussion of related issues, see Ashenfelter and Smith (1979) and Basu, Chau, and Kanbur (2015).

achieve that employment, the wage assigned to the highest-type firm must be $\hat{w}(\bar{\theta}) = \bar{\theta}$, which would leave the highest type no profit. As a result, if any lower type is allowed to operate—which would require that it operates at some wage $\hat{w} < \bar{\theta}$—the highest type would deviate to that other type. Hence, the only way to ensure that employment is efficient for the highest type would be to prevent other types from operating, which reduces the regulator's surplus. Instead, the optimal mechanism leaves the highest type some information rent, so its wage falls below $\bar{\theta}$ and employment is distorted.

Second, the optimal mechanism involves rationing, in the sense that some types are assigned a wage-employment pair $(\hat{L}, \hat{w})$ with the property that labor supply at wage $\hat{w}$ exceeds employment: $G(\hat{w}) > \hat{L}$. The intuition is that a uniform minimum wage encourages firms to pay higher wages only through the threat of shutdown; a size-contingent schedule can instead raise wages by rationing firms that pay less. The regulator prefers the latter because it replaces the discrete efficiency cost of exit with the smaller distortion of reduced employment among low-productivity firms. These and other features of the optimal mechanism are discussed in more detail in Online Appendix D.

6. Employment subsidies and taxes

This discussion highlights our assumption that the regulator's only instrument is the minimum wage. I now allow for employment subsidies, financed by lump-sum taxes on workers. I restrict attention to uniform policies: a common minimum wage $\underline{w}$ and a constant per-worker subsidy s, where $s < 0$ corresponds to an employment tax.[25] The regulator's problem is then:

$$\max_{\underline{w},\, s} \int_{\theta \in \Theta} \left\{ \underbrace{W\big(\hat{w}(\theta), \hat{L}(\theta)\big) - s\hat{L}(\theta)}_{\text{Worker surplus net of taxes}} + \lambda \underbrace{\left[\pi_\theta(\hat{w}(\theta), \hat{L}(\theta)) + s\hat{L}(\theta)\right]}_{\text{Firm surplus}} \right\} \mathrm{d}F(\theta), \text{ where}$$

$$\big(\hat{w}(\theta), \hat{L}(\theta)\big) = \underset{w \geq \underline{w},\ L \leq G(w)}{\arg\max} \left\{ \pi_\theta(w, L) + sL \right\}, \text{ breaking ties in the regulator's favor.}$$

It is simple to show that there is a solution to the regulator's problem, following the argument in Lemma 1. I assume throughout that the regulator's problem is concave with respect to the subsidy level. I will say that the regulator *subsidizes employment* if her chosen subsidy is weakly positive.

[25] I assume the subsidies are paid to employers, rather than workers. I show in Online Appendix B that it is without loss to consider subsidies paid only to one side of the market. In Online Appendix D, I consider size-contingent subsidies, though in practice these are less common.

6.1. Employment subsidies/taxes in the absence of a minimum wage. As a benchmark case, consider first the setting without a minimum wage. Even absent a minimum wage, the regulator may use a subsidy (or tax) to combat monopsony distortions. Since s effectively raises productivity from θ to $\theta + s$, the marginal pass-through of subsidy s into wages for type θ is $\frac{dw^m(\theta+s)}{ds}$. Define expected employment by $\bar{L}(s) := \mathbb{E}_{\theta \sim F}[L^m(\theta + s)]$, and worker-weighted marginal pass-through by $\eta(s) := \mathbb{E}_{\theta \sim f}[L^m(\theta + s)\frac{dw^m(\theta+s)}{ds}]/\bar{L}(s) \in (0, 1)$.

Lemma 5. Absent a minimum wage, the optimal subsidy s^* satisfies:

$$\eta(s^*)\bar{L}(s^*) + \lambda\bar{L}(s^*) = \bar{L}(s^*) + s^*\bar{L}'(s^*) \tag{10}$$

$$\iff s^* = [\eta(s^*) + \lambda - 1]\left[\frac{d[\ln \bar{L}(s^*)]}{ds^*}\right]^{-1}. \tag{11}$$

The LHS of (10) is the marginal benefit of the subsidy: the expected increase in the wage bill for inframarginal workers plus the increase in expected firm profit, equal to expected employment by the envelope theorem and valued by the regulator at weight λ. The RHS is the marginal cost: the mechanical cost $\bar{L}(s^*)$ plus the behavioral response.

Rearranging into (11) gives a simple formula for the optimal subsidy. Since $\frac{d[\ln \bar{L}(s)]}{ds} \geq 0$, the sign of the optimal subsidy depends on $[\eta(0) + \lambda - 1]$. Intuitively, at $s = 0$, a \$1 subsidy costs workers \$1 in taxes and (by the envelope theorem) increases firm profit by \$1, which the regulator values at λ; it is therefore worthwhile only if the pass-through to wages is at least $(1 - \lambda)$. In consequence, since pass-through $\eta(s) \in (0, 1)$, the regulator subsidizes employment if she equally values worker and firm surplus ($\lambda = 1$), and taxes employment if she values only worker surplus ($\lambda = 0$).

6.2. Employment subsidies/taxes in the presence of a minimum wage. This analysis absent a minimum wage sets up the main question: when the regulator can use *both* a minimum wage, *and* employment subsidies (or taxes), which should she use? Under perfect competition, Lee and Saez (2012) show that a minimum wage should be accompanied by employment subsidies. Proposition 4 shows that this need not hold under monopsony power. To simplify, I consider the case in which labor supply is *isoelastic*: that is, $G(w) = \kappa w^{\varepsilon}$, for some constant κ and supply elasticity ε.[26]

Proposition 4. *Say labor supply is isoelastic.*

[26] I consider isoelastic labor supply a leading case of interest, even though it does not meet all of the assumptions on G described in section 2: in particular, it is not the case that $g(\underline{o}) = g(0) = 0$, and labor supply is not bounded above.

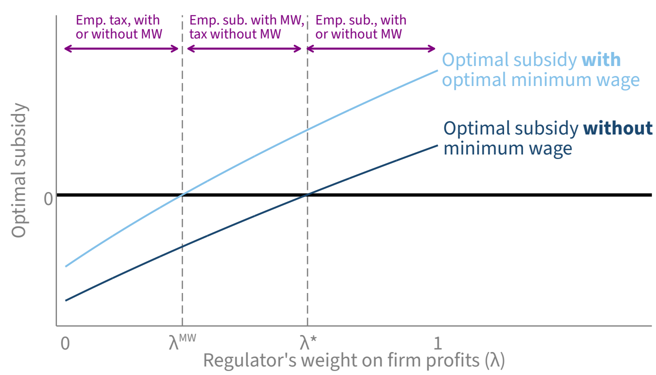
(i) In the absence of a minimum wage, the regulator subsidizes employment if and only if $\lambda \geq \lambda^*$, for some threshold weight on firm profits $\lambda^* \in (0, 1)$.

(ii) Under the optimal minimum wage $\underline{w}$, the regulator subsidizes employment if and only if $\lambda \geq \lambda^{MW}$, for some threshold weight on firm profits $\lambda^{MW} \in (0, 1)$.

(iii) The regulator is more likely to subsidize employment under the minimum wage, in the sense that the weight on firm profits required to subsidize employment is lower in the presence of a minimum wage: $\lambda^* > \lambda^{MW}$.

Figure 4 displays the conclusions of Proposition 4 graphically.[27] In the absence of a minimum wage, the regulator uses a subsidy ($s \geq 0$) whenever her weight on firm profits is sufficiently higher ($\lambda \geq \lambda^*$). The presence of an optimally-chosen minimum wage lowers the required weight on firm profits to $\lambda^{MW} < \lambda^*$. Thus, the space of values for the regulator's weight is divided into three regions: when the regulator's weight on firm profits is very low ($\lambda < \lambda^{MW}$), she will *tax* rather than subsidize employment; when her weight on firm profits is very high ($\lambda \geq \lambda^*$), she will subsidize employment; for intermediate weight on firm profits ($\lambda \in [\lambda^{MW}, \lambda^*)$), whether the regulator subsidizes employment depends on whether she has access to a minimum wage: the regulator will subsidize employment if she can use a minimum wage, but otherwise will tax employment.

FIGURE 4. Optimal employment subsidies and taxes under Proposition 4



Note: *This figure graphically illustrates Proposition 4. The optimal subsidy under an optimal minimum wage, as a function of the regulator's weight on firm profits λ, is displayed in light blue. The optimal subsidy under no minimum wage, as a function of the regulator's weight on firm profits λ, is displayed in navy.*

[27]To construct the figure, I explicitly compute the optimal subsidy, with and without a minimum wage, in the case of isoelastic labor supply and normally-distributed productivity. In addition to showing the conclusions from Proposition 4, the curves in Figure 4 indicates that, in this case, the optimal subsidy is increasing in the regulator's weight on firm surplus λ, and that the optimal subsidy with a minimum wage is strictly above the optimal subsidy without a minimum wage. I do not claim these are general properties of the optimal subsidy.

Intuitively, under competition, the zero-profit condition ensures that a subsidy does not transfer surplus from workers to firms: any subsidy is fully passed through to workers as higher wages (Lee and Saez 2012). Under monopsony, the pass-through of a subsidy is incomplete: for firms unconstrained by the minimum wage, wages rise by only $\varepsilon/[\varepsilon+1]$ per dollar of subsidy, the residual $1/[\varepsilon+1]$ accrues to firms.[28] If the regulator has no redistributive preferences ($\lambda = 1$), then she is indifferent to the distribution of surplus, and the minimum wage should always be accompanied by a subsidy, as in Lee and Saez (2012). On the other hand, if the regulator gives sufficiently little weight to firm surplus (λ sufficiently low), she will prefer to tax rather than subsidize employment. Proposition 4 thus shows that, under monopsony power, minimum wages are *not* generally accompanied by employment subsidies, unlike in Lee and Saez (2012).[29]

On the other hand, the two instruments are complements in the more limited sense that permitting the regulator to use a minimum wage only makes her more likely to also use an employment subsidy, as part (iii) states. The intuition is that, when there is a minimum wage $\underline{w}$, the subsidy has the additional benefit of allowing firms with productivity marginally below $\underline{w}$ to operate. Because these firms earn approximately zero profits, the surplus their operation generates accrues to workers, not owners. As a result, the minimum wage makes employment subsidies more attractive, though the regulator may still tax rather than subsidize employment under her optimal policy.

7. Discussion

In this paper, I have taken the view that a central challenge in setting minimum wages is the absence of information. If the regulator knew the monopsonist's productivity, and could set firm-specific minimum wages, she could implement the first-best allocation. The difficulty is that the monopsonist's productivity is *unknown*, or, equivalently, that the minimum wage will apply to a heterogeneous population of monopsonists. This perspective suggests several new insights for minimum wages: that dis-

[28] To see this expression, note a firm of productivity type θ receiving an employment subsidy s has effective productivity $(\theta + s)$, and hence the markdown expression for wages in (2) gives $w^m(\theta + s) = \frac{\varepsilon}{\varepsilon+1}(\theta + s)$.

[29] This discussion raises the question of whether the regulator can tax firm profits. Following Baron and Myerson (1982), I assume the regulator cannot observe firm profits. This assumption is essential here: if profits were observable, then, given observed wages and employment, the regulator could infer productivity and eliminate the firm's private information. One justification is that profits depend partly on unobservable managerial effort; Laffont and Tirole (1986) show that this yields conclusions similar to assuming profits are unobservable, as in Baron and Myerson (1982). Alternatively, Appendix Lemma A5 shows that the regulator's weight on firm profits, λ, can be interpreted as incorporating a linear profit tax.

persion tends to drag down minimum wages; that workers prefer local wage-setting; that minimum wages should be smaller for smaller firms; and that minimum wages are not generally accompanied by employment subsidies, though they do make subsidies more attractive at the margin.

Of course, the model is stylized in important respects. First, I have assumed that workers are homogeneous (except for having different outside options): a natural direction is to examine heterogeneous productivity among workers, not just firms. I have also assumed that the regulator knows the productivity distribution (though not the realization), and that she faces no uncertainty about other aspects of the labor market—notably, she knows the labor supply curve. Another natural direction for future work is to extend the analysis to richer environments in which the regulator also faces uncertainty about these aspects of firms. For instance, if a regulator sets policy over many periods, and initially does not know the productivity distribution F, there may be a case for initially setting low minimum wages in order to infer F from the firm size distribution, before gradually increasing the minimum wage.[30] More generally, the analysis suggests that minimum-wage policy should depend not only on the level of productivity, but also on how much regulators know about its distribution.

[30]In particular, note that any uniform minimum wage $\underline{w}$ that binds on some firms effectively censors the firm productivity distribution below $\underline{w}$, leaving the regulator less informed in future periods.

Appendix

Proof of Lemma 1. I will consider the three problems in turn.

The problem of setting a uniform minimum wage. The problem is continuous in $\underline{w}$. It is without loss to rule out minimum wages below $\underline{o}$, which clearly never bind and hence do not affect surplus. Under our assumption that employment is at least sometimes jointly efficient ($F(\underline{o}) < 1$), setting a minimum wage of $\underline{o}$ ensures strictly positive surplus. Since there is a unit mass of workers, the payoff from setting a minimum wage $\underline{w}$ is at most $\mathbb{E}[\theta \mathbb{1}\{\theta \geq \underline{w}\}]$. Since f is log-concave, we can take this quantity to be arbitrarily small by taking $\underline{w}$ sufficiently large, so we can rule out minimum wages above M, for some constant M. Hence, we can restrict the solution space to $[\underline{o}, M]$. We are then maximizing a continuous function on a compact interval, so there is a solution.

Next, we have to show that the regulator strictly prefers a minimum wage to no minimum wage. By the argument above, setting no minimum wage is equivalent to setting a minimum wage of $\underline{w} = \underline{o}$. The derivation in the proof Lemma 2 will show, starting from $\underline{w} = \underline{o}$, the marginal cost of an increase in the minimum wage is $f(\underline{o})W(\underline{o}, G(\underline{o})) = 0$, since $W(\underline{o}, G(\underline{o})) = 0$. On the other hand, the marginal benefit of an increase in the minimum wage is $\lambda g(\underline{o}) \int_{\tilde{\theta}=\underline{o}}^{\theta^\dagger(\underline{o})} [\tilde{\theta} - \underline{o}] \, \mathrm{d}F(\tilde{\theta})$. If $\lambda > 0$, this latter quantity is strictly positive, since we have assumed $g(\underline{o}) > 0$, and our assumption that there are some firms that sit at the very bottom of the distribution of outside options ($f(\underline{o}) > 0$) and that f has full support implies that $\int_{\tilde{\theta}=\underline{o}}^{\theta^\dagger(\underline{o})} [\tilde{\theta} - \underline{o}] \, \mathrm{d}F(\tilde{\theta}) > 0$. Hence, in the $\lambda > 0$ case, the regulator strictly benefits from a small increase in the minimum wage.

This argument does not apply to the $\lambda = 0$ case, in which the marginal benefit from a small increase in the minimum wage is also zero. In that case, we can use the following alternative argument. Consider a minimum wage $\underline{o} + \varepsilon$, for $\varepsilon \geq 0$ small. Note that, for $\lambda = 0$, the regulator places weight only on worker surplus. Such a minimum wage shuts down firms with productivities in the interval $[\underline{o}, \underline{o}+\varepsilon)$, and binds firms with productivities in the interval $[\underline{o}+\varepsilon, \theta^\dagger(\underline{o}+\varepsilon))$, where $\theta^\dagger(\underline{w})$ is defined as the productivity level at which the minimum wage just binds, implicitly defined by $w^m(\theta^\dagger(\underline{w})) = \underline{w}$.[31] Rearranging this definition gives $\theta^\dagger(\underline{w}) = \underline{w} + G(\underline{w})/g(\underline{w})$. Hence, for $\underline{w} = \underline{o} + \varepsilon$, and for $\varepsilon > 0$, we have $\theta^\dagger(\underline{w}) = \underline{o} + \varepsilon + g(\underline{o})\varepsilon/g(\underline{o}) + O(\varepsilon^2) = \underline{o} + 2\varepsilon + O(\varepsilon^2)$. In light of the expressions in Lemma 2, the marginal benefit of an increase in the minimum wage is then $G(\underline{o} + \varepsilon)[F(\underline{o} + 2\varepsilon) - F(\underline{o})] = 2g(\underline{o})f(\underline{o})\varepsilon^2 + O(\varepsilon^3)$. The marginal cost of an increase in the minimum wage is $f(\underline{o} + \varepsilon)W(\underline{o} + \varepsilon, G(\underline{o} + \varepsilon))$, which is at most $f(\underline{o} + \varepsilon)G(\underline{o} + \varepsilon)\varepsilon =$

[31]The discussion in section 3 explains that we can always find such a productivity level.

$f(\underline{o})g(\underline{o})\varepsilon^2 + O(\varepsilon^3)$. Comparing the costs and benefits shows that, for $\varepsilon > 0$ small, the benefit of a marginal increase in the minimum wage strictly exceeds its cost.

The problem of setting local minimum wages. The same argument applies conditional on each c.

The problem of setting minimum wages by firm size. Since a uniform minimum wage is a trivial example of a minimum wage by firm size, the argument above establishes that the regulator strictly prefers setting minimum wages by firm size to not having any wage. I defer the existence of an optimal minimum wage by firm size to Online Appendix D. □

Proof of Lemma 2. Denote by $V_f(\underline{w})$ the regulator's objective function given firm productivity distribution f and minimum wage $\underline{w}$. Substituting in the firm's best responses to the uniform minimum wage:

$$
\begin{aligned}
V_f(\underline{w}) &= \mathbb{E}_{\theta \sim f}[S_\theta(\underline{w})] \\
&= \int_{\theta=\underline{w}}^{\theta^\dagger(\underline{w})} S_\theta\left(\underline{w}, G(\underline{w})\right) \, \mathrm{d}F(\theta) + \int_{\theta=\theta^\dagger(\underline{w})}^{\infty} S_\theta\left(w^m(\theta), L^m(\theta)\right) \, \mathrm{d}F(\theta).
\end{aligned}
$$

Taking a first-order condition and applying the envelope theorem gives the result. □

Proof of Proposition 1. Adopt the notation in Lemma 2. Lemma A3 establishes that the density ratio $f'(\theta)/f(\theta)$ is strictly decreasing for $\theta < q$. The discussion in text establishes that this implies that, for any $\underline{w} \in (\underline{o}, q_f)$, we have $V_f'(\underline{w})r(\underline{w}) \geq V_{f'}'(\underline{w})$, for $r(\underline{w})$ strictly positive. By the restatement of the results of Milgrom and Shannon (1994) in Quah and Strulovici (2009, pp. 1952–1953), this implies that the set of optimal minimum wages under f dominates the set of optimal minimum wages under f' in the strong set order. Under our assumption that there is a unique optimal minimum wage, we can conclude that the optimal minimum wage under f is weakly higher than the optimal minimum wage under f'. Finally, to prove strictness, note that our argument in the text established that, if $\underline{w}$ is the optimal minimum wage under f, then it cannot satisfy the first-order condition in Lemma 2 under f', and hence $\underline{w}$ cannot be optimal under f'. We thus conclude that the minimum wage is strictly lower under f'. □

Proof of Lemma 3. The argument is very similar to the argument for Proposition 1. Fix some minimum wage $\underline{w}$. The marginal cost of the minimum wage under distribution f is $f(\underline{w})W(\underline{w}, G(\underline{w}))$. The marginal benefit of the minimum wage is $[1-\lambda]G(\underline{w})[F(\theta^\dagger(\underline{w})) - F(\underline{w})] + \lambda g(\underline{w}) \int_{\tilde{\theta}=\underline{w}}^{\theta^\dagger(\underline{w})} [\tilde{\theta} - \underline{w}] \, \mathrm{d}F(\tilde{\theta})$. When the distribution has been shifted to f', the marginal cost and benefit expressions are obtained by replacing f with f' (or F with F'). Following the argument in the proof of Proposition 1, this increases the marginal benefit of $\underline{w}$ more than it increases the marginal cost; as a result, the optimal minimum wage rises. $\square$

Proof of Proposition 2. Follows from discussion in text, using the expressions for the national mode and variance in Lemma A4. $\square$

Proof of Lemma 4 and Proposition 3. See Online Appendix D. $\square$

Proof of Lemma 5. Assuming there is no minimum wage and substituting in the firm's best responses to the uniform minimum wage, the regulator's objective function is

$$\mathbb{E}_{\theta\sim f}[S_\theta(s)] = \int_{\theta\in\Theta} W\big(\hat{w}^m(\theta,s), \hat{L}^m(\theta,s)\big) - s\hat{L}^m(\theta,s) \\ + \lambda\hat{L}^m(\theta,s)\big[\theta + s - w^m(\theta,s)\big]\,\mathrm{d}F(\theta),$$

$$\text{where } \hat{w}^m(\theta,s) := \arg\max_w G(w)\big[\theta + s - w\big] \text{ and } \hat{L}^m(\theta,s) := G\big(\hat{w}^m(\theta,s)\big).$$

Taking a first-order condition and applying the envelope theorem gives the result. $\square$

Proof of Proposition 4. For (i), rearranging (11) gives $s^* \geq 0 \iff \lambda \geq \lambda^* := 1-\eta(0)$. In the isoelastic case, $\eta(s) = \frac{\varepsilon}{\varepsilon+1}$, and hence $\lambda^* = 1/[\varepsilon+1] \in (0,1)$. For (ii), define expected social welfare given minimum wage $\underline{w}$ and subsidy s as $V(\underline{w}, s)$. Define maximized expected social welfare given subsidy s as $\hat{V}(s) := \max_{\underline{w}} V(\underline{w}, s)$. This is well-defined, by the argument we made in Lemma 1, which is simple to extend to the case of the subsidy. Denote by $\underline{w}^0$ the optimal minimum wage when $s = 0$. By the envelope theorem, $\frac{\mathrm{d}\hat{V}(s)}{\mathrm{d}s}\,|_{s=0} = \frac{\partial V(s,\underline{w})}{\partial s}\,|_{s=0,\underline{w}=\underline{w}_0}$. Evaluating this derivative, and using the first-order condition for the no-subsidy minimum wage to simplify:

$$\left.\frac{\mathrm{d}\hat{V}(s)}{\mathrm{d}s}\right|_{s=0} = \lambda g(\underline{w}^0)\int_{\tilde{\theta}=\underline{w}^0}^{\underline{w}^0+\underline{w}^0/\varepsilon} [\tilde{\theta} - \underline{w}^0]\mathrm{d}F(\tilde{\theta}) \\ + \left[\lambda - \frac{1}{\varepsilon+1}\right]\int_{\tilde{\theta}=\underline{w}^0+\underline{w}^0/\varepsilon}^{\infty} G(w^m(\tilde{\theta}))\,\mathrm{d}F(\tilde{\theta}).$$

For an interior minimum wage, both integrals are positive, and hence the expression is increasing in λ. By inspection, for λ sufficiently small (but above 0), the whole expression is negative; for λ sufficiently large (but below 1), it is positive. Since the expression is continuous and increasing, we conclude that the regulator uses a subsidy so long as $\lambda \geq \lambda^{\mathrm{MW}}$, for some $\lambda^{\mathrm{MW}} \in (0,1)$. For (iii), evaluating the expression at $\lambda = \lambda^* = 1/[\varepsilon+1]$, the second term is zero, and hence the integral is positive, so $\lambda^{\mathrm{MW}} < \lambda^*$. $\square$

References

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

ActivPayroll. 2025. "Malaysia's Minimum Wage Increase 2025: What Employers and Employees Need to Know." https://www.activpayroll.com/news-articles/malaysias-minimum-wage-increase. Accessed: 2026-03-29.

Adhvaryu, Achyuta, Emir Murathanoglu, and Anant Nyshadham. 2023. "On the allocation and impacts of managerial training."Technical report, National Bureau of Economic Research.

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

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

Almunia, Miguel, and David Lopez-Rodriguez. 2018. "Under the radar: The effects of monitoring firms on tax compliance." *American Economic Journal: Economic Policy* 10 (1): 1–38.

Amador, Manuel, and Kyle Bagwell. 2022. "Regulating a monopolist with uncertain costs without transfers." *Theoretical Economics* 17 (4): 1719–1760.

Amirapu, Amrit, and Michael Gechter. 2020. "Labor regulations and the cost of corruption: Evidence from the Indian firm size distribution." *Review of Economics and Statistics* 102 (1): 34–48.

Andrews, Dan, Chiara Criscuolo, and Peter N Gal. 2016. "The best versus the rest: the global productivity slowdown, divergence across firms and the role of public policy."

Anxo, Dominique. 2024. "The EU Directive on Adequate Minimum Wages: Preliminary Assessment of Potential Impacts." Report 2024:2, Swedish Institute for European Policy Studies (SIEPS). Stockholm.

ASEAN. 2026. "ASEAN Briefing: Salary and Minimum Wage in Indonesia." https://www.aseanbriefing.com/doing-business-guide/indonesia/human-resources-and-payroll/minimum-wage. Accessed: 2026-03-28.

Ashenfelter, Orley, and Robert S Smith. 1979. "Compliance with the minimum wage law." *Journal of Political Economy* 87 (2): 333–350.

Autor, David H., Alan Manning, and Christopher L. Smith. 2016. "The Contribution of the Minimum Wage to US Wage Inequality over Three Decades: A Reassessment." *American Economic Journal: Applied Economics* 8 (1): 58–99.

Bagnoli, Mark, and Ted Bergstrom. 2005. "Log-concave probability and its applications." *Economic theory* 26 (2): 445–469.

Baron, David P, and Roger B Myerson. 1982. "Regulating a monopolist with unknown costs." *Econometrica: Journal of the Econometric Society*: 911–930.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. 2013. "Cross-country differences in productivity: The role of allocation and selection." *American Economic Review* 103 (1): 305–334.

Basu, Arnab K, Nancy H Chau, and Ravi Kanbur. 2015. "Contractual dualism, market power and informality." *The Economic Journal* 125 (589): 1534–1573.

Berger, David, Kyle Herkenhoff, and Simon Mongey. 2022. "Labor market power." *American Economic Review* 112 (4): 1147–1193.

Berger, David, Kyle Herkenhoff, and Simon Mongey. 2025. "Minimum wages, efficiency, and welfare." *Econometrica* 93 (1): 265–301.

Bernard, Andrew B, Jonathan Eaton, J Bradford Jensen, and Samuel Kortum. 2003. "Plants and productivity in international trade." *American Economic Review* 93 (4): 1268–1290.

Bijnens, Gert, Helene Hall, Hugo Monnery, and Laura Nicolae. 2024. "The Effect of Inflation Indexation On Employment: Evidence From Belgium." *Available at SSRN 4934002*.

Bijnens, Gert, Simon Jäger, and Benjamin Schoefer. 2025. "What Does Consulting Do?"Technical report, National Bureau of Economic Research.

Bils, Mark, Peter J Klenow, and Cian Ruane. 2021. "Misallocation or mismeasurement?" *Journal of Monetary Economics* 124: S39–S56.

Birnbaum, Zygmunt William. 1948. "On random variables with comparable peakedness." *The Annals of Mathematical Statistics* 19 (1): 76–81.

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J Terry. 2018. "Really uncertain business cycles." *Econometrica* 86 (3): 1031–1065.

Boal, William M, and Michael R Ransom. 1997. "Monopsony in the labor market." *Journal of economic literature* 35 (1): 86–112.

Boeri, Tito. 2012. "Setting the minimum wage." *Labour Economics* 19 (3): 281–290.

Bossavie, Laurent, Ayşenur Acar Erdoğan, and Mattia Makovec. 2019. "The impact of the minimum wage on firm destruction, employment and informality."Technical report.

Brown, Charles. 1999. "Minimum Wages, Employment, and the Distribution of Income." In *Handbook of Labor Economics*, edited by Orley C. Ashenfelter and David Card, vol. 3, chap. 32, 2101–2163: Elsevier.

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

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

Card, David, and Alan B. Krueger. 1994. "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania." *American Economic Review* 84 (4): 772–793.

Casella, George, and Roger Berger. 2024. *Statistical inference*: Chapman and Hall/CRC.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux. 2012. "The productivity advantages of large cities: Distinguishing agglomeration from firm selection." *Econometrica* 80 (6): 2543–2594.

Comisión Nacional de los Salarios Mínimos. 2025. "Incremento a los Salarios Mínimos para 2026." https://www.gob.mx/conasami/articulos/incremento-a-los-salarios-minimos-para-2026?idiom=es. Accessed: 2026-03-28.

Cunningham, Cindy, Lucia Foster, Cheryl Grim, John Haltiwanger, Sabrina Wulff Pabilonia, Jay Stewart, and Zoltan Wolf. 2023. "Dispersion in dispersion: measuring establishment-level differences in productivity." *Review of Income and*

Wealth 69 (4): 999–1032.
Dautović, Ernest, Harald Hau, and Yi Huang. 2024. "Consumption response to minimum wages: evidence from Chinese households." *Review of Economics and Statistics*: 1–47.
Davis, Lucas W, and Lutz Kilian. 2011. "The allocative cost of price ceilings in the US residential market for natural gas." *Journal of Political Economy* 119 (2): 212–241.
Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda. 2020. "Changing business dynamism and productivity: Shocks versus responsiveness." *American Economic Review* 110 (12): 3952–3990.
Dentons. 2025. "Costa Rica Approves Minimum Wage Increase Effective 2026." https://www.dentons.com/en/insights/articles/2025/december/17/costa-rica-approves-minimum-wage-increase-effective-2026. Accessed: 2026-03-28.
Dharmapala, Dhammika, Joel Slemrod, and John Douglas Wilson. 2011. "Tax policy and the missing middle: Optimal tax remittance with firm-level administrative costs." *Journal of Public Economics* 95 (9-10): 1036–1047.
Dickens, Richard. 2023. "How are minimum wages set?" *IZA World of Labor*.
Doligalski, Paweł, Piotr Dworczak, Mohammad Akbarpour, and Scott Duke Kominers. 2025. *Optimal redistribution via income taxation and market design*: School of Economics, University of Bristol.
Dube, Arindrajit, and Attila Lindner. 2021. "City limits: What do local-area minimum wages do?" *Journal of Economic Perspectives* 35 (1): 27–50.
Dube, Arindrajit, and Attila Lindner. 2024. "Chapter 4 - Minimum wages in the 21st century." vol. 5 of Handbook of Labor Economics, 261–383: Elsevier.
Dworczak, Piotr, Scott Duke Kominers, and Mohammad Akbarpour. 2021. "Redistribution through markets." *Econometrica* 89 (4): 1665–1698.
Eurofound. 2023. *Minimum Wages in 2023: Annual Review*, Minimum Wages in the EU Series. Luxembourg: Publications Office of the European Union.
European Parliament and Council of the European Union. 2022. "Directive (EU) 2022/2041 of the European Parliament and of the Council of 19 October 2022 on Adequate Minimum Wages in the European Union." OJ L 275, 25.10.2022, pp. 33–59.
Eurostat. 2026. "Now available: first 2026 data on minimum wages." https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20260130-2. Eurostat News Article. Accessed March 13, 2026.
Fair Work Commission. 2026. "Which award applies." https://www.fwc.gov.au/which-award-applies. Accessed: 2026-04-06.
Fortin, Nicole M, Thomas Lemieux, and Neil Lloyd. 2021. "Labor market institutions and the distribution of wages: The role of spillover effects." *Journal of Labor Economics* 39 (S2): S369–S412.
Garicano, Luis, Claire Lelarge, and John Van Reenen. 2016. "Firm size distortions and the productivity distribution: Evidence from France." *American Economic Review* 106 (11): 3439–3479.
Garnero, Andrea, Stephan Kampelmann, and François Rycx. 2015. "Minimum wage systems and earnings inequalities: Does institutional diversity matter?" *European Journal of Industrial Relations* 21 (2): 115–130.

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

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

Gopalan, Radhakrishnan, Barton H. Hamilton, Ankit Kalda, and David Sovich. 2021. "State Minimum Wages, Employment, and Wage Spillovers: Evidence from Administrative Payroll Data." *Journal of Labor Economics* 39 (3): 673–707.

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

Holmström, Bengt. 1979. "Moral hazard and observability." *The Bell journal of economics*: 74–91.

Hsieh, Chang-Tai, and Peter J Klenow. 2009. "Misallocation and manufacturing TFP in China and India." *The Quarterly journal of economics* 124 (4): 1403–1448.

International Labour Organization. 2015. "2.1 The Logic Behind Different Minimum Wage Systems." Chapter 2 in "How Many Different Minimum Wage Rates Should There Be?".

International Labour Organization. 2016a. "Minimum Wage Policy Guide."Technical report, International Labour Office, Geneva.

International Labour Organization. 2016b. "Learn More About Minimum Wage Rates by Enterprise Size." ILO web resource.

Japanese Ministry of Health, Labour and Welfare. 2024. "Types of Minimum Wage." https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/koyou_roudou/roudoukijun/chingin/newpage_43885.html. Japanese page. Accessed: 2026-03-28.

Jardim, Ekaterina, and Emma van Inwegen. 2023. "Business churn, labor intensity, and the minimum wage."Technical report, working paper.

Johnson, Justin P, and David P Myatt. 2006. "On the simple economics of advertising, marketing, and product design." *American Economic Review* 96 (3): 756–784.

Kang, Zi Yang, and Mitchell Watt. 2024. "Optimal In-Kind Redistribution." *arXiv preprint arXiv:2409.06112*.

Kaplow, Louis. 2019. "Optimal regulation with exemptions." *International Journal of Industrial Organization* 66: 1–39.

Katz, Lawrence F, and Alan B Krueger. 1992. "The effect of the minimum wage on the fast-food industry." *ILR Review* 46 (1): 6–21.

Keen, Michael, and Jack Mintz. 2004. "The optimal threshold for a value-added tax." *Journal of Public Economics* 88 (3-4): 559–576.

Kehrig, Matthias et al. 2011. "The cyclicality of productivity dispersion." *US Census Bureau Center for Economic Studies Paper No. CES-WP-11-15*.

Kehrig, Matthias, and Nicolas Vincent. 2025. "Good dispersion, bad dispersion." *Review of Economic Studies*: rdaf042.

Kline, Patrick. 2025. "Labor market monopsony: Fundamentals and frontiers." *Handbook of Labor Economics* 6: 655–728.

Kolotilin, Anton, and Andriy Zapechelnyuk. 2025. "Persuasion meets delegation." *Econometrica* 93 (1): 195–228.

Laffont, Jean-Jacques, and Jean Tirole. 1986. "Using cost observation to regulate firms." *Journal of Political Economy* 94 (3, Part 1): 614–641.

Lee, David, and Emmanuel Saez. 2012. "Optimal minimum wage policy in com-

petitive labor markets." *Journal of Public Economics* 96 (9-10): 739–749.
Lehmann, Erich Leo, and George Casella. 1998. *Theory of point estimation*: Springer.
Lentz, Rasmus, and Dale T Mortensen. 2008. "An empirical model of growth through product innovation." *Econometrica* 76 (6): 1317–1373.
Levin-Waldman, Oren M. 2001. *The case of the minimum wage: Competing policy models*: State University of New York Press.
Lewis, Tracy R, and David EM Sappington. 1988a. "Regulating a monopolist with unknown demand." *The American Economic Review*: 986–998.
Lewis, Tracy R, and David EM Sappington. 1988b. "Regulating a monopolist with unknown demand and cost functions." *The RAND Journal of Economics*: 438–457.
Loertscher, Simon, and Ellen V Muir. 2023. "Optimal labor procurement under minimum wages and monopsony power."Technical report, MIT Sloan working paper.
Loertscher, Simon, and Ellen V Muir. 2024. "Market power, randomization and regulation." *International Journal of Industrial Organization* 96: 103081.
Luca, Dara Lee, and Michael Luca. 2019. "Survival of the fittest: The impact of the minimum wage on firm exit."Technical report, National Bureau of Economic Research.
Luttmer, Erzo FP. 2007. "Does the minimum wage cause inefficient rationing?."
Manning, Alan. 2003. "A Generalised Model of Monopsony." CEP Discussion Paper 499, Centre for Economic Performance, London School of Economics.
Manning, Alan. 2011. "Imperfect competition in the labor market." In *Handbook of labor economics*, vol. 4, 973–1041: Elsevier.
Manning, Alan. 2021. "The elusive employment effect of the minimum wage." *Journal of Economic Perspectives* 35 (1): 3–26.
Maskin, Eric, and John Riley. 2000. "Asymmetric auctions." *The review of economic studies* 67 (3): 413–438.
Meyer, Jack. 1987. "Two-moment decision models and expected utility maximization." *The American economic review*: 421–430.
Milgrom, Paul, and Chris Shannon. 1994. "Monotone comparative statics." *Econometrica: Journal of the Econometric Society*: 157–180.
Ministerio de Trabajo y Desarrollo Laboral. 2024. "Nuevo Salario Mínimo se hará efectivo en la segunda quincena de enero." https://www.mitradel.gob.pa/nuevo-salario-minimo-se-hara-efectivo-en-la-segunda-quincena-de-enero/. Accessed: 2026-03-28.
Ministerio de Trabajo y Desarrollo Laboral. 2025. "Trabajadores y empleadores logran consenso y acuerdan nuevas tasas de salario mínimo." https://www.mitradel.gob.pa/trabajadores-y-empleadores-logran-consenso-y-acuerdan-nuevas-tasas-de-salario-minimo/. Accessed: 2026-03-28.
Myerson, Roger B. 1981. "Optimal auction design." *Mathematics of Operations Research* 6 (1): 58–73.
Neumark, David, and William L. Wascher. 2008. *Minimum Wages*. Cambridge, MA: MIT Press.
Nigai, Sergey. 2017. "A tale of two tails: Productivity distribution and the gains from trade." *Journal of International Economics* 104: 44–62.

NZCTU Te Kauae Kaimahi. 2025. “A New Deal for Workers.”Technical report, NZCTU Te Kauae Kaimahi.
Onji, Kazuki. 2009. “The response of firms to eligibility thresholds: Evidence from the Japanese value-added tax.” *Journal of Public Economics* 93 (5-6): 766–775.
Philippines National Wages and Productivity Commission. 2025. “Daily Minimum Wage Rates.” https://nwpc.dole.gov.ph/. Official website of the NWPC. Accessed: 2026-03-28.
Presidencia de la República Dominicana. 2026. “Ministerio de Trabajo llama a empresarios a cumplir con el pago del aumento del 8% del salario mínimo.” https://presidencia.gob.do/noticias/ministerio-de-trabajo-llama-empresarios-cumplir-con-el-pago-del-aumento-del-8-del-salario. Accessed: 2026-03-28.
Quah, John K-H, and Bruno Strulovici. 2009. “Comparative statics, informativeness, and the interval dominance order.” *Econometrica* 77 (6): 1949–1992.
Rao, Nirupama L, and Max Risch. 2026. “Who’s afraid of the minimum wage? Measuring the impacts on independent businesses using matched US tax returns.” *The Quarterly Journal of Economics* 141 (1): 373–427.
Restuccia, Diego. 2025. “On Productivity and Distortions across Countries.”Technical report, National Bureau of Economic Research.
Robinson, Joan. 1933. *The Economics of Imperfect Competition*. London: Macmillan.
Rocheteau, Guillaume, and Murat Tasci. 2008. “Positive and Normative Effects of a Minimum Wage.” Working Paper 08-01, Federal Reserve Bank of Cleveland.
Rohlin, Shawn M. 2011. “State minimum wages and business location: Evidence from a refined border approach.” *Journal of Urban Economics* 69 (1): 103–117.
Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Grace Cooper, Julia A. Rivera Drew, Stephanie Richards, Renae Rogers, Jonathan Schroeder, and Kari C.W. Williams. 2025. “IPUMS USA: Version 16.0.” Dataset.
Sandmo, Agnar. 1971. “On the theory of the competitive firm under price uncertainty.” *The American Economic Review* 61 (1): 65–73.
Schivardi, Fabiano, and Roberto Torrini. 2008. “Identifying the effects of firing restrictions through size-contingent differences in regulation.” *Labour Economics* 15 (3): 482–511.
Schulten, Thorsten, and Torsten Müller. 2021. “A paradigm shift towards Social Europe? The proposed Directive on adequate minimum wages in the European Union.” *Italian Labour Law e-Journal* 14 (1): 1–19.
Schwellnus, Cyrille. 2026. “Strengthening formal employment and social protection.” In *OECD Economic Surveys: Philippines 2026*, Paris: OECD Publishing. Accessed: 2026-03-29.
Secretaría de Trabajo y Seguridad Social. 2025. “Tabla de Salario Mínimo Vigente a partir del 01 de enero de 2025.” https://www.trabajo.gob.hn/tabla-de-salario-minimo/. Tabla de Salario Mínimo y Bono Educativo en Honduras 2025. Accessed: 2026-03-28.
Seltzer, Andrew J. 1995. “The political economy of the Fair Labor Standards Act of 1938.” *Journal of Political Economy* 103 (6): 1302–1342.
Simon, Andrew, and Matthew Wilson. 2021. “Optimal minimum wage setting in a federal system.” *Journal of Urban Economics* 123: 103336.
Sinn, Hans-Werner. 1983. *Economic decisions under uncertainty*: Springer Science & Business Media.

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

Syverson, Chad. 2004. "Product substitutability and productivity dispersion." *Review of Economics and Statistics* 86 (2): 534–550.

Treskon, Mark, John Marotta, Prasanna Rajasekaran, Kriti Ramakrishnan, Aaron Shroyer, and Solomon Greene. 2021. "Do the Effects of a Regulatory Patchwork Justify State Preemption of Local Laws?"Technical report, Urban Institute.

UK Trades Union Congress. 2013. "Bring back wages councils to tackle living standards crisis." https://www.tuc.org.uk/news/bring-back-wages-councils-tackle-living-standards-crisis.

U.S. Census Bureau. 1998. *Statistical Abstract of the United States: 1998*. Washington, DC: U.S. Census Bureau.

U.S. Department of Labor, Wage and Hour Division. 2023. "Handy Reference Guide to the Fair Labor Standards Act." https://www.dol.gov/agencies/whd/compliance-assistance/handy-reference-guide-flsa. Revised November 2023. Accessed: 2026-03-29.

U.S. Department of Labor, Wage and Hour Division. 2026. "State Minimum Wage Laws." https://www.dol.gov/agencies/whd/minimum-wage/state. Updated January 1, 2026. Accessed: 2026-03-29.

Vietnam Social Security. 2025. "Regional minimum wages to increase from January 1, 2026." https://vss.gov.vn/english/news/Pages/vietnam-social-security.aspx?CateID=0&ItemID=12822. Accessed: 2026-03-28.

White, T Kirk, Jerome P Reiter, and Amil Petrin. 2018. "Imputation in US manufacturing data and its implications for productivity dispersion." *Review of Economics and Statistics* 100 (3): 502–509.

Wong, Wing-Keung, and Chenghu Ma. 2008. "Preferences over location-scale family." *Economic Theory* 37 (1): 119–146.

Zhou, Mi. 2016. "Minimum wage setting, implementation and working conditions in formal and informal sectors of the garment industry in Pakistan."Technical report, International Labour Organization. Accessed: 2026-03-29.

Online Appendix

The online appendix is in four parts. Online Appendix A contains proofs. Online Appendix B discusses extending the results in three ways: to relax the assumption of efficient rationing; to allow for other inputs; and to accommodate decreasing returns to scale. Online Appendix B also shows that the effects of an increase in the dispersion of productivity are ambiguous without some restrictions on the distributions. Online Appendix C contains additional empirical details. Online Appendix D describes the optimal mechanism in section 5 in greater detail.

Online Appendix A. Proofs

A.1 Intermediate steps for proofs of statements in the main text

Lemma A1. Say that a minimum wage schedule $\underline{w}^*(\cdot)$ solves

$$\max_{\underline{w}:\ \mathbb{R}_+\to\mathbb{R}_+} \int_{\theta\in\Theta} S_\theta\big(\hat{w}(\theta,\underline{w}),\hat{L}(\theta,\underline{w})\big)\,\mathrm{d}F(\theta),\ \text{where}$$

$$\big(\hat{w}(\theta,\underline{w}),\hat{L}(\theta,\underline{w})\big)=\operatorname*{arg\,max}_{w\geq\underline{w}(L),\ L\leq G(w)}\Big\{\pi_\theta(w,L)\Big\},\ \text{breaking ties in the regulator's favor.}$$

Then for each $\theta\in\Theta$, define $\hat{w}^*(\theta):=\hat{w}(\theta,\underline{w}^*)$, $\hat{L}^*(\theta):=\hat{L}(\theta,\underline{w}^*)$. Then $(\hat{w}^*(\cdot),\hat{L}^*(\cdot))$ solves

$$\max_{\hat{w}:\ \Theta\to\mathbb{R},\ \hat{L}:\ \Theta\to\mathbb{R}} \int_{\theta\in\Theta} S_\theta\big(\hat{w}(\theta),\hat{L}(\theta)\big)\,\mathrm{d}F(\theta),$$

$$\text{subject to (incentive compatibility): } \theta\in\operatorname*{arg\,max}_{\hat{\theta}}\pi_\theta(\hat{w}(\hat{\theta}),\hat{L}(\hat{\theta})).$$

$$\text{(participation): } \pi_\theta(\hat{w}(\theta),\hat{L}(\theta))\geq 0.$$

$$\text{(labor supply): } \hat{L}(\theta)\leq G(\hat{w}(\theta)),\ \text{for all } \theta\in\Theta.$$

Conversely, say some $(\hat{w}^*(\cdot),\hat{L}^*(\cdot))$ solves the latter problem. Then we can define $\underline{w}^*:\mathbb{R}_+\to\mathbb{R}$ such that for each type $\theta\in\Theta$, $\big(\hat{w}(\theta,\underline{w}^*),\hat{L}(\theta,\underline{w}^*)\big)=(\hat{w}^*(\theta),\hat{L}^*(\theta))$, defining $\hat{w}(\cdot,\cdot)$ as in the former problem. Moreover, $\underline{w}^*$ solves the former problem.

Proof of Lemma A1. For the first statement, I proceed in two steps. First, I show that $(\hat{w}^*(\cdot),\hat{L}^*(\cdot))$ satisfies the constraints in the latter problem. Labor supply holds by construction. The firm can always choose $L=0$, which gives zero profit, and hence at their optimal employment-wage pair they must make non-negative profit; hence, the participation constraint holds. Incentive compatibility holds because, for any type θ, the allocation assigned to θ' is feasible in the original problem. We therefore conclude that $(\hat{w}^*(\cdot), \hat{L}^*(\cdot))$ is indeed a candidate solution to the latter problem. Second, I show that $(\hat{w}^*(\cdot), \hat{L}^*(\cdot))$ is indeed the optimal solution to the latter problem. Suppose by way of contradiction that there were some $(\hat{w}(\cdot), \hat{L}(\cdot))$ which ensured strictly higher payoff for the regulator while still satisfying the constraints. Denote by $\hat{\mathcal{L}}$ the set of employment levels assigned by the allocation: that is, $\hat{\mathcal{L}} = \{\ell \in \mathbb{R} : \ell = \hat{L}(\theta) \text{ for some } \theta \in \Theta\}$. Note that it must be the case that, if $\hat{L}(\theta) = \hat{L}(\theta') > 0$, then $\hat{w}(\theta) = \hat{w}(\theta')$: otherwise, the allocation would not be incentive-compatible, since the type assigned the higher wage would deviate to mimic the other type. In consequence, for any $L > 0$, we can write $\hat{w}(\theta) = \hat{w}(\hat{L}(\theta))$. Then consider the indirect mechanism

$$\underline{\hat{w}}(L) = \begin{cases} \hat{w}(L), \text{ if } L > 0, L \in \hat{\mathcal{L}} \\ \text{arbitrary, if } L = 0, \\ Q \text{ if } L \notin \hat{\mathcal{L}}, \end{cases} \tag{A1}$$

for some $Q > \bar{\theta}$: that is, Q is large enough that no firm would be willing to operate a minimum wage Q. This indirect mechanism ensures $(\hat{w}(\theta, \underline{\hat{w}}), \hat{L}(\theta, \underline{\hat{w}})) = (\hat{w}(\theta), \hat{L}(\theta))$. As a result, under our assumption that $(\hat{w}(\cdot), \hat{L}(\cdot))$ ensures strictly higher payoff for the regulator than $(w^*(\cdot), L^*(\cdot))$, we would also have that $\underline{\hat{w}}(\cdot)$ delivers a strictly higher payoff for the regulator than $\underline{w}^*(\cdot)$, which contradicts that $\underline{w}^*(\cdot)$ was the solution to the former problem.

Now we go in the opposite direction. Say some $(\hat{w}^*(\cdot), \hat{L}^*(\cdot))$ solves the latter problem. Construct the indirect mechanism $\underline{w}^*(\cdot)$ as in equation (A1), replacing $\hat{w}(\cdot)$ with $\hat{w}^*$ and $\hat{\mathcal{L}}$ with $\hat{\mathcal{L}}^*$ (defined in the corresponding way). Then the same argument shows that $\hat{w}^*$ solves the former problem. □

Lemma A2. For a PDF f with a unique mode q_f, and some other PDF f', the following are equivalent:

(i) f and f' belong to the same location-scale family, and have the same mode, but f' has strictly larger variance.

(ii) $f'(x) = \frac{1}{s} f\left(q_f + \frac{x - q_f}{s}\right)$, for some $s > 1$.

Proof of Lemma A2. First, I show (i) implies (ii). Denoting the base density of the location-scale family to which f and f' belong by h:

$$f(x) = \frac{1}{\sigma} h\left(\frac{x - \mu}{\sigma}\right), \quad f'(x) = \frac{1}{\sigma'} h\left(\frac{x - \mu'}{\sigma'}\right),$$

where (μ, σ) are the location and scale of f, and (μ', σ') are the location and scale of f'. Denote by q_h the mode of h. Since q_f is the mode of f and f', $q_f = \mu + \sigma q_h = \mu' + \sigma' q_h$. As a result, we have $\mu = q_f - \sigma q_h$ and $\mu' = q_f - \sigma' q_h$. The variance of the distribution f is $\sigma^2 \mathrm{Var}(h)$, and the variance of f' is $\sigma'^2 \mathrm{Var}(h)$, so since f' has strictly larger variance, and hence $s := \frac{\sigma'}{\sigma} > 1$. Then:

$$
\begin{aligned}
f'(x) &= \frac{1}{\sigma'} h\left(\frac{x-\mu'}{\sigma'}\right) \\
&= \frac{1}{\sigma'} h\left(\frac{x - q_f + \sigma' q_h}{\sigma'}\right) \\
&= \frac{1}{s\sigma} h\left(\frac{x - q_f}{s\sigma} + q_h\right) \\
&= \frac{1}{s\sigma} h\left(\frac{q_f + \frac{x-q_f}{s} - \mu}{\sigma}\right) = \frac{1}{s} f\left(q_f + \frac{x - q_f}{s}\right),
\end{aligned}
$$

where the first line uses the location-scale expression for the density f', the second line substitutes in our expression for μ', the third line substitutes in $\sigma' = s\sigma$, and the fourth line substitutes in $q_h = [q_f - \mu]/\sigma$.

In the opposite direction, it is clear that f and f' belong to the same location-scale family, that f and f' have the same maximizer (and hence the same mode), and that f' has a strictly larger variance. □

Lemma A3. Say f' is a shape-preserving spread of f, and that f satisfies the restrictions described in section 2. For $\theta \in \Theta$, define the density ratio $r(\theta) := f'(\theta)/f(\theta)$. The ratio $r(\theta)$ is strictly decreasing for $\theta < q_f$.

Proof of Lemma A3. It suffices to show that $\ln r(\theta)$ is strictly decreasing. Write $\ell(\theta) := \log f(\theta)$ and for $\theta < q_f$, define $\hat{\theta}(\theta) := q_f + [\theta - q_f]/s$, so that $\ln r(\theta) = -\ln s + \ell(\hat{\theta}(\theta)) - \ell(\theta)$. Since f is log-concave, ℓ is concave, and hence has decreasing one-sided derivatives, so $D_+\ell(\theta) > D_+\ell(\hat{\theta}(\theta))$; since f is unimodal and $\hat{\theta}(\theta) < q_f$, $D_+\ell(\hat{\theta}(\theta)) > 0$, and hence $D_+\ell(\theta) \geq D_+\ell(\hat{\theta}(\theta)) > 0$. As a result, $D_+[\ell(\hat{\theta}(\theta)) - \ell(\theta)] = \frac{1}{s} D_+\ell(\hat{\theta}(\theta)) - D_+\ell(\theta) < 0$, and hence $\ell(\hat{\theta}(\theta)) - \ell(\theta)$ is strictly decreasing in θ for $\theta < q_f$. As a result, $\ln r(\theta)$ is strictly decreasing. □

Lemma A4. In the setting described in Proposition 2, the mode of the national distribution is $\mathbb{E}[m_c]$, and the variance of the national distribution is $\sigma^2 + \mathrm{Var}(m_c)$.

Proof of Lemma A4. Since the distributions are symmetric and unimodal, the mode m_c of each conditional distribution $f(\cdot \mid c)$ is also its mean. Since the national distribution is in the same location-scale family, the same is true for the national distribution. Since the national distribution is a weighted average of these distributions, its average and therefore its mode is $\mathbb{E}[m_c]$. By the law of total variance, the variance of national distribution is $\bar{\sigma}^2 = \sigma^2 + \mathrm{Var}\big(\mathbb{E}[\theta \mid c]\big) = \sigma^2 + \mathrm{Var}(m_c)$, where the second equality uses the fact that the symmetry of the distributions implies the mode of each distribution is also the mean. □

Lemma A5. Consider the settings described in section 2 and 6, with no tax on firm profits, and weight on firm profits λ. If we allow for a linear tax on firm profits $\tau \in (0, 1)$ that is remitted to workers, then the problem coincides with the original problem, with the regulator's weight on firm profits reparameterized as $\tilde{\lambda} := \tau + \lambda[1 - \tau]$.

Proof of Lemma A5. I will suppose that the firm pays taxes on any subsidy it receives; a similar argument applies to the case in which subsidies are untaxed. First, note that the monopsonist's problem is unchanged: the monopsonist previously maximized $\pi_\theta(w, L) + sL$, and now maximizes $(1 - \tau)(\pi_\theta(w, L) + sL)$, so its labor and employment choices are the same. Second, consider the regulator's payoff under linear tax τ:

$$S_\theta(w, L, \tau) = \underbrace{W(w, L) - sL + \tau[\pi_\theta(w, L) + sL]}_{\text{Worker payoff}} + \lambda \underbrace{(1 - \tau)[\pi_\theta(w, L) + sL]}_{\text{Firm payoff}}, \tag{A2}$$

where the third term in the expression labelled "worker payoff" in (A2) is the tax amount remitted to workers. Rearranging this expression:

$$S_\theta(w, L, \tau) = W(w, L) - sL + [\tau + \lambda(1 - \tau)][\pi_\theta(w, L) + sL] \tag{A3}$$

$$= W(w, L) - sL + \tilde{\lambda}[\pi_\theta(w, L) + sL], \tag{A4}$$

where (A4) coincides with the regulator's payoff absent the tax, reparameterizing her weight on firm profits. □

Online Appendix B. Additional results

B.1 Alternative rationing assumptions

I now discuss the effects of alternative rationing assumptions. Note that a uniform minimum wage (or a local minimum wage) does not create any rationing, under our assumptions that (1) the monopsonist's production technology has constant returns to scale, and (2) ties are broken in the regulator's favor. That is, if the regulator sets a constant minimum wage $\underline{w}$, firms with productivity $\theta < \underline{w}$) will not operate, firms with productivity

$\theta \in [\underline{w}, \theta^{\dagger}(\underline{w}))$ will set wage $\underline{w}$ and employ $G(\underline{w})$ workers, and firms with productivity $\theta \geq \theta^{\dagger}(\underline{w})$ will set monopsony wages and employment. In none of these cases is there rationing. In consequence, our assumption of efficient rationing does not have any effect on the results in sections 3 or 4. Adding uniform subsidies also does not create rationing: we can think of each firm's productivity as $\theta + s$, and then the same logic as above applies. As a result, our rationing assumptions do not affect the results in section 6.

Now consider the results on size-dependent minimum wages in section 5. There, Lemma A7 showed that the optimal mechanism does in general use rationing, and hence alternative assumptions on rationing may affect the results. Lemma 4 showed that wages are weakly increasing without using the regulator's objective at all, and hence it is clear that the optimal schedule will still assign weakly higher wages to firms, but Proposition 3 arguing that the schedule is strictly increasing relied on efficient rationing. There are alternative ways to specify the costs of rationing: for instance, we could assume that workers are rationed maximally *inefficiently*, or that there is some mixture of efficient and inefficient rationing. Another alternative is that the regulator assigns further costs to rationing even *beyond* labor supply costs—for instance, the costs arising from workers taking other steps to get to the front of job queues.

As an extreme case, suppose the regulator has an arbitrarily strong preference to *never* ration employment. In this case, we can think of the regulator's problem as corresponding to the problem in section 5, with the additional constraint that, at each assigned wage $\hat{w}(\theta)$, the corresponding labor supply is $\hat{L}(\theta) = G(\hat{w}(\theta))$. This has the effect that the regulator's assigned wage-employment pairs must lie along the labor supply curve: her only remaining tool is to choose a menu of admissible wage levels and allow the delegate to select among them. Then the problem is formally very similar to the one considered in section 7 of Kolotilin and Zapechelnyuk (2025). The results in Corollary 3 of that paper imply that, under some conditions on the productivity distribution f and labor supply curve G, the regulator will impose only a minimum wage, and no further regulation.

B.2 Allowing for other inputs in the production function

In the main text, I claimed that the model described in section 2, in which the only input is labor, can be viewed as a reduced-form representation of a model in which the monopsonist has a CRS production technology including labor and other inputs, for which the monopsonist is a price-taker. I now make that claim precise.

Suppose a firm of total factor productivity $\tilde{\theta}$ chooses labor L and a vector $Z \in \mathbb{R}_{+}^{K}$ of other inputs, and that its output is $\tilde{\theta}\mathcal{F}(L, Z)$, where $\mathcal{F} \colon \mathbb{R}_{+}^{K+1} \rightarrow$

$\mathbb{R}_+$ is the CRS production technology.[32] Denote by $p \in \mathbb{R}_+^K$ the vector of input prices for non-labor inputs, and by w the wage. The firm's problem is:

$$\max_{w \geq \underline{w}, L \leq G(w), Z} \tilde{\theta}\mathcal{F}(L, Z) - wL - p \cdot Z \iff \max_{w \geq \underline{w}, L \leq G(w)} \{\max_{Z} \tilde{\theta}\mathcal{F}(L, Z) - wL - p \cdot Z\}.$$

Denote the solution to the inner maximization problem by $Z^*(L, \tilde{\theta})$. Since $\mathcal{F}$ is CRS, $Z^*(L, \tilde{\theta}) = Lz^*(\tilde{\theta})$, for $z^*(\tilde{\theta}) \in \arg\max_z \{\tilde{\theta}\mathcal{F}(1, z) - p \cdot z\}$, and corresponding output is then $L\tilde{\theta}\mathcal{F}(1, z^*(\tilde{\theta}))$. As a result, for any labor input L and wage w, the firm's maximized profit is

$$\tilde{\theta}\mathcal{F}(L, Z^*(L, \tilde{\theta})) - wL - p \cdot Z^*(L, \tilde{\theta}) = L[\tilde{\theta}\mathcal{F}(1, z^*(\tilde{\theta})) - p \cdot z^*(\tilde{\theta}) - w].$$

Defining $\theta := \tilde{\theta}\mathcal{F}(1, z^*(\tilde{\theta})) - p \cdot z^*(\tilde{\theta})$, the firm's problem is $\max_{w \geq \underline{w}, L \leq G(w)} L[\theta - w]$, as before.

Finally, I argue that, in this reformulation, θ can be interpreted as labor productivity, at least if we assume that $\mathcal{F}$ is differentiable and satisfies the Inada conditions. In that case, the firm's optimal choice of each k satisfies $\tilde{\theta}\mathcal{F}_{Z_k}(1, z^*(\tilde{\theta})) = p_k$. By Euler's Theorem, under our assumption that $\mathcal{F}$ is CRS, we have $\mathcal{F}(1, z^*) = \mathcal{F}_L(1, z^*) + \sum_{k=1}^K \mathcal{F}_{Z_k}(1, z^*) z_k^*$; our definition for θ then gives

$$\begin{aligned}\theta &:= \tilde{\theta}\mathcal{F}(1, z^*(\tilde{\theta})) - p \cdot z^*(\tilde{\theta}) \\ &= \tilde{\theta}\mathcal{F}_L(1, z^*(\tilde{\theta})) + \left[\sum_{k=1}^K \tilde{\theta}\mathcal{F}_{Z_k}(1, z^*(\tilde{\theta})) z_k^*(\tilde{\theta})\right] - \left[\sum_{k=1}^K \tilde{\theta}\mathcal{F}_{Z_k}(1, z^*(\tilde{\theta})) z_k^*(\tilde{\theta})\right] = \tilde{\theta}\mathcal{F}_L(1, z^*(\tilde{\theta})).\end{aligned}$$

That is, θ can be interpreted as the marginal product of labor given that the firm has optimally chosen all other inputs.

B.3 Ambiguous effects of a mean-preserving spread, without further restrictions

In the main text, I considered a particular, global sense in which the productivity distribution can become more dispersed. Would it suffice to consider the more standard notion of a mean-preserving spread? I now show that our conclusion in Proposition 1 would *not* hold under any mean-preserving spread. To illustrate the main ideas, I use a discrete example. I first show a discrete example which is consistent with the conclusion that dispersion tends to lower the optimal minimum wage, and then I show an example in the opposite direction.

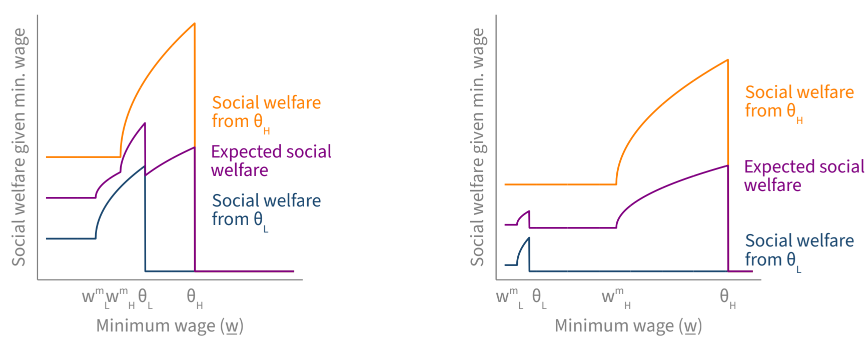
Consider a regulator who assigns probability $1/2$ to $\theta_L = \hat{\theta} - \Delta$, and

[32] I denote the production function by $\mathcal{F}$ since we have already defined F as the productivity CDF.

probability $1/2$ to $\theta_H = \hat{\theta}+\Delta$, for $\Delta \in (0, \hat{\theta})$. Appendix Figure A1(a) depicts social welfare as a function of the minimum wage $\underline{w}$, for each realized firm type (shown in navy and orange), and in expectation (shown in purple). Clearly, the optimal minimum wage must be either θ_L or θ_H. Setting $\underline{w} = \theta_L$ allows both firm types to operate, but leaves some surplus with the high-productivity firm. Setting $\underline{w} = \theta_H$ extracts the full surplus from the high-productivity firm, but causes the low-productivity firm to shut down. As $\Delta \to 0$, this trade-off becomes one-sided: setting $\underline{w} = \theta_L$ extracts almost the full surplus from the H-type monopsonist, while allowing both types to operate. Hence, for Δ sufficiently small, $\underline{w} = \theta_L$ is optimal: the optimal minimum wage falls *below* expected productivity $\hat{\theta}$. Thus, as in Proposition 1, dispersion in the minimum wage lowers the minimum wage: if the firm type was known, the average minimum wage would be $\hat{\theta}$, but the optimal minimum wage $\theta_L < \hat{\theta}$.

FIGURE A1. The effects of uncertainty about productivity on the optimal uniform minimum wage

(a) Social welfare in the two-type case, for θ_L and θ_H close together **(b) Social welfare in the two-type case, for θ_L and θ_H far apart**



Note: *Panel (a) depicts social welfare as a function of $\underline{w}$ for a regulator who assigns probability $1/2$ to the event that productivity is $\theta_L = \hat{\theta}-\Delta$, and probability $1/2$ to $\theta_H = \hat{\theta}+\Delta$, for $\hat{\theta}, \Delta > 0$, taking Δ to be small. Panel (b) replicates (b), taking Δ to approach $\hat{\theta}$.*

This argument is not fully general. To see this, take Δ to be large: in particular, say $\Delta \to \hat{\theta}$, so that $(\theta_L, \theta_H) \to (0, 2\hat{\theta})$. Appendix Figure A1(b) depicts social welfare. As before, the relevant comparison is between $\underline{w} = \theta_L$, which allows both types to operate, and $\underline{w} = \theta_H$, which extracts more surplus from the high-productivity firm but shuts down the low-productivity firm. As $\Delta \to \hat{\theta}$, the benefit of allowing the L-type firm to operate approaches zero. In consequence, for Δ sufficiently large, the regulator sets $\underline{w} = \theta_H$: the optimal minimum wage lies *above* expected productivity $\hat{\theta}$.

The example thus illustrates that the effect of uncertainty about productivity will depend on the distribution of firm types. Nonetheless, there is something unintuitive about the case in Appendix Figure A1(b), which essentially requires that the regulator "gambles" on receiving a high-productivity type. Proposition 1 describes further structure which rules out these cases.

B.4 Extension to allow for decreasing returns to scale

In the main text, I assumed the production function exhibited constant returns to scale. Now I consider the possibility of decreasing returns. Suppose revenue for a type-θ firm with L workers is $\theta\mathcal{F}(L)$. Suppose that $\mathcal{F}$ is twice continuously differentiable, strictly increasing, and strictly concave. Impose the same conditions on the productivity distribution F as before, continue to assume that the regulator's problem of setting a uniform minimum wage is concave in $\underline{w}$, and continue to assume that the optimal minimum wage $\underline{w}$ only binds firms with productivity below modal productivity q_f. I will say that $\mathcal{F}$ satisfies *decreasing sensitivity* if $\eta(L) := -\mathcal{F}'(L)/[L\mathcal{F}''(L)]$ is weakly decreasing in L. To interpret this condition, note that $\eta(L)$ is the demand elasticity under competition. To give some examples of production functions which satisfy this condition, note that under Cobb-Douglas production with a single input ($\mathcal{F}(L) = L^\alpha$, for $\alpha \in (0,1)$) we have $\eta(L) = 1/[1-\alpha]$, which is constant in L and hence trivially exhibits diminishing sensitivity. Similarly, if $\mathcal{F}(L) = \ln(L)$, then $\eta(L) = 1$. If $\mathcal{F}(L) = \ln(L+1)$, then $\eta(L) = [1+L]/L$, which again exhibits diminishing sensitivity.

Lemma A6. Say $\mathcal{F}$ satisfies diminishing sensitivity. Then Proposition 1 holds under the specification of production above.

Proof of Lemma A6. *Firm's problem.* Fix a minimum wage $\underline{w}$. The problem facing each firm type θ is concave in w, and the constraint $w \geq \underline{w}$ is linear, so either the constraint binds, or it can be ignored. In the latter case, the firm sets its monopsony wage $w^m(\theta)$, with associated employment $L^m(\theta)$. Denote by $\theta^\dagger(\underline{w})$ the type such that the minimum wage just binds: $w^m(\theta^\dagger(\underline{w})) = \underline{w}$. In the former case, the firm sets wage $\underline{w}$. Let $D(\theta, \underline{w})$ denote labor demand under wage $\underline{w}$, implicitly defined by $\theta\mathcal{F}'(D(\theta,\underline{w})) = \underline{w}$ whenever $\theta\mathcal{F}'(0) > \underline{w}$, and define $D(\theta,\underline{w}) = 0$ otherwise. Define $\theta_0(\underline{w})$ as the highest type such that types below $\theta_0(\underline{w})$ hire no workers: that is, $\theta_0(\underline{w})$ is implicitly defined by $\theta_0(\underline{w})\mathcal{F}'(0) = \underline{w}$ (and set $\theta_0 = \underline{\theta}$ if there is no such type, or 0 if f is not bounded below). The firm's employment when the constraint binds is then $L^D(\theta,\underline{w}) = \min\{G(\underline{w}), D(\theta,\underline{w})\}$. Denote by $\theta^D(\underline{w})$ the type such that $G(\underline{w}) = D(\theta^D(\underline{w}),\underline{w})$ if there is such a type; to handle edge cases, formally, define $\theta^D(\underline{w}) := \inf\{\theta \in [\theta_0(\underline{w}), \theta^\dagger(\underline{w})] : D(\theta,\underline{w}) \geq G(\underline{w})\}$, with the convention that $\theta^D(\underline{w}) = \theta^\dagger(\underline{w})$ if the set is empty.

Regulator's problem. Then the regulator's payoff under minimum wage $\underline{w}$ is

$$\begin{aligned}\mathbb{E}_{\theta\sim f}[S_\theta(\underline{w})] &= \int_{\theta=\theta_0}^{\theta^\dagger(\underline{w})} S_\theta(\underline{w}, L^D(\theta,\underline{w}))\,\mathrm{d}F(\theta) + \int_{\theta=\theta^\dagger(\underline{w})}^{\bar\theta} S_\theta(w^m(\theta), L^m(\theta))\,\mathrm{d}F(\theta)\\ &= \int_{\theta=\theta_0}^{\theta^D(\underline{w})} S_\theta(\underline{w}, D(\theta,\underline{w}))\,\mathrm{d}F(\theta) + \int_{\theta=\theta^D(\underline{w})}^{\theta^\dagger(\underline{w})} S_\theta(\underline{w}, G(\underline{w}))\,\mathrm{d}F(\theta) +\\ &\qquad \int_{\theta=\theta^\dagger(\underline{w})}^{\bar\theta} S_\theta(w^m(\theta), L^m(\theta))\,\mathrm{d}F(\theta).\end{aligned}$$

Differentiating with respect to $\underline{w}$, the boundary terms drop out and we are left with:

$$\frac{\mathrm{d}\left[\mathbb{E}_{\theta\sim f}[S_\theta(\underline{w})]\right]}{\mathrm{d}\underline{w}} = \int_{\theta=\theta_0(\underline{w})}^{\theta^D(\underline{w})} h_D(\theta,\underline{w})\,\mathrm{d}F(\theta) + \int_{\theta=\theta^D(\underline{w})}^{\theta^\dagger(\underline{w})} h_S(\theta,\underline{w})\,\mathrm{d}F(\theta),$$
$$\text{for } h_D(\theta,\underline{w}) := \frac{\mathrm{d}[S_\theta(\underline{w}, D(\theta,\underline{w}))]}{\mathrm{d}\underline{w}} \text{ and } h_S(\theta,\underline{w}) := \frac{\mathrm{d}[S_\theta(\underline{w}, G(\underline{w}))]}{\mathrm{d}\underline{w}}.$$

Separately computing each of these derivatives:

$$\text{(A5)}\qquad h_D(\theta,\underline{w}) = D(\theta,\underline{w})\left[(1-\lambda) - \eta\big(D(\theta,\underline{w})\big)\big(1 - \frac{G^{-1}(D(\theta,\underline{w}))}{\underline{w}}\big)\right]$$

$$\text{(A6)}\qquad h_S(\theta,\underline{w}) = (1-\lambda)G(\underline{w}) + \lambda g(\underline{w})\Big[\theta\mathcal{F}'(G(\underline{w})) - \underline{w}\Big].$$

Under our maintained assumption that the regulator's problem is concave, $\frac{\mathrm{d}\left[\mathbb{E}_{\theta\sim f}[S_\theta(\underline{w})]\right]}{\mathrm{d}\underline{w}}\big|_{\underline{w}=\underline{w}^*} = 0$ for the optimal minimum wage $\underline{w}^*$. Clearly, $h_S(\theta,\underline{w}) \geq 0$ for any $\underline{w}$, since for $\theta \in [\theta^D(\underline{w}, \theta^\dagger(\underline{w})]$, we must have $\theta\mathcal{F}'(G(\underline{w})) \geq \underline{w}$. Then it must be the case that $h_D(\theta,\underline{w}^*) < 0$ for some $\theta \in (\theta_0, \theta^D(\underline{w}^*)$. Define θ° as the highest value such that $h_D(\theta^\circ,\underline{w}^*) = 0$.

Single-crossing. Define $h(\theta,\underline{w}) = \mathbb{1}\{\theta \in [\theta_0, \theta^D(\underline{w}))\}h_D(\theta,\underline{w}) + \mathbb{1}\{\theta \in [\theta^D(\underline{w}), \theta^\dagger(\underline{w})]\}h_S(\theta,\underline{w})$. I now show that, for $\theta \in (\theta_0, \theta^\circ)$, we have $h(\theta,\underline{w}^*) < 0$, whereas for $\theta \in (\theta^\circ, \theta^\dagger(\underline{w}))$, we have $h(\theta,\underline{w}^*) > 0$. That is, our claim is that $h(\theta,\underline{w}^*)$ crosses 0 exactly once, and from below. Intuitively, noting that $h(\theta,\underline{w})$ is the marginal benefit to the regulator of a higher minimum wage for type θ, we are claiming that higher minimum wages hurt low-type firms, and help high-type firms.

Suppose first that $\theta \in (\theta_0, \theta^\circ)$. In that case, $h(\theta,\underline{w}^*) = h_D(\theta,\underline{w}^*)$. Since $\theta^\circ > \theta_0$, it must be the case that $D(\theta^\circ,\underline{w}^*) > 0$. As a result, since $h_D(\theta^\circ,\underline{w}^*) = 0$, it must be the case that the bracketed term in equation (A5) is zero. Next, note that the bracketed term is increasing in θ. This is because $D(\theta,\underline{w})$ is increasing in θ, and hence (by diminishing sensitivity)

$-\eta(D(\theta,\underline{w}))$ is increasing in θ, $G^{-1}(D(\theta,\underline{w})) < \underline{w}$, and $G^{-1}(D(\theta,\underline{w}))$ is increasing in θ. Since $D(\theta,\underline{w}^*) > 0$ for all $\theta \in (\theta_0, \theta^D(\underline{w}^*))$, we can therefore conclude that $h_D(\theta,\underline{w}^*) < 0$ for $\theta \in (\theta_0,\theta^\circ)$. On the other hand, consider $\theta \in (\theta^\circ, \theta^\dagger(\underline{w}^*))$. For $\theta \in (\theta^\circ, \theta^D(\underline{w}))$, the same logic as above shows that $h_D(\theta,\underline{w}^*) > 0$. For $\theta \in [\theta^\circ, \theta^\dagger(\underline{w}))$, $h(\theta,\underline{w}^*) = h_s(\theta,\underline{w}^*) > 0$. This shows the claim.

Consequences of single-crossing property for shape-preserving spread. Now consider the effects of a shape-preserving spread from f to f'. Following the logic in the discussion of Proposition 1, since $\underline{w}$ was optimal, we must have $\int_{\tilde\theta=\theta_0}^{\theta^\circ} \mathrm{MC}(\underline{w}^*,\tilde\theta)f(\tilde\theta)\mathrm{d}\tilde\theta = \int_{\tilde\theta=\theta^\circ}^{\theta^\dagger(\underline{w}^*)} \mathrm{MB}(\underline{w}^*,\tilde\theta)f(\tilde\theta)\mathrm{d}\tilde\theta$, for $\mathrm{MC}(\underline{w},\tilde\theta) := \mathbb{1}\{\tilde\theta \in (\theta_0,\theta^\circ)\}h(\underline{w},\tilde\theta)$ and $\mathrm{MB}(\underline{w},\tilde\theta) := \mathbb{1}\{\tilde\theta \in (\theta^\circ,\theta^\dagger(\underline{w}))\}h(\underline{w},\tilde\theta)$. The marginal cost terms depend on the density between θ_0 and θ°, whereas the marginal benefit terms depend on the density between θ° and $\theta^\dagger(\underline{w})$. Now consider replacing f with the shape-preserving spread f'. The marginal cost or benefit to type $\tilde\theta$ is multiplied by the density ratio $f'(\tilde\theta)/f(\tilde\theta)$. Since this density ratio is decreasing (see Lemma A3), it must be the case that the marginal costs now exceed marginal benefits, which (under our assumption that the regulator's problem is concave) implies that $\underline{w}^*$ is strictly higher than the optimal minimum wage under f'. □

B.5 Extension to allow for firm-side and worker-side subsidies

In the main text, I claimed that it was without loss to consider only subsidies paid to firms. I now show this claim formally. Suppose that the regulator can use a minimum wage $\underline{w}$, a subsidy to firms s_F, and a subsidy to workers s_W. The regulator's problem is then:

$$\max_{\underline{w},\, s_W,\, s_F} \int_{\theta\in\Theta} \Big\{ \underbrace{W\big(\hat w(\theta)+s_W, \hat L(\theta)\big) - [s_F+s_W]\hat L(\theta)}_{\text{Worker surplus net of taxes}} + \lambda \underbrace{\big[\pi_\theta(\hat w(\theta),\hat L(\theta)) + s_F \hat L(\theta)\big]}_{\text{Firm surplus}} \Big\}\, \mathrm{d}F(\theta), \text{ where}$$

$$\big(\hat w(\theta), \hat L(\theta)\big) = \underset{w\geq \underline{w},\; L\leq G(w+s_W)}{\arg\max} \Big\{\pi_\theta(w,L) + s_F L\Big\}, \text{ breaking ties in the regulator's favor.}$$

I claim that the regulator's problem depends only on the total subsidy amount $s := s_W + s_F$ and the effective minimum wage $\tilde w := \underline{w} + s_W$. To see this, note that defining $z := w + s_W$ as total worker compensation, we can rewrite the firm's problem as $\max_{z\geq\underline{w}+s_W, L\leq G(z)} L[\theta - z + s_W + s_F]$, which is then $\max_{z\geq\tilde w, L\leq G(z)} L[\theta - z + s]$, which depends only on $(\tilde w, s)$ and not on $(\underline{w}, s_W, s_F)$ directly. Similarly, fixing compensation z and employment L, we can rewrite worker surplus as $zL - c(L) - sL$, and type-θ firm profit as $L[\theta + s - z]$, which do not depend on our original instruments.

This demonstrates that any $(\underline{w}, s_W, s_F)$ with the same $(\tilde w, s)$ will induce the same allocations and social welfare as any $(\underline{w}', s_W', s_F')$. As a result, we can ignore subsidies paid to workers: if the optimal mechanism involves some non-zero subsidy to workers, then there is another optimal mechanism in which s_W is set to be equal to 0, and $(s_F, \underline{w})$ are adjusted to hold constant $(\tilde{w}, s)$.

Online Appendix C. Additional empirical details

This section contains additional empirical details on two topics: the classification of regimes in Table 1; and the empirical distribution of firm productivity and wages.

Construction of Table 1. All of the countries in the first column are described by Eurostat (2026) as having "national minimum wages". In the second column:

- Australia differentiates minimum wages by both industry and occupation (Fair Work Commission 2026).
- Costa Rica differentiates minimum wages by occupation and industry (Dentons 2025).
- Indonesia differentiates minimum wages by province (and, within province, sometimes also by city/district) (ASEAN 2026).
- Japan differentiates minimum wages by both geography (prefecture) and industry (Japanese Ministry of Health, Labour and Welfare 2024).
- Mexico differentiates minimum wages by both geography (*Zona Libre de la Frontera Norte vs.* the remainder of Mexico) and occupation (Comisión Nacional de los Salarios Mínimos 2025).
- Panama differentiates minimum wages by region and industry (Ministerio de Trabajo y Desarrollo Laboral 2025).
- The Philippines differentiates minimum wages by both region (and, within region, sometimes also by city or municipality) and sector (*e.g.*, domestic *vs.* private-sector workers) (Philippines National Wages and Productivity Commission 2025).
- Vietnam differentiates minimum wages by region (Vietnam Social Security 2025).

In the third column:

- In parts of the U.S., smaller firms are permitted to pay lower minimum wages than larger firms. For instance, very small farms are exempt from the minimum wage provisions of the Fair Labor Standards Act (FLSA) (U.S. Department of Labor, Wage and Hour Division 2023). Firms with low sales are also not covered on an enterprise basis by the FLSA (though employees at those firms may be covered individually, under some conditions) (U.S. Department of Labor, Wage and Hour Division 2023). At the state level, small firms in Arkansas, Georgia, and Illinois are exempt from state minimum wage laws (U.S. Department of Labor, Wage and Hour Division 2026). Missouri exempts firms with very low sales from state minimum wages, and Montana has an explicitly lower state minimum wage for firms with very low sales (U.S. Department of Labor, Wage and Hour Division 2026).

- The Dominican Republic differentiates minimum wages by firm size (in four firm size classes) (Presidencia de la República Dominicana 2026).
- Honduras differentiates minimum wages by firm size (Secretaría de Trabajo y Seguridad Social 2025).
- Panama differentiates minimum wages by firm size (Ministerio de Trabajo y Desarrollo Laboral 2024).
- Indonesia exempts micro and small firms from the standard schedule of minimum wages, though they are separately subject to other wage restrictions (ASEAN 2026).
- The Philippines differentiates minimum wages by firm size in some regions (Schwellnus 2026).
- Malaysia permitted small firms to pay lower wages until August 2025, when all firms were transitioned to a uniform schedule (ActivPayroll 2025).
- In parts of Pakistan (*e.g.,* Khyber Pakhtunkhwa), small enterprises are exempt from the minimum wage (Zhou 2016).

Empirical productivity and wage distribution and wages. In Section 2, I assumed that the optimal uniform minimum wage binds only firms with below-modal productivity. I stated that this assumption is empirically mild. Two back-of-the-envelope exercises support that view.

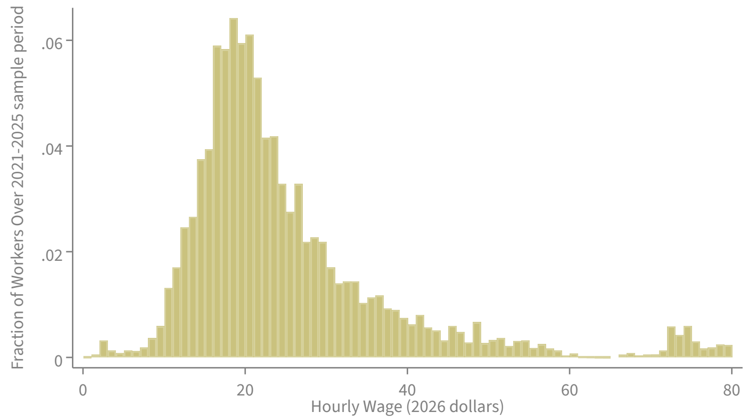
The first exercise makes use of the empirical distribution of firm productivity. Figure 2A in Bernard et al. (2003) indicates that the modal manufacturing plant has a labor productivity between 42% and 84% of the overall mean labor productivity. To be conservative, I take the lower end of this estimate. Since Bernard et al. (2003) do not report an estimate of mean labor productivity, in order to translate this into dollar terms, I combine that ratio with manufacturing value added per worker in 1992. U.S. Census Bureau (1998) reports that, in 1992, there were 16.9 million manufacturing employees, and value added by manufacturing was $1.42 trillion (in 1992 dollars) (row 1, Table 1233, p. 739). This implies that value added per worker was around $42 per hour, under an assumption that the average worker is employed 40 hours per week, 50 weeks per year. At the lower end of the range from Bernard et al. (2003), this implies that the modal labor productivity is around \$17.6; adjusting for inflation gives the result of \$42 in 2026 dollars.

Of course, there are some difficulties with this evidence. First, measured value added per worker may differ from average labor productivity. Second, the evidence is confined to manufacturing; unfortunately, I am not aware of similar estimated distributions of labor productivity among other sectors with substantial exposure to the minimum wage, including retail and hospitality.

For this reason, it is also informative to consider the empirical distribution of wages. Appendix Figure A2 presents a histogram of the U.S. hourly wage distribution (for workers paid hourly) over the period 2021–2025, computed using data from the Current Population Survey (CPS) accessed through IPUMS (Ruggles et al. 2025). The modal wage is approximately \$20. Inferring a distribution of labor productivities from the distribution of wages requires an assumption about markdowns. Berger, Herkenhoff, and Mongey (2022) estimated that the payroll-weighted average markdown in the U.S. is approximately 0.78 (p. 1184). If this markdown were constant, it would imply that the modal labor productivity is around \$25.6.

Of course, there are also difficulties inferring labor productivity from wages: under at least the assumption that markdowns are positive, the distribution of labor productivities must stochastically dominate the distribution of wages, but there is no guarantee that the modal labor productivity is equal to the modal wage, multiplied by the average markdown. Nonetheless, the fact that both approaches give values for the modal labor productivity substantially above even the most ambitious proposals for minimum wages in the U.S. suggest that our assumption is reasonable.

FIGURE A2. U.S. hourly wage distribution, 2021–2025



Note: *This figure presents a histogram of hourly wages in the U.S. for the period 2021–2025, adjusted for inflation to be in 2026 dollars. The histogram is weighted using CPS sampling weights. The sample is restricted to workers paid hourly; salaried workers are excluded.*

Online Appendix D. Detailed description of the optimal mechanism in section 5

I now describe in more detail the optimal mechanism in section 5. Our goals are to prove Lemma 4 and Proposition 3, to prove the parts of Lemma 1 referring to size-contingent minimum wages, and more generally to describe the optimal mechanism.

I begin with Lemma 4.

Proof of Lemma 4. Fix a direct mechanism $(\hat{w}(\cdot), \hat{L}(\cdot))$ that satisfies the incentive, participation, and labor supply constraints. By the argument in the text, $\hat{L}(\cdot)$ and $\hat{w}(\cdot)$ are weakly increasing on the set of types assigned positive employment. By the same argument as the one we made in the proof of Lemma A1, if two positive-employment types are assigned the same employment level, incentive compatibility requires them to receive the same wage. Thus the assigned wage is a well-defined weakly increasing function of assigned employment (for types assigned positive employment). Now consider the indirect mechanism

$$\underline{w}(L) = \begin{cases} 0, \text{ if } L = 0, \\ \inf\{\hat{w}(\theta') : \hat{L}(\theta') \geq L\}, \text{ if } L \in (0, \hat{L}(\bar{\theta})] \\ Q, \text{ if } L > \hat{L}(\bar{\theta}), \end{cases}$$

for $Q > \bar{\theta}$. Since $\hat{w}(\cdot)$ is weakly increasing (for positive employment levels), the schedule $\underline{w}(\cdot)$ is weakly increasing. To see that $\underline{w}$ implements the allocation $(\hat{w}(\cdot), \hat{L}(\cdot))$, fix a type θ. Choosing $L = 0$ gives zero profit, and $L > \hat{L}(\bar{\theta})$ gives negative profit. By the incentive constraint, a firm of type θ also prefers employment $\hat{L}(\theta)$ and wage $\underline{w}(\hat{L}(\theta))$ to any employment level in $(0, \hat{L}(\bar{\theta})]$. $\square$

I now turn to the optimal mechanism. Define *virtual productivity* $\phi(\theta) := \theta - (1-\lambda)\frac{1-F(\theta)}{f(\theta)}$. If the regulator places no weight on firm profits ($\lambda = 0$), ϕ coincides with the standard Myersonian virtual value; under our assumption that f is log-concave, ϕ is increasing. Say that labor is *rationed* for type θ if $\hat{L}(\theta) < G(\hat{w}(\theta))$.

I begin by summarizing some features of the optimal mechanism.

Lemma A7. The optimal mechanism has the following properties:

(i) The lowest-productivity types are rationed, whereas the highest-productivity types are not.

(ii) A firm of type θ is assigned zero employment if $\phi(\theta) < \underline{o}$.

(iii) For almost every type with positive first-best employment, the optimal mechanism induces strictly less than first-best employment.

Intuitively, a uniform minimum wage encourages firms to pay higher wages only through the threat of shutdown; a size-contingent schedule can instead raise wages by rationing firms that pay less (Lemma A7(i)). The regulator prefers the latter because it replaces the discrete efficiency cost of exit with the smaller distortion of reduced employment among low-productivity firms. Even so, the optimal mechanism still shuts down some firms (Lemma A7(ii)) and does not restore first-best employment (Lemma A7(iii)). A notable feature is distortion *even for the highest type*, contrary to the usual intuition. Intuitively, the minimum wage simultaneously screens firms and determines employment: since employment is bound by labor supply, achieving first-best employment $\hat{L}(\bar{\theta}) = G(\bar{\theta})$ would require $\hat{w}(\bar{\theta}) = \bar{\theta}$, which leaves the firm no profit. The optimal mechanism instead leaves the highest type some information rent, so $\hat{w}(\bar{\theta}) < \bar{\theta}$ and employment is distorted.

To prove Proposition 3 and Lemma A7, I proceed in four steps. For simplicity, I assume the mechanism is differentiable.[33]

Step 1: rewritten problem using employment and indirect utility as controls. Fix a mechanism $(\hat{L}, \hat{w})$. Denote by $\hat{V}(\theta) := \hat{L}(\theta)[\theta - \hat{w}(\theta)]$ the profit assigned to type θ by that mechanism. We will rewrite the problem using $(\hat{L}, \hat{V})$ as controls rather than $(\hat{L}, \hat{w})$. Since $\hat{V}(\theta) = \hat{L}(\theta)[\theta - \hat{w}(\theta)]$, we have

[33]Note that the allocation induced by a uniform minimum wage is itself *not* differentiable (nor even continuous). The argument presented later that the optimal minimum wage is strictly increasing in employment does not rely on our assumption of differentiability.

$\hat{w}(\theta) = \theta - \hat{V}(\theta)/\hat{L}(\theta)$ for $\hat{L}(\theta) > 0$. Moreover, incentive compatibility is equivalent to (1) $\hat{V}'(\theta) = \hat{L}(\theta)$, and hence $\hat{V}(\theta) = \hat{V}(\underline{\theta}) + \int_{\tilde{\theta}=\underline{\theta}}^{\theta} \hat{L}(\tilde{\theta})\mathrm{d}\tilde{\theta}$, and (2) $\hat{L}(\theta)$ is weakly increasing. Define the cost to workers of providing L units of labor as $c(L) := \int_{\tilde{o}=\underline{o}}^{G^{-1}(L)} \tilde{o}\ \mathrm{d}G(\tilde{o})$. Then the regulator's objective function is:

$$\begin{aligned}
\int_{\theta\in\Theta} S_\theta\big(\hat{w}(\theta), \hat{L}(\theta)\big)\ \mathrm{d}F(\theta) &= \int_{\theta\in\Theta} \Big[\theta\hat{L}(\theta) - c\big(\hat{L}(\theta)\big)\Big]\mathrm{d}F(\theta) - (1-\lambda)\int_{\theta\in\Theta} \hat{V}(\theta)\ \mathrm{d}F(\theta) \\
&= \int_{\theta\in\Theta} \Big[\theta\hat{L}(\theta) - c\big(\hat{L}(\theta)\big)\Big]\ \mathrm{d}F(\theta) \\
&\qquad - (1-\lambda)\Big[\hat{V}(\underline{\theta}) + \int_{\theta\in\Theta} \hat{L}(\theta)[1-F(\theta)]\ \mathrm{d}\theta\Big] \\
&= \int_{\theta\in\Theta} \Big[\hat{L}(\theta)\phi(\theta) - c\big(\hat{L}(\theta)\big)\Big]\ \mathrm{d}F(\theta) - (1-\lambda)\hat{V}(\underline{\theta}).
\end{aligned}$$

where the second line comes from integrating by parts. Then the regulator's problem is:

$$\max_{\hat{L},\hat{V}}\Big\{\int_{\theta\in\Theta} \Big[\hat{L}(\theta)\phi(\theta) - c\big(\hat{L}(\theta)\big)\Big]\ \mathrm{d}F(\theta) - (1-\lambda)\hat{V}(\underline{\theta})\Big\},$$

subject to (1) labor supply: $\hat{V}(\theta) \leq \hat{\Gamma}(\theta, \hat{L}(\theta))$,

(2) monotonicity: $\hat{L}(\theta)$ is increasing,

(3) participation: $\hat{V}(\underline{\theta}) \geq 0$,

(4) incentive constraint: $\hat{V}'(\theta) = \hat{L}(\theta)$,

defining $\hat{\Gamma}(\theta, \hat{L}(\theta)) := \hat{L}(\theta)[\theta - G^{-1}(\hat{L}(\theta))]$. If $\lambda < 1$, this immediately implies $\hat{V}(\underline{\theta}) = 0$; for $\lambda = 1$, there is an optimal mechanism with $\hat{V}(\underline{\theta}) = 0$. From now on, I will take the $\hat{V}(\underline{\theta}) = 0$ case. In addition, it implies $\hat{L}(\theta) = 0$ for $\phi(\theta) < \underline{o}$. Define $\theta_0 = \min\{\theta \in \Theta\colon \phi(\theta) \geq \underline{o}\}$. (Note that, under our assumption that f is log-concave, ϕ is regular.)

Step 2. Role of the labor supply constraint. There is some temptation to ignore the labor supply constraint and maximize pointwise by choice of $\hat{L}(\theta)$, which would lead to $\hat{L}(\theta) = G(\phi(\theta))$ (for $\theta \geq \theta_0$). Since ϕ is increasing (under our assumption that f is log-concave), this satisfies monotonicity. However, it would fail the labor supply constraint. To see this, note that under $\hat{L}(\theta) = G(\phi(\theta))$, we have $\hat{\Gamma}(\theta, \hat{L}(\theta)) = \hat{L}(\theta)[\theta - G^{-1}(\hat{L}(\theta))] = G(\phi(\theta))[\theta - \phi(\theta)]$, and hence $\Gamma(\bar{\theta}, \hat{L}(\bar{\theta})) = 0$. The labor constraint would then require $\hat{V}(\bar{\theta}) \leq 0$, which contradicts the incentive constraint for $\bar{\theta} > \theta_0$. We therefore conclude that we cannot ignore the constraints and focus on the relaxed problem.

Step 3: optimal control problem. Now return to the problem with constraints. Since $\hat{L}(\theta) = 0$ for $\phi(\theta) < \underline{o}$, we are left with:

$$\max_{\hat{L}\colon [\theta_0,\bar{\theta}]\to\mathbb{R},\ \hat{V}\colon [\theta_0,\bar{\theta}]\to\mathbb{R}} \int_{\theta=\theta_0}^{\bar{\theta}} \left[\hat{L}(\theta)\phi(\theta) - c\big(\hat{L}(\theta)\big)\right] \mathrm{d}F(\theta),$$

$$\begin{aligned} \text{subject to (1) labor supply: } & \hat{V}(\theta) \leq \hat{\Gamma}(\theta, \hat{L}(\theta)), \\ \text{(2) monotonicity: } & \hat{L}(\theta) \text{ is increasing}, \\ \text{(3) incentive constraint: } & \hat{V}'(\theta) = \hat{L}(\theta) \\ \text{(4) participation constraint: } & \hat{V}(\theta_0) = 0. \end{aligned}$$

Since c is convex and G is log-concave, the objective is concave in each $\hat{L}(\cdot)$, and the constraints define a convex feasible set.[34] We then form a Hamiltonian. To account for the monotonicity constraint, rather than work with $\hat{L}$ directly, I will work with $\hat{u}(\theta) := \hat{L}'(\theta)$. (As before, this is a concave problem: the objective is linear in $u(\cdot)$, and the constraints are concave.) That is, the type θ is the independent variable; the control is $\hat{u}$; the state variables are $\hat{L}$ and $\hat{V}$; and the law of motion is $\hat{L}'(\theta) = \hat{u}(\theta)$, with initial condition $\hat{L}(\theta_0) = 0$ and $\hat{V}'(\theta) = \hat{L}(\theta)$, with initial condition $\hat{V}(\theta_0) = 0$:

$$\mathcal{H}(\theta, \hat{u}, \hat{V}, \hat{L}, p, \mu, \nu) = [\hat{L}\phi(\theta) - c(\hat{L})]f(\theta) + p\hat{L} + q\hat{u} + \mu[\Gamma(\theta, L) - \hat{V}(\theta)],$$

where p is the costate on $V' = L$, q is the costate on $L' = u$, and $\mu \geq 0$ is the multiplier on the labor supply constraint, and the control constraint is that $u \geq 0$. By the maximum principle, for almost every θ, the optimal $\hat{u}(\theta)$ solves $\max_{u\geq 0} \mathcal{H}(\theta^*, \hat{u}, \hat{V}, \hat{L}, p^*, \mu^*, \nu^*)$.

Step 4: main properties. This implicitly characterizes the optimal mechanism, but does not say much explicitly. I now describe some properties of the optimal mechanism.

1. First, note that a solution exists. To see this, define $\mathcal{L} := \{\hat{L}\colon \Theta \to \mathbb{R}_+$ is weakly increasing$\}$. For any $\hat{L} \in \mathcal{L}$, denote by $V_{\hat{L}}\colon \Theta \to \mathbb{R}$ corresponding indirect utility, pinned down by the incentive and participation constraints: $\hat{V}_{\hat{L}}(\theta) = \int_{\tilde{\theta}=\underline{\theta}}^{\theta} \hat{L}(\tilde{\theta})\,\mathrm{d}\tilde{\theta}$. Since there is a unit mass of workers, we can restrict $\mathcal{L}$ to functions bounded above by 1 without loss. Endow $\mathcal{L}$ with the L^1 topology and note that the regulator's objective is continuous in $(\hat{L})$. By Helly's selection theorem and dominated convergence, $\mathcal{L}$ is compact. Moreover, the regulator's objective is continuous in $\hat{L}$, since the integrand is continuous in $\hat{L}$ and uniformly bounded. Since the labor supply constraint is continuous, the subset of $\mathcal{L}$ satisfying the labor supply constraint is closed, and

[34] This is clear for all the constraints other than the labor supply constraint. For that constraint, note that $\Gamma_{LL}(\theta, L) = [-Lg'(G^{-1}(L)) - 2[g(G^{-1}(L))]^2]/[g(G^{-1}(L))]^3$. This is weakly negative, since G is log-concave.

hence compact. Finally, the subset of $\mathcal{L}$ satisfying the labor constraint is non-empty: trivially, one such allocation is $\hat{L}(\theta) = \hat{V}(\theta) = 0$ for all θ. Hence, the regulator is maximizing a continuous objective on a compact, non-empty set; we therefore conclude that an optimal mechanism exists.

2. Second, the regulator strictly prefers some size-contingent minimum wage to no minimum wage. This follows from noting that the proof of Lemma 1 showed that the regulator strictly prefers some uniform minimum wage to no minimum wage, and the uniform minimum wage is trivially a size-contingent minimum wage.
3. Third, the optimal minimum wage is strictly increasing in employment. To see this, note that by Lemma 4, the only other possibility is that, on some segment, the optimal minimum wage is constant with employment: that is, there are two types θ and θ' with $\hat{L}(\theta) > \hat{L}(\theta') > 0$ but $\hat{w}(\theta) = \hat{w}(\theta')$. For concision, refer to this shared wage as $\hat{w}$. By Lemma 4, it must then be the case that $\hat{w}(\theta) = \hat{w}$ for every $\hat{L} \in [\theta', \theta]$. To satisfy the participation constraint of type θ', we must have $\hat{w} \leq \theta'$. On the other hand, if $\theta' > \hat{w}$, then type θ' would report type θ in order to strictly increase its payoff. We then conclude that we must have $\theta' = \hat{w}$, and hence type θ' makes no profit. Moreover, it must be the case that no other type below θ' is assigned positive employment: otherwise, it must be at a wage strictly below θ', and hence type θ' would deviate to mimic such a type to receive a strictly positive profit.

   Now fix some $\kappa > 0$ and $\varepsilon > 0$, and suppose we altered the minimum wage schedule to assign minimum wage $w^* - \varepsilon$ for firms with employment no more than $\kappa\varepsilon$. That is, under the new minimum wage schedule $(\tilde{w}_\varepsilon, \tilde{L}_\varepsilon)$, we have:

$$\tilde{w}_\varepsilon(\theta) = \begin{cases} \hat{w}(\theta), \text{ if } \theta \in [\theta^*_\varepsilon, \bar{\theta}], \\ w^* - \varepsilon, \text{ if } \theta \in [w^* - \varepsilon, w^*), \\ 0, \text{else}, \end{cases} \quad \tilde{L}_\varepsilon(\theta) = \begin{cases} \hat{L}(\theta), \text{ if } \theta \in [\theta^*_\varepsilon, \bar{\theta}], \\ \kappa\varepsilon, \text{ if } \theta \in [w^* - \varepsilon, w^*), \\ 0, \text{else}, \end{cases}$$

   where $\theta^*_\varepsilon = w^* + \frac{\kappa\varepsilon^2}{\hat{L}(\theta') - \kappa\varepsilon}$. I will show that, for ε small enough, this schedule satisfies the incentive and participation constraints and delivers a higher expected payoff for the regulator. It is simple to verify that, by construction, $(\tilde{w}_\varepsilon, \tilde{L}_\varepsilon)$ satisfies the participation and incentive constraints. I now additionally show that the regulator's payoff is higher. The increase in the regulator's payoff from schedule

$(\tilde{w}, \tilde{L})$ relative to $(\hat{w}_\varepsilon, \hat{L}_\varepsilon)$ is:

$$\Delta(\varepsilon) = \int_{\tilde{\theta}=w^*-\varepsilon}^{w^*} S_{\tilde{\theta}}(w^* - \varepsilon, \kappa\varepsilon)\, \mathrm{d}F(\tilde{\theta}) - \int_{\tilde{\theta}=w^*}^{\theta^*_\varepsilon} \left[S_{\tilde{\theta}}(w^*, \hat{L}(\theta')) - S_{\tilde{\theta}}(w^* - \varepsilon, \kappa\varepsilon)\right] \mathrm{d}F(\tilde{\theta})$$

Evaluating as $\varepsilon \to 0$, and examining only terms of order ε^2:

$$\begin{aligned}\Delta(\varepsilon) &= \kappa f(w^*)\left[(w^* - \underline{o}) - \frac{W(w^*, \hat{L}(\theta))}{\hat{L}(\theta)}\right]\varepsilon^2 + o(\varepsilon^2)\\ &= \kappa f(w^*)\left[\mathbb{E}[\tilde{o} \mid \tilde{o} \leq w^*] - \underline{o}\right]\varepsilon^2 + o(\varepsilon^2) > 0.\end{aligned}$$

This delivers an improvement that increases the regulator's payoff.

4. Fourth, the argument in step 2 that the labor supply constraint cannot be ignored also establishes that there is inefficiency at the top: $\hat{w}(\bar{\theta}) < \bar{\theta}$, and hence $\hat{L}(\bar{\theta}) \leq G(\hat{w}(\bar{\theta})) < G(\bar{\theta}) = L^c(\bar{\theta})$.
5. Fifth, a firm of type θ is assigned zero employment if $\phi(\theta) < \underline{o}$, as is immediate from our discussion of step 1.
6. Sixth, to see that almost every type with positive first-best employment is assigned strictly less than first-best employment, note that if some type θ is assigned first-best employment $G(\theta)$, by the labor supply constraint we must have $w \geq \theta$, and by the participation constraint $w \leq \theta$; this allows us to conclude that $w = \theta$. Hence, such a type earns no profit: $\hat{V}(\theta) = 0$. Since the incentive constraint is $\hat{V}'(\theta) = \hat{L}(\theta)$, any interval of such types with positive measure must be assigned zero employment.

This establishes the claims in Proposition 3 and Lemma A7. The first and second parts also establish the statements about size-contingent minimum wages in Lemma 1.

Benchmark: size-dependent subsidies. Finally, as a point of comparison, say the regulator can set subsidies that vary by firm size. This takes us into an environment very similar to Baron and Myerson (1982):

$$\max_{\underline{w}:\, \mathbb{R}_+\to\mathbb{R}_+,\, s:\, \mathbb{R}_+\to\mathbb{R}} \int_{\theta\in\Theta} \left\{ \underbrace{W\big(\hat{w}(\theta), \hat{L}(\theta)\big) - s(\hat{L}(\theta))}_{\text{Worker surplus}} + \lambda \underbrace{[\pi_\theta(\hat{w}(\theta), \hat{L}(\theta)) + s(\hat{L}(\theta))]}_{\text{Firm surplus}} \right\} \mathrm{d}F(\theta), \text{ where}$$

$$\big(\hat{w}(\theta), \hat{L}(\theta)\big) = \operatorname*{arg\,max}_{w\geq\underline{w}(L),\, L\leq G(w)} \left\{\pi_\theta(w, L) + s(L)\right\}, \text{ breaking ties in the regulator's favor.}$$

Applying the revelation principle:

$$\max_{\hat{w}:\, \Theta\to\mathbb{R},\, \hat{L}:\, \Theta\to\mathbb{R},\, s:\, \Theta\to\mathbb{R}} \int_{\theta\in\Theta} \left\{W\big(\hat{w}(\theta), \hat{L}(\theta)\big) - \hat{s}(\theta) + \lambda[\pi_\theta(\hat{w}(\theta), \hat{L}(\theta)) + \hat{s}(\theta)]\right\},$$

$$\text{subject to (incentive compatibility): } \theta \in \arg\max_{\hat{\theta}} \pi_\theta(\hat{w}(\hat{\theta}), \hat{L}(\hat{\theta})) + \hat{s}(\hat{\theta}).$$
$$\text{(participation): } \pi_\theta(\hat{w}(\theta), \hat{L}(\theta)) + \hat{s}(\theta) \geq 0.$$
$$\text{(labor supply): } \hat{L}(\theta) \leq G(\hat{w}(\theta)), \text{ for all } \theta \in \Theta.$$

First, note that the two payment arms $\hat{w}$ and $\hat{s}$ can be combined into one total payment from the firm to workers $\hat{P}$, where $\hat{P}(\theta) = \hat{w}(\theta)\hat{L}(\theta) - \hat{s}(\theta)$. This is because any $\hat{w}, \hat{s}$ producing the same $\hat{P}$ gives the same value of the objective, and has the same effect for the incentive compatibility and participation constraints; moreover, we can always ensure the labor supply constraint holds by setting $\hat{w}'(\theta) = G^{-1}(\hat{L}(\theta))$. We can then rewrite the problem as:

$$\max_{\hat{P}\colon \Theta\to\mathbb{R},\, \hat{L}\colon \Theta\to\mathbb{R}} \int_{\theta\in\Theta} \left\{\hat{P}(\theta) - C(\hat{L}(\theta)) + \lambda\left(\hat{L}(\theta)\theta - \hat{P}(\theta)\right)\right\},$$
$$\text{subject to (incentive compatibility): } \theta \in \arg\max_{\hat{\theta}} \hat{L}(\hat{\theta})\theta - \hat{P}(\hat{\theta}).$$
$$\text{(participation): } \hat{L}(\theta)\theta - \hat{P}(\theta) \geq 0,$$

where $C(L) = \int_{\tilde{o}=\underline{o}}^{G^{-1}(L)} \tilde{o}\,\mathrm{d}G(\tilde{o})$ is the cost to workers of providing labor L. Define the indirect utility of type θ under the mechanism $(\hat{P}, \hat{L})$ as $\hat{V}(\theta) := \theta\hat{L}(\theta) - \hat{P}(\theta)$. Then the incentive compatibility constraint can be replaced with the requirement that $\hat{V}'(\theta) = \hat{L}(\theta)$ almost everywhere, and that $L(\theta)$ is non-decreasing. That $\hat{V}'(\theta) = \hat{L}(\theta)$ implies $\hat{V}(\theta) = \hat{V}(\underline{\theta}) + \int_{\tilde{\theta}=\underline{\theta}}^{\theta} \hat{L}\mathrm{d}\tilde{\theta}$. Rewriting the objective:

$$\max_{\hat{V}(\underline{\theta}),\, \hat{L}\colon \Theta\to\mathbb{R}} \int_{\theta\in\Theta} \left\{\theta\hat{L}(\theta) - (1-\lambda)\left[\hat{V}(\underline{\theta}) + \int_{\tilde{\theta}=\underline{\theta}}^{\theta} \hat{L}\mathrm{d}\tilde{\theta}\right] - C(\hat{L}(\theta))\right\},$$
$$\text{subject to (participation): } \hat{V}(\underline{\theta}) \geq 0.$$
$$\text{(monotonicity): } \hat{L} \text{ is non-decreasing.}$$

We must then set $\hat{V}(\underline{\theta}) = 0$. Integrating by parts and exchanging the order of the integrals:

$$\max_{\hat{L}\colon \Theta\to\mathbb{R}} \int_{\theta\in\Theta} \left\{\phi(\theta)\hat{L}(\theta) - C(\hat{L}(\theta))\right\},$$
$$\text{subject to (monotonicity): } \hat{L} \text{ is non-decreasing.}$$

Ignoring the monotonicity constraint, we can maximize pointwise at each $\hat{L}$, which gives $\hat{L}(\theta) = G(\phi(\theta))$. This allocation is indeed monotone, since ϕ is increasing (under our assumption that f is log-concave, and hence regular). As discussed in the main text, this allocation is efficient at the top: $\hat{L}(\theta) = G(\theta) = L^C(\theta)$. Finally, note that wages are not separately identifiable from subsidies in this set-up: only the *total* payment $\hat{P}$ can be determined, and the regulator is indifferent to any pair $(\hat{w}, \hat{s})$ that induces the total payment $\hat{P}$, so long as $\hat{w}$ satisfies the labor supply constraint.