Wage dispersion, minimum wages and involuntary unemployment: a mechanism design perspective *

Simon Loertscher[†] Ellen V. Muir[‡]

November 11, 2021

Preliminary and incomplete

Abstract

We provide a theory of wage dispersion and involuntary unemployment based on optimal monopsony pricing. A wage schedule that includes a wage exceeding the market-clearing wage is optimal whenever the cost of procurement under a market-clearing wage is not convex at the optimal level of employment. Introducing a minimum wage between the lowest wage offered in equilibrium and the market-clearing wage decreases involuntary unemployment and increases employment. Whenever there is wage dispersion and involuntary unemployment at a given minimum wage, a sufficiently small increase in that wage increases employment and decreases involuntary unemployment. If there is no involuntary unemployment at a given minimum wage, a sufficiently small increase in that wage increases employment, generically. Introducing a model of quantity competition in which the aggregate quantity is procured at minimal cost, we show that setting a minimum wage above—but sufficiently close to—the lowest wage offered in equilibrium absent wage regulation still increases employment.

Keywords: monopsony power, efficiency wages, quantity competition, wage regulation, employment.
JEL-Classification: C72, D47, D82

*We thank Simon Anderson, David Byrne, Mick Coelli and seminar audiences at ALETS, Korea University, Singapore Management University, the University of Melbourne, the University of Virginia, Universitat Autònoma de Barcelona, Peking University and UCLA for useful discussions and feedback. Zhichong Lu provided excellent research assistance. Financial support through a Visiting Research Scholar grant from the Faculty of Business and Economics at the University of Melbourne, and from the Samuel and June Hordern Endowment is also gratefully acknowledged.

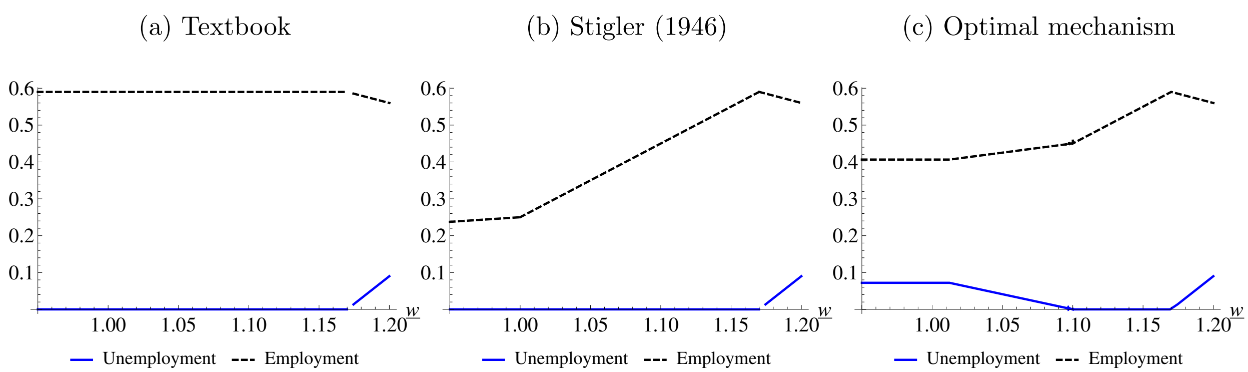
†Department of Economics, Level 4, FBE Building, 111 Barry Street, University of Melbourne, Victoria 3010, Australia. Email: simonl@unimelb.edu.au.

‡Department of Economics, Stanford University. Email: evmuir@stanford.edu.

1 Introduction

Minimum wage legislation, which has recently regained a prominent role in public policy proposals in the United States, has been around for over a century.[1] So too have debates among economists and policy makers concerning the effects of minimum wages on total employment, involuntary unemployment, and workers' pay. In models with price-taking firms and workers, minimum wages have either no effect, or induce involuntary unemployment and inefficiently low employment. However, as pointed out by Stigler (1946), if employers exert monopsony power over the labor market then an appropriately chosen minimum wage can increase workers' pay and employment without creating involuntary unemployment. The effects Stigler identified are consistent with the empirical findings of Card and Krueger (1994).

In this paper, we offer a novel perspective on the effects of minimum wages that nests the aforementioned approaches. Analyzing a model with a monopsony employer facing a continuum of workers, we first show that the monopsony's optimal procurement mechanism involves wage dispersion and induces involuntary unemployment *without* a minimum wage if the procurement cost function is not convex at the optimal level of employment. This mechanism is optimal because it minimizes the employer's labor cost subject to the workers' incentive and individual rationality constraints. We then show that any minimum wage between the lowest wage offered in equilibrium and the market-clearing wage given the optimal quantity absent wage regulation increases total employment and decreases involuntary unemployment. In fact, it is always possible to eliminate involuntary unemployment through an appropriately chosen minimum wage.



The above figure illustrates the effects of minimum wages for the textbook model that assumes price-taking behaviour; for Stigler's analysis of minimum wages under monopsony

[1] While precursors to minimum wage legislation date back to the Hammurabi Code (c. 1755–1750 BC), New Zealand became the first country to implement a minimum wage in 1894, followed by the Australian state of Victoria in 1896, and the United Kingdom in 1909 (Starr, 1981).

power; and for the case analyzed in this paper, in which the monoposony is allowed to use the optimal mechansim subject to workers' incentive compatibility and individual rationality constraints. The optimality of wage dispersion and involuntary unemployment in our analysis arises when the cost of hiring labor at the market-clearing wage is not convex in the level of employment. In such cases randomly rationing the associated excess supply of labor allows the employer to procure labor at the lowest marginal cost, subject to the monotonicity constraint implied by the incentive compatibility constraints that workers with lower costs cannot be hired with lower probability than workers with higher costs.

The optimal mechanism when the procurement cost function is not convex involves an efficiency wage—a wage exceeding the market clearing wage—and low wage at which the workers with the lowest opportunity cost of working are employed.[2] The richness and at first glance counterintuitive nature of the minimum wage effects when the employer is allowed to use an optimal mechanism raise the question of how a regulator or legislator could tell whether the problem at hand is such that increasing the minimum wage increases employment and decreases involuntary unemployment or achieves the contrary effects. As we show, the answer relates to whether or not there is wage dispersion before the minimum wage is increased marginally: If there is involuntary unemployment and wage dispersion, then a sufficiently small increase of the minimum wage will increase employment and decrease involuntary unemployment. Similarly, if there is no wage dispersion and no involuntary unemployment, then a sufficiently small increase in the minimum wage will, generically, increase employment.[3]

Proposing a model of quantity competition in which the aggregate quantity is procured at minimal cost, we also show that total employment and involuntary unemployment can move in the same direction and that there is no intrinsic relationship between the intensity of competition and the level of involuntary unemployment. Indeed, perfect competition is consistent with a positive level of involuntary unemployment. With horizontally differentiated workers and jobs, optimal procurement may involve deliberate and inefficient mismatches of workers and jobs, in addition to involuntary unemployment.

Our model also sheds new light on the effects of migration from one region, country or sector to another on labor markets. To fix ideas, consider geographical migration and

[2] The fact that the mechanism involving an efficiency wage and involuntary unemployment resonates with the dictum often attributed to Henry Ford that the Five-Dollar Day was "the best cost-cutting measure ever undertaken." Contrary to perceived wisdom, a wage of five dollars per day was not uniformly applied across all workers from the time of its introduction in 1914. See, for example, Sward (1948) who notes that according to the company's financial statement 30 percent of the overall workforce were paid less than that in 1916.

[3] The non-generic case arises when the minimum wage is equal to the wage that would prevail under price-taking behaviour.

assume that all local labor markets exhibit a convex cost of procurement. This implies that without migration, cost-minimizing procurement in each market involves a market-clearing wage and no involuntary unemployment. Further assume that migration involves a fixed cost for every worker. If demand in one local market increases, it will eventually become optimal for workers from other regions to bear that fixed cost and move to the market in which demand is stronger. This means that, prior to workers moving, the cost of procurement becomes non-convex in the destination market. Hence, with the prospect of migration, an efficiency wage and involuntary unemployment may become optimal under the mechanism that minimizes the procurement cost of labor. While this may resonate with perceived wisdom, the reasons for these effects are subtle and quite different from traditional explanations since no frictions— such as wage rigidity, unions, resistance of local workers to wage decreases—are involved. It is the prospect of attracting workers from other regions that makes the cost of procurement non-convex and may in turn make it optimal for the monopsony to use an efficiency wage and induce involuntary unemployment.

Our paper is closely related to three strands of literature: efficiency wage theory, monopoly and monopsony models under price regulation, and mechanism design problems that involve ironing. That involuntary unemployment is beneficial for businesses and detrimental for workers is a popular idea whose origins date back at least to Friedrich Engels' and Karl Marx' notion of a *reserve army of labor*.[4] More recently, it appears in the guise of the efficiency-wage theory of involuntary unemployment. According to this theory firms deliberately offer wages that exceed their market-clearing level so that the resulting excess supply of labor (and corresponding level of involuntary unemployment) can be used to discipline their workforce. For example, firms may offer efficiency wages to increase workers' effort or reduce churn. The collection of essays in Akerlof and Yellen (1986) provides an overview of the early literature that formalized these ideas, while Krueger and Summers (1988) provide empirical evidence on industry wage structure. Notwithstanding their popular appeal, one major drawback of shirking and labor market turnover models of efficiency wages is that they rest on implicit or explicit restrictions on the contracting space. As Yellen (1984, p.202) put it: "All these models suffer from a similar theoretical difficulty-that employment contracts more ingenious than the simple wage schemes considered, can reduce or eliminate involuntary unemployment." Our paper contributes to this literature by developing a model in which an efficiency wage that is optimal, subject only to individual rationality and incentive compatibility constraints, induces involuntary unemployment. Because the mechanism design approach we use is free of institutional assumptions and does not restrict the contracting space, in our setting efficiency wages and involuntary unemployment arise from the

[4]See Engels (1845) and Marx (1867).

primitives of the problem.

Stigler (1946) observed that equilibrium employment can be increased with a minimum wage in the presence of monopsony power. The basic logic extends to imperfectly competitive markets, as shown, for example, by Bhaskar et al. (2002). As mentioned, our paper shares the feature that minimum wages can increase employment. However, while these models can explain *underemployment* due to market power on the demand side, they cannot say anything about effects on involuntary unemployment because all unemployment is voluntary in models with market-clearing wages. By allowing the monopsony to use an optimal procurement mechanism, we obtain wage dispersion and involuntary unemployment absent wage regulation in equilibrium, thereby combining insights from Stigler's analysis and the mechanism design approach pioneered by Roger Myerson. From a methodological perspective, our paper is thus most closely related to the literature on monopoly pricing and mechanism design that fail to satisfy the regularity condition of Myerson (1981) and involve ironing. Of course, the idea that monopolies may benefit from bunching when faced with non-concave optimization problems is not novel and dates back to Hotelling (1931), with subsequent contributions by Mussa and Rosen (1978), Myerson (1981), Bulow and Roberts (1989), and a recent upsurge of interest driven by the applications considered in Condorelli (2012), Dworczak et al. (2021), Loertscher and Muir (2021a) and Akbarpour et al. (2020). That said, to the best of our knowledge, the connection between irregular mechanism design problems, involuntary unemployment and minimum wage effects that are made in this paper have never been touched upon before.[5]

Our model of quantity competition in Section 5 is related to the literature on Cournot competition (Cournot, 1838), while our discussion of optimal mechanisms in the Hotelling model builds on Balestrieri et al. (2021) and Loertscher and Muir (2021b). It also relates to mechanism design problems involving endogenous worst-off types (see, for example, Loertscher and Wasser, 2019).

The remainder of this paper is structured as follows. Section 2 introduces the baseline procurement setup. In Section 3, we relate the monopsony's optimal procurement mechanism to efficiency wages and involuntary unemployment. In Section 4, we analyze the effects of minimum wages. Section 5 extends the model to quantity competition. Section 6 discusses the effects of policies aiming at reducing the inefficiency arising from random allocations and provides extensions to horizontally and vertically differentiated jobs and costly migration. Section 7 concludes the paper.

[5]The analysis of the model with heterogeneous tasks and endogenous multi-tasking in Section 6.2 is related to the seminal multi-tasking model of Holmström and Milgrom (1991) and the mechanism design analyses of Condorelli (2012) and Loertscher and Muir (2021a), which involve the allocation of heterogeneous goods.

2 Setup

We consider the procurement problem of a monopsony whose willingness to pay for $Q \in [0,1]$ units of labor is $V(Q)$. For simplicity the function V is assumed to be a strictly decreasing and continuously differentiable function on $Q \in [0,1]$.[6,7] Let W denote the inverse supply function faced by the the monopsony so that $W(Q)$ is then the market-clearing wage for procuring the quantity $Q \in [0,1]$. We then let

$$C(Q) := W(Q)Q$$

denote the cost of procuring $Q \in [0,1]$ units at the market-clearing wage. We assume that the supply side consists of a continuum of workers of mass 1 each of whom supplies one unit of labor inelastically, with $W(Q)$ representing the opportunity cost of working for the worker with Q-th lowest opportunity cost. We further assume that the opportunity cost of working is the private information of each worker and that W is a strictly increasing (so that the monopsony faces an upwards sloping labor supply schedule) and continuously differentiable function.[8] This in turn implies that the function C is strictly increasing and continuously differentiable. The input (or labor) supply function is denoted S so that $S(w) = W^{-1}(w)$ holds for all wages $w \in [W(0), W(1)]$.

We assume that $V(0) > W(0)$ and $V(1) < W(1)$, so that under the optimal procurement mechanism there is a strictly positive mass of workers employed by the monopsony and a strictly positive mass of workers that are not employed. The efficient employment level Q^p that would emerge under price-taking behaviour satisfies $V(Q^p) = W(Q^p)$. Because V is decreasing while W is increasing and $V(0) > W(0)$ and $V(1) < W(1)$, Q^p exists and is unique.

We also allow the monopsony to offer multiple wages $(w_1, \ldots, w_n)$ and to specify the amount that it is willing to procure at each wage w_i. We refer to $(w_1, \ldots, w_n)$ as the *wage schedule*. We say that the monopsony uses an *efficiency wage* to procure a quantity Q if its wage schedule involves a wage w_i that is larger than the market-clearing wage (that is, $w_i > W(Q)$) and if it procures a positive quantity at w_i.

[6] As will become clear, by assuming that V is strictly decreasing we avoid having to deal with the possibility that the optimal quantity procured is not unique. Introducing this assumption ensures that whenever a procurement mechanism involving an efficiency wage is optimal, it is the uniquely optimal mechanism.

[7] If the firm uses the input to generate revenue $R(Q)$, where R is concave and increasing for Q sufficiently small, then the firm's willingness to pay for the Q-th unit of input is given by $V(Q) = R'(Q)$. The firm could be a monopoly on the output market with a technology that transforms one unit of input into one unit of output or a price-taking firm, in which case the concavity is derived from a production function that exhibits decreasing marginal products in the input.

[8] The assumption that W is continuously differentiable is made purely for expositional convenience.

In short, the setting we consider makes two important departures from an otherwise completely standard monopsony pricing problem. First, we do not restrict the monopsony to setting the market-clearing wage $w = W(Q)$ when it procures the quantity Q. Second, we do not assume that the cost of procurement function C is convex. As we shall see, these assumptions go hand-in-hand: It is without loss of generality to restrict attention to market-clearing wages when the cost function C is convex. However, when the cost function C fails to be convex, the monopsony may strictly benefit from offering an efficiency wage and inducing involuntary unemployment.

Given the important role that non-convex cost functions play in our analysis, this may beg the question of why such functions might arise in practice. In Section 6.3 we show that non-convexities arise naturally when workers face a fixed cost of moving, changing occupation or participating in the labor market. The assumption that a monopsony faces a convex procurement cost is analogous to the assumption that a monopoly faces a concave revenue function. As discussed in Loertscher and Muir (2021a), while this assumption is widely maintained in both theoretical and empirical work in Industrial Organization it is frequently rejected when tested empirically. However, more fundamentally, there is simply no theoretical reason for why convex cost functions should arise in the first place.

3 Optimal procurement mechanism

We begin by introducing a function that will play a central role in our analysis: the *convexification* $\underline{C}$ of the cost function C, which is the largest convex function that is weakly less than C at every point $Q \in [0, 1]$. If C is a convex function then we have $\underline{C} = C$. If C fails to be convex then its convexification is characterized by a countable set $\mathcal{M}$ and a set of disjoint open intervals $\{(Q_1(m), Q_2(m))\}_{m \in \mathcal{M}}$ such that

$$\underline{C}(Q) = \begin{cases} C(Q_1(m)) + \frac{(Q-Q_1(m))(C(Q_2(m))-C(Q_1(m)))}{Q_2(m)-Q_1(m)}, & \exists\, m \in \mathcal{M} \text{ s.t. } Q \in (Q_1(m), Q_2(m)) \\ C(Q), & Q \notin \bigcup_{m \in \mathcal{M}} (Q_1(m), Q_2(m)) \end{cases}$$

and, for each $m \in \mathcal{M}$, $Q_1(m)$ and $Q_2(m)$ satisfy the first-order condition

$$C'(Q_1(m)) = \frac{C(Q_2(m)) - C(Q_1(m))}{Q_2(m) - Q_1(m)} = C'(Q_2(m)). \tag{1}$$

Observe that on each interval $(Q_1(m), Q_2(m))$, $\underline{C}$ is a linear function given by a convex combination of $C(Q_1(m))$ and $C(Q_2(m))$ that exhibits constant marginal cost. In particular, if $Q \in (Q_1(m), Q_2(m))$ for some $m \in \mathcal{M}$, then $\underline{C}(Q)$ can equivalently be written

$$\underline{C}(Q) = (1 - \alpha_m(Q))C(Q_1(m)) + \alpha_m(Q)C(Q_2(m)),$$

where $\alpha_m(Q) = \frac{Q - Q_1(m)}{Q_2(m) - Q_1(m)}$. Since C is increasing, so too is $\underline{C}$. Without loss of generality, if $|\mathcal{M}| > 1$, we can index the intervals $(Q_1(m), Q_2(m))$ in increasing order so that, for all $m \geq 2$, we have $Q_2(m-1) < Q_1(m)$. Note that since W is strictly increasing, we must have $Q_1(1) > 0$.[9]

For the monopsony problem with a fixed marginal benefit function V, the focus on the case where the function C fails to be convex on a single interval is without loss of generality.[10] For the remainder of this section, we assume that C exhibits only one interval of non-convexity, that is $|\mathcal{M}| = 1$, and simply write Q_1 and Q_2 in lieu of $Q_1(1)$ and $Q_2(1)$. As an illustration, consider the piecewise linear input supply function

$$W(Q) = \begin{cases} 4Q, & Q \in [0, 1/4) \\ Q/2 + 7/8, & Q \in [1/4, 1] \end{cases}, \tag{2}$$

which gives rise to the non-convex cost function[11]

$$C(Q) = \begin{cases} 4Q^2, & Q \in [0, 1/4) \\ Q^2/2 + 7Q/8, & Q \in [1/4, 1] \end{cases}. \tag{4}$$

Straightforward computations show that

$$Q_1 = \frac{4 + \sqrt{2}}{32} \approx 0.169 \quad \text{and} \quad Q_2 = \frac{1 + 2\sqrt{2}}{8} \approx 0.478.$$

Figure 1 displays C in panel (a), C and $\underline{C}$ in panel (b), and C' and $\underline{C}'$ in panel (c).

Intuitively, the convexification $\underline{C}$ is constructed from C by replacing C' with the average

[9]To see this, assume to the contrary that $Q_1(1) = 0$. Because $C'(0) = W(0)$ and $C(Q_2(1)) = Q_2(1)W(Q_2(1))$, the first equality in the first-order condition (1) becomes $W(0) = W(Q_2(1))$, which contradicts the assumption that W is strictly increasing.

[10]The reason for this sufficiency is is that the function V is decreasing, which implies a unique point of intersection of the functions $\underline{C}'$ and V.

[11]The example in (3) is a special case of the piecewise linear specification in which, for $a > b > 0$ and $\underline{q} \in (0, 1)$, $W(Q)$ is given by

$$W(Q) = \begin{cases} aQ, & Q \in [0, \underline{q}) \\ bQ + (a-b)\underline{q}, & Q \in [\underline{q}, 1] \end{cases}, \tag{3}$$

which gives rise to $C(Q) = aQ^2$ for $Q \in [0, \underline{q})$ and $C(Q) = bQ^2 + Q(a-b)\underline{q}$ for $Q \in [\underline{q}, 1]$. Example (3) arises by setting $a = 4$, $b = 1/2$ and $\underline{q} = 1/4$.

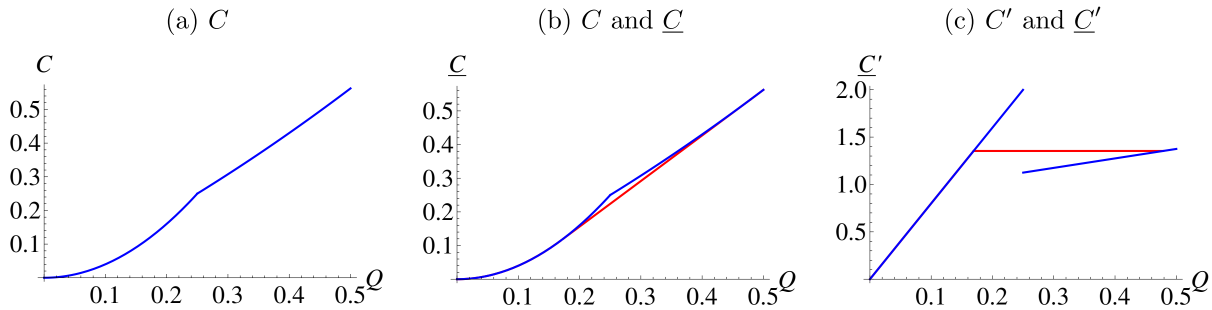


Figure 1: Illustration of C (blue), $\underline{C}$ (red), C' (blue) and $\underline{C}'$ (red) for the piecewise linear specification.

slope of C over the interval (Q_1, Q_2) since

$$\frac{\int_{Q_1}^{Q_2} C'(Q)dQ}{Q_2 - Q_1} = \frac{C(Q_2) - C(Q_1)}{Q_2 - Q_1}.$$

Therefore, $\underline{C}$ is the smallest function that can be constructed from C in this way subject to the monotonicity constraint of not giving more weight to $C'(\hat{Q})$ than to $C'(\tilde{Q})$ for $\hat{Q} > \tilde{Q}$ in the averaging process. This is achieved by assigning all $Q \in (Q_1, Q_2)$ equal weight.

The relevance of the convexification $\underline{C}$ becomes clear in the following lemma and proposition, after which we also discuss the origin and interpretation of the monotonicity constraint.

Lemma 1. *The monopsony can procure the quantity Q at cost $\underline{C}(Q)$. Moreover, if $\underline{C}(Q) < C(Q)$, this is achieved by using an efficiency wage.*

Because the proof of this lemma is instructive, we provide it in the main body of the paper. The first part of the lemma is obviously true if $Q \notin (Q_1, Q_2)$ for then $\underline{C}(Q) = C(Q)$, in which case Q can be procured at cost $C(Q)$ by setting the market-clearing wage $w = W(Q)$. So assume $Q \in (Q_1, Q_2)$, in which case $\underline{C}(Q) < C(Q)$. Let the monopsony set the wage schedule (w_1, w_2) with $w_2 = W(Q_2)$ at which it is willing to procure $Q - Q_1$ units, and w_1, at which it procures Q_1 units. Incentive compatibility for the marginal worker with opportunity cost $W(Q_1)$ who is indifferent between working for sure at the low wage w_1 and taking the gamble of working at the higher wage w_2 with probability α (and being unemployed with probability $1 - \alpha$) is

$$w_1 - W(Q_1) = \alpha(\underbrace{W(Q_2)}_{=w_2} - W(Q_1)),$$

yielding

$$w_1 = (1 - \alpha)W(Q_1) + \alpha W(Q_2),$$

which increases in Q because α increases in Q and $W(Q_1) < W(Q_2)$.

Notice also that if the worker with opportunity cost $W(Q_1)$ is indifferent, all workers with lower opportunity costs strictly prefer to work for sure at w_1 to taking the gamble.[12]

We are left to show that this results in the cost $\underline{C}(Q)$. To see that this is the case, notice first that the total wage payment to the low-wage workers is $Q_1 w_1 = (1-\alpha)C(Q_1) + \alpha W(Q_2)Q_1$, while the total wage bill to the high-wage workers is $(Q - Q_1)w_2 = W(Q_2)(Q - Q_1)$. Because $\alpha Q_1 + Q - Q_1 = \alpha Q_2$, adding up yields

$$Q_1 w_1 + (Q - Q_1)w_2 = (1-\alpha)C(Q_1) + \alpha C(Q_2) = \underline{C}(Q)$$

as required.

Lemma 1 shows that for $Q \in (Q_1, Q_2)$ the monopsony can do better using an efficiency wage rather than a market-clearing wage. Moreover, in such cases we have also constructed an explicit mechanism, parameterized by the quantities (Q_1, Q, Q_2), that achieves a procurement cost of $\underline{C}(Q)$. We will refer to this class of mechanisms as *two-price mechanisms*. The next proposition shows that an efficiency wage is optimal in this case in the sense that $\underline{C}(Q)$ is the minimum cost for procuring Q in an incentive compatible and individually rational mechanism.

Proposition 1. *In the incentive compatible and individually rational mechanism that minimizes the cost of procuring the quantity Q the cost of procurement is $\underline{C}(Q)$.*

Proposition 1 can be established using the mechanism design approach of Myerson (1981). Together with Lemma 1, it implies that for $Q \in (Q_1, Q_2)$ efficiency wages are without loss of generality.

The monotonicity constraint that $C'(\hat{Q})$ obtains no more weight in the averaging process than $C'(\tilde{Q})$ for $\hat{Q} > \tilde{Q}$ corresponds to the incentive compatibility constraint in the mechanism design problem. This constraint implies that a worker with a lower opportunity cost of working, say $\tilde{w} = W(\tilde{Q})$, cannot be employed with lower probability than a worker with the

[12] This two-wage scheme is robust to the introduction of risk aversion on behalf of workers in the following sense. Suppose all workers have the same initial wealth level, which without loss of generality can be normalized to zero, and the same, strictly concave utility function $U(.)$. So a worker with opportunity cost $W(Q)$ working at wage $w \geq W(Q)$ has a utility of $U(w - W(Q))$ while an unemployed worker has a utility of $U(0)$. To replicate the equilibrium above, the participation constraint for the marginal worker still requires $w_2 = W(Q_2)$ as in the risk neutral case. In contrast, the wage $\hat{w}_1$ that makes workers with opportunity cost $W(Q_1)$ indifferent now satisfies $U(\hat{w}_1 - W(Q_1)) = \alpha U(W(Q_2) - W(Q_1)) + (1-\alpha)U(0)$. Since U is strictly concave, $\hat{w}_1 < w_1 = (1-\alpha)W(Q_1) + \alpha W(Q_2)$ follows. Moreover, the single-crossing condition is satisfied a fortiori because U is concave. Not surprisingly, the additional benefit of insurance offered by certain employment works in favor of the firm's scheme. However, it is not clear whether with risk averse agents, the optimal mechanism only involves two wages. Moreover, the mechanism likely varies with the curvature of the agents' utility function. For these reasons, it seems preferable to stick to the model with risk-neutral agents.

higher opportunity cost $\hat{w} = W(\hat{Q})$. Intuitively, the worker of type $\tilde{w}$ could always profitably imitate the worker of type $\hat{w}$ if $\hat{w}$ were employed with higher probability.[13] Because the firm would prefer to hire workers with high types, whose marginal cost is small, to hiring workers with low types and high marginal costs, the best it can do, subject to incentive compatibility, is to hire them with equal probability. In a nutshell, this is why the monoposony uses an efficiency wage that induces involuntary unemployment to procure $Q \in (Q_1, Q_2)$.

Let Q^* be such that

$$V(Q^*) = \underline{C}'(Q^*).$$

Observe that Q^* is unique because V is strictly decreasing by assumption and $\underline{C}'$ is weakly increasing because $\underline{C}$ is a convex function.

Proposition 2. *The monopsony optimally employs Q^* workers. It optimally uses wage dispersion and induces involuntary unemployment if and only if $Q^* \in (Q_1, Q_2)$.*

Proposition 2 can be proven by adapting arguments from Loertscher and Muir (2021a), which analyzes optimal monopoly pricing when the revenue function is not concave. When selling homogeneous goods, randomization takes the form of rationing, which because of its inefficiency offers scope for resale among consumers if the objects sold can be transferred between them. In a labor market context, resale, or subcontracting, among workers is not an issue if the jobs have to be performed on site and the employer can control who gets access to that site. Proposition 2 provides a formalization of why a firm can benefit from, using the Marx's and Engels' term, a reserve army of the unemployed. The excess supply induced by the efficiency wage allows it to randomize over workers and thereby to reduce its procurement cost.

If the monopsony optimally uses an efficiency wage, then the level of involuntary unemployment is $Q_2 - Q^*$, and the mass of workers who are employed is Q^*. The rate of involuntary unemployment, measured as fraction of unemployed over the total number of individuals willing to work, is $(Q_2 - Q^*)/Q_2$. We then denote by $w_1^* = w_1(Q^*; m)$ the equilibrium low wage, that is, the value of w_1 when $Q = Q^*$ and hence $\alpha = \alpha^*$. For consistency, even though it does not vary with $Q \in (Q_1, Q_2)$, we also write $w_2^* = W(Q_2)$.

[13] The formal proof is elementary (see e.g. Börgers, 2015). Denote by $t(w)$ the expected transfer an agent receives when, in a direct mechanism, he reports that his type is w, and by $q(w)$ the probability that he has to work upon the same report. Incentive compatibility for types w and $\hat{w}$ then implies $t(w) - q(w)w \geq t(\hat{w}) - q(\hat{w})w$ and $t(w) - q(w)\hat{w} \leq t(\hat{w}) - q(\hat{w})\hat{w}$, respectively. Subtracting the second from the first implies $q(w)(\hat{w} - w) \geq q(\hat{w})(\hat{w} - w)$, which for $\hat{w} > w$ holds if and only if $q(w) \geq q(\hat{w})$.

4 Minimum wage effects

Minimum wages are often considered a cause of involuntary unemployment. In models with price-taking firms, a minimum wage affects equilibrium outcomes only if it is set above the the market-clearing wage, in which case it induces involuntary unemployment. However, as shown by Stigler (1946), an appropriately chosen minimum wage can increase employment without causing involuntary unemployment in the presence of monopsony power. We are now going to show that when involuntary unemployment occurs in equilibrium in our model, appropriately chosen minimum wages increase employment and decrease involuntary unemployment. Moreover, an appropriately chosen minimum wage eliminates involuntary unemployment. We first discuss the effects of a minimum wage equal to the efficiency wage, which are relatively simple and provide an illustration of the mechanics at work in general. Then we turn to these general effects.

4.1 Setting a minimum wage equal to the efficiency wage

As we shall see shortly, our general analysis investigating the effects of minimum wages is quite involved and technical. To build intuition we start by discussing the special case in which a regulator sets a minimum wage equal to an efficiency wage that prevails absent regulation. Naturally, this case is covered by the formal propositions stated in Section 4.2. However, here we can more or less immediately see that such a minimum wage will increase employment. In particular, since both $C'(Q) > W(Q)$ and $V(Q^*) = C'(Q_2(m))$ hold, we know that $V(Q^*) = C'(Q_2(m)) > W(Q_2(m))$. Consequently, a minimum wage equal to the efficiency wage $W(Q_2(m))$ will increase employment in equilibrium since the firm will demand the quantity $Q^*(\underline{w})$ such that $V'(Q^*(\underline{w})) = \underline{w}$ with $\underline{w} = W(Q_2(m))$.

Whether or not this eliminates involuntary unemployment depends on whether Q^p is larger or smaller than $Q_2(m)$. If $Q^p \geq Q_2(m)$, then we have $V(Q_2(m)) \geq W(Q_2(m))$. This means that a firm facing a minimum wage of $\underline{w} = W(Q_2(m))$ will optimally hire $Q_2(m)$ workers at $\underline{w}$, which is the market-clearing wage for the quantity $Q_2(m)$. It will not hire any additional workers because we have $\underline{C}'(Q) > V(Q_2(m))$ for $Q > Q_2(m)$. In contrast, if $Q^p < Q_2(m)$, then $V(Q_2(m)) < W(Q_2(m))$, and a minimum wage equal to $W(Q_2(m))$ will induce involuntary unemployment insofar as for, say, the minimum wage $\underline{w} = W(Q^p)$, there would be no involuntary unemployment.

Figure 2 illustrates the effects of a minimum wage $\underline{w} = W(Q_2(m))$ for the case when $Q^p < Q_2(m)$ and when $Q^p > Q_2(m)$. Without minimum wage regulation, there is involuntary unemployment. If a minimum wage $\underline{w} = W(Q_2)$ is imposed, the monopsony hires more workers. The case $Q^p < Q_2(m)$ is displayed in panel (a), in which case the firm hires less than $Q_2(m)$ workers because $V(Q_2) < W(Q_2(m))$. Consequently, while involuntary unemployment is reduced, it is not eliminated. In contrast, if $Q^p \geq Q_2(m)$, which is displayed in panel (b), $V(Q_2(m)) \geq W(Q_2(m))$ holds. Hence, given $\underline{w} = W(Q_2(m))$, the firm will optimally employ $Q_2(m)$ workers, and involuntary unemployment is eliminated.

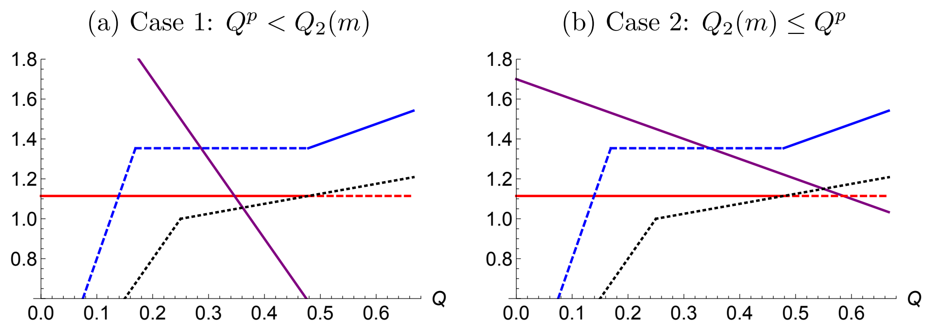


Figure 2: Illustration of the effects associate with imposing a minimum wage of $W(Q_2)$ (red). The quantity Q^p is given the intersection of $W(Q)$ (black) with $V(Q)$ (purple). On the left $Q^p < Q_2$ and $\underline{w} = W(Q_2)$ exceeds the wage $W(Q^p)$ that would prevail under price-taking behaviour. On the right, $Q^p > Q_2$, and $\underline{w} = W(Q_2)$ eliminates involuntary unemployment.

4.2 General minimum wage effects

We now systematically analyze the effects of minimum wages on employment, wage dispersion, and involuntary unemployment. We will show that, roughly speaking, these effects vary depending on whether the minimum wage lies within one of three regions, as illustrated in Figure 3, where we drop the index m for notational simplicity. In the first region, which is characterized by $\underline{w} \in (w_1(Q^*), W(\hat{Q}))$ and plotted in red in Figure 3, the minimum wage is accompanied by wage dispersion and involuntary unemployment. In this region, increasing the minimum wage will decrease involuntary unemployment and wage dispersion and increase employment. The second region, plotted in blue and characterized by $\underline{w} \in [W(\hat{Q}), W(Q^p))$, has the pure effects identified by Stigler (1946) that increasing the minimum wage increases employment without causing involuntary unemployment. In this region, and beyond, there is no wage dispersion. The last region, plotted in black and characterized by $\underline{w} \geq W(Q^p)$, corresponds to the textbook model with price-taking behaviour in which increasing the minimum wage increases involuntary unemployment and decreases employment. Figure 3 provides a rough or schematic summary insofar as there may be additional regions inside the interval $(W(Q^*), W(Q^p)]$ with and without wage dispersion, and $W(\hat{Q})$ need not be strictly less than $W(Q^p)$ if $Q^p < Q_2$. However, everything else is precise. In particular, at $\underline{w} = W(Q^*)$ there will still be wage dispersion and involuntary unemployment, and, of course, $w_1(Q^*) < W(Q^*)$

is the case for any $Q^* \in (Q_1, Q_2)$. If $Q_2 \leq Q^p$, then $\hat{Q} < Q_2$ holds, in which case there is no wage dispersion for any $\underline{w} \in (W(\hat{Q}), W(Q_2)]$.

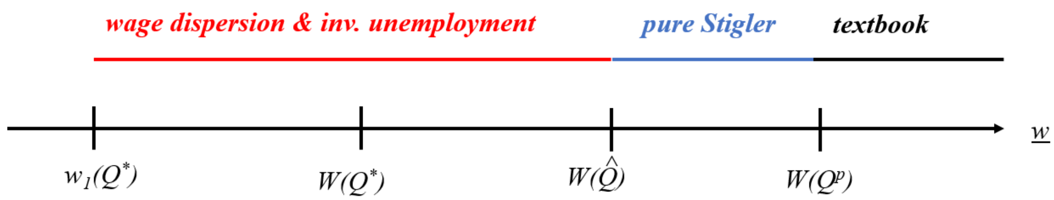


Figure 3: Three regions of minimum wage effects (schematic).

In order to state and prove our results, we first need to determine the minimum cost $\underline{C}(Q, \underline{w})$ of procuring the quantity Q—and the associated optimal procurement mechanism—for a given minimum wage $\underline{w}$. Recall that S denotes the labor supply function. For any $Q \leq S(\underline{w})$, the minimum cost of procuring the quantity Q is $\underline{w}Q$ because this cost cannot be reduced by randomizing over wages that are all at least as high as $\underline{w}$. Likewise, when $S(\underline{w}) \notin (Q_1(m), Q_2(m))$ for any $m \in \mathcal{M}$ (or, equivalently, when $\underline{w} \notin (W(Q_1(m)), W(Q_2(m)))$ for any $m \in \mathcal{M}$) the minimum cost of procuring the quantity $Q > S(\underline{w})$ is simply $\underline{C}(Q)$.[14] Thus, if $\underline{w} \notin (W(Q_1(m)), W(Q_2(m)))$ for any $m \in \mathcal{M}$, the minimum cost of procuring the quantity Q is $\underline{C}(Q, \underline{w}) = \underline{w}Q$ if $Q \in [0, S(\underline{w})]$ and $C(Q, \underline{w}) = \underline{C}(Q)$ if $Q > S(\underline{w})$.

Thing become more complicated when $S(\underline{w}) \in (Q_1(m), Q_2(m))$ for some $m \in \mathcal{M}$ (or, equivalently, when $\underline{w} \in (W(Q_1(m)), W(Q_2(m))))$. Given such an $m \in \mathcal{M}$, let $\alpha_m(Q) := \frac{Q - Q_1(m)}{Q_2(m) - Q_1(m)}$ and let

$$w_1(Q; m) := (1 - \alpha_m(Q))W(Q_1(m)) + \alpha_m(Q)W(Q_2(m)), \tag{5}$$

denote the lower wage that is paid in equilibrium absent wage regulation. Clearly, the the minimum wage does not constrain the optimal wages the monopsony uses absent wage regulation and the minimum cost of procuring Q given $\underline{w}$ is still $\underline{C}(Q)$ if

$$\underline{w} \leq w_1(Q; m).$$

Since $Q_1(m) < S(\underline{w}) < Q_2(m)$ and $w_1(Q; m)$ is an increasing and continuous function in Q on $[Q_1(m), Q_2(m)]$ satisfying $w_1(Q_i(m); m) = W(Q_i(m))$, for any $\underline{w} \in (W(Q_1(m)), W(Q_2(m)))$, $w_1^{-1}(\underline{w}; m)$ is well-defined. Consequently, we have $\underline{C}(Q, \underline{w}) = \underline{C}(Q)$ for any $Q \geq w_1^{-1}(\underline{w}; m)$ and $Q \in (S(\underline{w}), w_1^{-1}(\underline{w}; m))$ is the only case that requires further analysis.

[14]If $Q \notin (Q_1(m), Q_2(m))$ for any $m \in \mathcal{M}$, then the minimum cost of procuring Q absent a minimum wage is $\underline{C}(Q) = W(Q)Q$ with $W(Q) > \underline{w}$. Alternatively, if $Q \in (Q_1(m), Q_2(m))$ for some $m \in \mathcal{M}$, then the lower wage absent wage regulation is no smaller than $W(Q_1(m)) \geq \underline{w}$. In either case the minimum wage does not affect the minimum cost of procurement.

We first state in general terms what the solution is. Then we discuss its key technical properties and their economic implications. For $\underline{w} \in (W(Q_1(m)), W(Q_2(m)))$, consider the cost function given by

$$\underline{C}(Q, \underline{w}) = \begin{cases} \underline{w}Q, & Q \in [0, S(\underline{w})] \\ \mathcal{L}^*(Q, \underline{w}), & Q \in \big(S(\underline{w}), w_1^{-1}(\underline{w}; m)\big) \\ \underline{C}(Q), & Q \geq w_1^{-1}(\underline{w}; m), \end{cases} \tag{6}$$

where $\mathcal{L}^*(Q, \underline{w})$ is the value (which is written in terms of a Lagrangian in the proof Lemma 2) of the cost-minimization problem

$$\mathcal{L}^*(Q, \underline{w}) := \min_{q_1 \in [0,Q],\, q_2 \geq Q} \{(1-\alpha)C(q_1) + \alpha C(q_2)\} \tag{7}$$
$$\text{s.t.} \quad (1-\alpha)\, W(q_1) + \alpha W(q_2) \geq \underline{w}, \quad \alpha = \frac{Q - q_1}{q_2 - q_1}.$$

Whenever $Q \in \big(S(\underline{w}), w_1^{-1}(\underline{w}; m)\big)$ the function $\underline{C}(Q, \underline{w})$ computes the cost-minimizing two-wage procurement mechanism, subject to the constraint that the lower wage be no less than the minimum wage. As we have already discussed, under the minimum wage $\underline{w}$, the minimal cost of procuring the quantity $Q \in [0, S(\underline{w})]$ is $\underline{w}Q$ and the minimum cost of procuring the quantity $Q \geq w_1^{-1}(\underline{w}; m)$ is $\underline{C}(Q)$. Putting all of this together, provided it is without loss of generality to restrict attention to two-price procurement mechanisms when $Q \in \big(S(\underline{w}), w_1^{-1}(\underline{w}; m)\big)$, by construction $\underline{C}(Q, \underline{w})$ specifies the minimal cost of procuring the quantity Q under the minimum wage $\underline{w}$. The following lemma shows that this is indeed the case and establishes a number of useful properties of this function, as well as the marginal cost function $\underline{C}'(Q, \underline{w}) := \frac{\partial \underline{C}(Q,\underline{w})}{\partial Q}$.

Lemma 2. *The minimal cost of procuring the quantity Q under the minimum wage $\underline{w} \in (W(Q_1(m)), W(Q_2(m)))$ is given by $\underline{C}(Q, \underline{w})$ in* (6)*. Moreover, $\underline{C}(Q, \underline{w})$ is continuous and increasing in both Q and $\underline{w}$ and convex in Q. It also satisfies* $\lim_{Q \downarrow S(\underline{w})} \underline{C}(Q, \underline{w}) = \underline{w}S(\underline{w})$, $\lim_{Q \uparrow w_1(\underline{w};m)} \underline{C}(Q, \underline{w}) = \underline{C}(Q)$ *and, for $Q \in \big(S(\underline{w}), w_1^{-1}(\underline{w}; m)\big)$,*

$$\frac{\partial \underline{C}'(Q, \underline{w})}{\partial Q} > 0 > \frac{\partial \underline{C}'(Q, \underline{w})}{\partial \underline{w}}.$$

We have now formally shown that for a given minimum wage $\underline{w}$ and quantity $Q \in \big(S(\underline{w}), w_1^{-1}(\underline{w}; m)\big)$, the optimal procurement mechanism is a two-price mechanism where workers are hired with certainty at the minimum wage and rationed at a higher wage. From a theoretical perspective, the fact that $\frac{\partial \underline{C}'(Q,\underline{w})}{\partial Q} > 0$ holds for this region of the parameter space is noteworthy. As illustrated in Figure 4, this implies that over the interval $Q \in \big(S(\underline{w}), w_1^{-1}(\underline{w}; m)\big)$, the minimum cost of procurement is strictly convex and the "ironed" marginal cost function is strictly increasing. In standard irregular mechanism design problems the ironed marginal cost functions are constant on such ironing intervals. Here, the slope of the function $\underline{C}(\cdot, \cdot)$ varies with Q over the interval $Q \in \big(S(\underline{w}), w_1^{-1}(\underline{w}; m)\big)$ because the Lagrange multiplier (i.e. shadow price) associated with the minimum wage constraint decreases as Q increases.

(a) Non-constant ironing

(b) Optimal quantity $Q^*(\underline{w})$

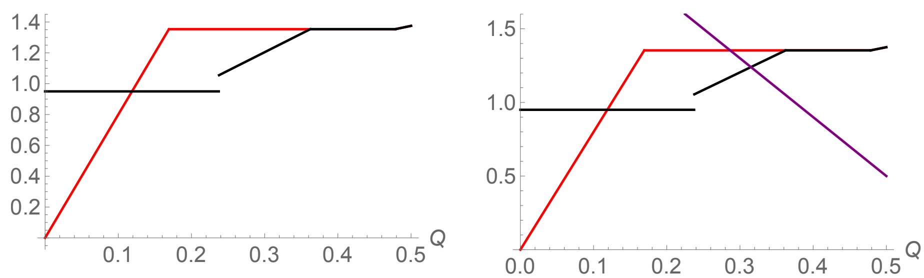


Figure 4: Panel (a) illustrates $\underline{C}'(Q, \underline{w})$ (black) and $\underline{C}'(Q)$ (red) for our piecewise linear leading example. Panel (b) illustrates the optimal quantity, which is given by the intersection of $V(Q)$ (purple) and $\underline{C}'(Q, \underline{w})$.

From an economic perspective, the fact that the marginal cost of procurement decreases in $\underline{w}$ for $Q \in (S(\underline{w}), w_1^{-1}(\underline{w}; m))$ (the second inequality in the display in Lemma 2) is key for the possibility that an increase in the minimum wage increases employment. This is illustrated in the right-hand panel in Figure 4. Here, the optimal quantity given the minimum wage $\underline{w}$, which we denote $Q^*(\underline{w})$, is given by the point of intersection of $V(Q)$ and $\underline{C}'(Q, \underline{w})$. Since $\underline{C}'(Q, \underline{w})$ decreases in $\underline{w}$ and $V(Q)$ decreases in Q, it follows that a local increase in the minimum wage increases employment. The following lemma characterizes the optimal quantity procured given $\underline{w} \in (W(Q_1(m)), W(Q_2(m)))$ when $Q^* \in (Q_1(m), Q_2(m))$ formally and completely. It also relates the minimum wage and the resulting optimal quantity procured to wage dispersion and involuntary unemployment.

Lemma 3. *Assume $Q^* \in (Q_1(m), Q_2(m))$. Then, for $\underline{w} \in (W(Q_1(m)), W(Q_2(m)))$, the optimal quantity procured, denoted $Q^*(\underline{w})$, is such that $V(Q^*(\underline{w})) = \underline{C}'(Q^*(\underline{w}), \underline{w})$, provided such a quantity exists. If there is no Q such that $V(Q) = \underline{C}'(Q, \underline{w})$, we have $Q^*(\underline{w}) = S(\underline{w})$. For $\underline{w} \in (W(Q_1(m)), W(Q_2(m)))$, the optimal procurement mechanism given $\underline{w}$ involves wage dispersion if and only if $Q^*(\underline{w}) > S(\underline{w})$. Moreover, provided $Q^*(\underline{w}) \neq S(\underline{w})$, this mechanism induces involuntary unemployment.*

Before we state more detailed comparative statics concerning these minimum wage effects, we need to introduce some additional notation. Specifically, we let $\gamma(Q;m)$ denote the marginal cost of procuring the quantity Q as the minimum wage $\underline{w}$ approaches $W(Q)$ from below. That is,

$$\gamma(Q;m) := \lim_{\underline{w} \uparrow W(Q)} \underline{C}'(Q, \underline{w}).$$

As is illustrated in Figure 4, the marginal cost function $\underline{C}'(Q, \underline{w})$ may be discontinuous at the point $Q = S(\underline{w})$, where the optimal procurement mechanism involves posting a market-clearing wage of $\underline{w}$. Given any sufficiently small $\epsilon > 0$, the optimal mechanism for procuring the quantity $Q - \epsilon$ is a two-price mechanism with rationing at the minimum wage $\underline{w}$ and the optimal mechanism for procuring the quantity $Q + \epsilon$ involves hiring some workers with certainty at the minimum wage $\underline{w}$ and rationing others at an efficiency wage. This difference between the left-hand and right-hand mechanisms explains why the marginal cost function $\underline{C}'(Q, \underline{w})$ is not necessarily continuous at $Q = S(\underline{w})$. The function $\gamma(Q;m)$ corresponds to left-hand value of $\underline{C}'(Q', \underline{w})$ when $Q' = Q$ and $\underline{w} = W(Q)$.

The significance of this function is illustrated in Figure 2 for the special case where $\underline{w} = W(Q_2(m))$. Here, we see that intersections between the functions γ and V dictate when we enter regions such as the one illustrated in the right-hand panel of Figure 2, where wage dispersion and involuntary unemployment are eliminated. In order to state general comparative statics concerning minimum wage effects, we have to account for these regions. To that end, we introduce two important quantity cutoffs in the following lemma.

Lemma 4. *Assuming that $Q^* \in (Q_1(m), Q_2(m))$ for some $m \in \mathcal{M}$, there exists quantity cutoffs $\hat{Q}_L(m)$ and $\hat{Q}_H(m)$ given by*

$$\hat{Q}_L(m) := \min\{Q : \gamma(Q;m) = V(Q)\} \quad \text{and} \quad \hat{Q}_H(m) := \max\{Q : \gamma(Q;m) = V(Q)\}$$

satisfying $\hat{Q}_H(m) \leq Q^p$ and $Q^* < \hat{Q}_L(m) \leq \hat{Q}_H(m) < Q_2(m)$.

We are now in a position to state and prove a series of propositions which state comparative statics that specify how wage dispersion, involuntary unemployment, and employment vary as the minimum wage $\underline{w} \in (w_1(Q^*;m), W(Q_2(m))]$ increases.

Proposition 3. *Suppose that $Q^* \in (Q_1(m), Q_2(m))$ for some $m \in \mathcal{M}$. Then for all $\underline{w} \in (w_1(Q^*;m), W(\hat{Q}_L(m)))$,*

(i) there is wage dispersion and involuntary unemployment; and

(ii) increasing $\underline{w}$ decreases involuntary unemployment and wage dispersion, and increases employment.

Proposition 3 covers the first region from Figure 3. The fact that an increase in the minimum wage at $\underline{w} = w_1(Q^*; m)$ has a positive effect on employment is noteworthy in itself because $w_1(Q^*; m) < W(Q^*)$. That is, $w_1(Q^*; m)$ is below the market-clearing wage for the quantity Q^*. In models in which market-clearing wages are imposed, minimum wages that are so low are typically ineffective as is a minimum wage equal to $W(Q^*)$. The reason for this positive quantity effect of such "small" minimum wages here is that even with the minimum wage, the optimal procurement contract involves wage dispersion and involuntary unemployment.

To gain intuition as to why wage dispersion decreases in $\underline{w}$ in this region, the following lemma is useful. It describes some formal properties of the optimal procurement mechanism in the region where the minimum wage constraint is binding and the monopsony optimally uses a two-price mechanism.

Lemma 5. *Given any $m \in \mathcal{M}$, suppose that $\underline{w} \in (W(Q_1(m)), W(Q_2(m)))$ and $Q \in \big(S(\underline{w}), w_1^{-1}(\underline{w}; m)\big)$. For $i \in \{1, 2\}$, let $q_i^*(Q, \underline{w})$ denote the solution value of q_i in (7). Then $q_1^*(Q, \underline{w})$ increases in $\underline{w}$ and decreases in Q and $q_2^*(Q, \underline{w})$ decreases in $\underline{w}$ and increases in Q.*

In the proof of Proposition 3 we show that an increase in $\underline{w}$ reduces wage dispersion by both decreasing the high wage paid in equilibrium and increasing the low wage paid in equilibrium, provided there is wage dispersion in equilibrium. That the low wage increases in $\underline{w}$ is trivial since this wage is simply the minimum wage itself. That the high wage $W(q_2^*(Q^*(\underline{w}), \underline{w}))$ decreases in $\underline{w}$ is less obvious due to the countervailing effects that changes in $\underline{w}$ have on $q_2^*(Q^*(\underline{w}), \underline{w})$.[15] However, the fact that $q_2^*(Q^*(\underline{w}), \underline{w})$ decreases in $\underline{w}$ and the fact that $Q^*(\underline{w})$ increases in $\underline{w}$ together imply that involuntary unemployment decreases in $\underline{w}$.[16]

Proposition 4. *Suppose that $Q^* \in (Q_1(m), Q_2(m))$ for some $m \in \mathcal{M}$. Then, for all $\underline{w} \in (W(\hat{Q}_H(m))), W(Q_2(m))]$, there is no wage dispersion. Moreover,*

(i) for all $\underline{w} \in [W(\hat{Q}_H(m)), \min\{W(Q^p), W(Q_2(m))\})$, there is no involuntary unemployment and employment increases in $\underline{w}$; and

(ii) for all $\underline{w} \in (\min\{W(Q^p), W(Q_2(m))\}, W(Q_2(m))]$, there is involuntary unemployment, which increases in $\underline{w}$, while employment decreases in $\underline{w}$, provided $\underline{w} < V(0)$.

[15] From Lemma 5, we know that $q_2^*(Q, \underline{w})$ increases in Q and decreases in $\underline{w}$. Since $Q^*(\underline{w})$ increases in $\underline{w}$ showing that that $q_2^*(Q^*(\underline{w}), \underline{w})$ is decreases in $\underline{w}$ requires showing that the latter effect dominates the former effect.

[16] It further implies that the unemployment rate, defined as $\frac{q_2^*(Q^*(\underline{w}), \underline{w}) - Q^*(\underline{w})}{q_2^*(Q^*(\underline{w}), \underline{w})} = 1 - \frac{Q^*(\underline{w})}{q_2^*(Q^*(\underline{w}), \underline{w}))}$, decreases in $\underline{w}$ because $Q^*(\underline{w})$ increases and $q_2^*(Q^*(\underline{w}), \underline{w})$ decreases.

Proposition 4 applies to the second and third regions from Figure 3. It distinguishes between whether the efficient quantity, which as mentioned is denoted by Q^p, is smaller or larger than $Q_2(m)$. If $Q^p \leq Q_2(m)$ then statement (a) describes the comparative statics for the second region (the "pure Stigler" region) where there is no wage dispersion, no involuntary unemployment and employment is increasing in the minimum wage $\underline{w}$.[17] Moreover, statement (b) describes the comparative statics for the third region (the "textbook" region) where there is no wage dispersion, involuntary unemployment increases in $\underline{w}$ and employment decreases in $\underline{w}$.

If in addition to W being piecewise linear V is weakly concave, then $\hat{Q}_L = \hat{Q}_H$. Consequently, Propositions 3 and 4 provide a complete characterization of the minimum wage effects for $\underline{w} \in (w_1(Q^*; m), W(Q_2(m))]$ and there is no region in which increases in the minimum wage induce wage dispersion and involuntary unemployment (while still increasing employment).[18] Figure 5 illustrates these effects and the comparative statics from Propositions 3 and 4 for our piecewise linear specification with $Q^*, Q^p \in (Q_1, Q_2)$. Here we can see that if there is wage dispersion in equilibrium, increasing the minimum wage will increase employment and decrease both involuntary unemployment and wage dispersion. Note that wage dispersion and involuntary unemployment vanish before the minimum wage reaches $W(Q^p)$, which is typically the case since $\hat{Q}_H(m) = Q^p$ is non-generic as described in Footnote 17. For $\underline{w} \in [W(\hat{Q}_H(m), W(Q^p))$, increasing the minimum wage has the effect of increasing employment as observed by Stigler (1946). For $\underline{w} > W(Q^p)$, increasing the minimum wage has the textbook effect of decreasing employment and increasing involuntary unemployment.

In Note A.1 we discuss the effects of minimum wages above $W(Q_2(m))$. These results are needed for the proof of the general theorem stated at the end of this subsection. When $Q_2(m) \geq Q^p$, statement (ii) from Proposition 4 still applies in this case. However, when $Q^p > Q_2(m)$, what happens past $W(Q_2(m))$ depends on whether there is another ironing range before one reaches $W(Q^p)$.

Proposition 5. *Suppose that $Q^* \in (Q_1(m), Q_2(m))$ for some $m \in \mathcal{M}$ and that $\hat{Q}_L(m) < \hat{Q}_H(m)$. Then the set $\{Q : \gamma(Q; m) = V(Q)\}$ contains an even number k of quantity cutoffs $\hat{Q}_j(m)$ that we index in increasing order so that $\hat{Q}_1(m) = \hat{Q}_L(m)$ and $\hat{Q}_k(m) = \hat{Q}_H(m)$.*

[17] If $W(Q)$ is piecewise linear of the form in (3) then, aside from knife-edge cases in which $V(\underline{q}) = W(\underline{q})$, we have $\hat{Q}_H < Q^p$ and such a region exists. (It can be shown that for W piecewise linear, $\underline{q}$ is the only point between Q_1 and Q_2 at which $\underline{C}'$ is continuous.)

[18] The function γ is piecewise linear and convex when W is piecewise linear. Moreover, if V is concave then these functions can only intersect once on $(Q_1(m), Q_2(m))$. (It might seem that V, if linear, could coincide with the downward sloping part of γ, which would mean that there is a continuum of points of overlap; but that is not possible because $Q^* \in (Q_1(m), Q_2(m))$ implies $V(Q^*) > \gamma(Q^*; m)$.)

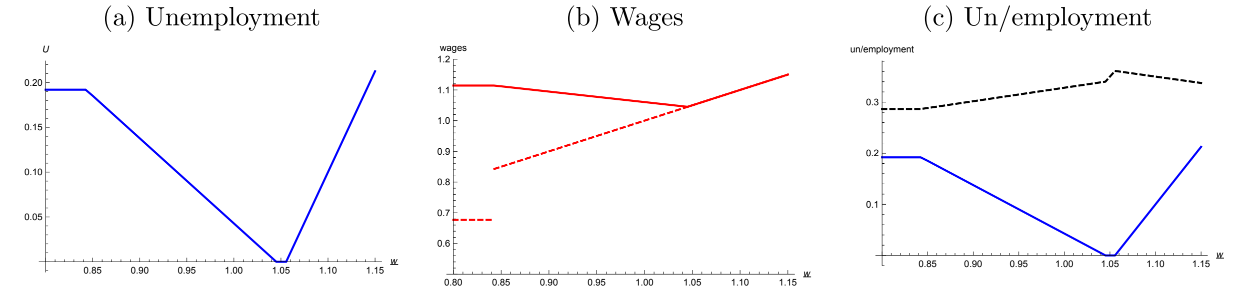


Figure 5: Equilibrium unemployment, wages and employment as a function of $\underline{w}$ for a case with $Q^* \in (Q_1(m), Q_2(m))$ and $Q^p < Q_2(m)$. All functions are constant for $\underline{w} \leq w_1(Q^*; m)$.

There is no wage dispersion for intervals with an upper quantity cutoff that corresponds to an even index and statements (i) and (ii) from Proposition 4 apply for intervals with an upper quantity cutoff that corresponds to an odd index.

The cutoff quantities that define each of the intervals identified in Proposition 5 each correspond to an intersection of the functions γ and V. In this region, which is not included in Figure 3, equilibrium behaviour alternates between regions where statements (i) and (ii) from Proposition 3 apply and where there is no wage dispersion. Consequently, this behaviour alternates between regions where there is and where there is no wage dispersion and involuntary unemployment. While employment continuously increases in the minimum wage over these intervals, at the point where one transitions from a region without involuntary unemployment and wage dispersion into one with involuntary unemployment, both involuntary unemployment and wage dispersion increase discontinuously. As the minimum wage increases further, both involuntary unemployment then decrease continuously and become zero at the end of the interval.

Figure 6 illustrates howa small increase in the minimum wage can lead to a discontinuous increase in involuntary unemployment, which is then followed by a continuous decrease of involuntary unemployment. The figure is plotted for a piecewise linear example in which Q^* is not part of the ironing range whereas Q^p is.[19] This implies that V first crosses γ from below on (Q_1, Q_2), which in turn implies that for $\underline{w}$ close to but above $W(Q_1)$ there is no wage dispersion but for larger values of $\underline{w}$ there is both.

Implications for regulators We conclude this section by addressing the question of how a regulator who observes wages and whether there is involuntary unemployment at a

[19] Specifically, the figure assumes $V(Q) = v$ for $Q \leq 1/4$ and $V(Q) = 0$ otherwise, with $v \in (W(Q_1), C'(Q_2))$, which can be thought of as the limit of a decreasing function, for a parameterization such that $W(Q_1) = 0.67$ and $W(Q_2) = 1.11$.

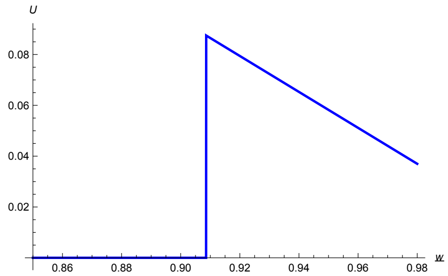


Figure 6: A discontinuous increase of unemployment is followed by a continuous decreases.

given minimum wage can gauge whether marginally increasing the minimum will increase employment.

One implication of the above analysis is that the relationship between involuntary unemployment and minimum wages is non-monotone. If $Q^* \in (Q_1(m), Q_2(m))$, there is involuntary unemployment of size $Q_2(m) - Q^*$ without a minimum wage, or equivalently, for any minimum wage $\underline{w} \leq w_1(Q^*; m)$. For $\underline{w} \in (w_1(Q^*; m), W(\hat{Q}_L(m))$, involuntary unemployment decreases with $\underline{w}$ and becomes 0 at $\underline{w} = W(\hat{Q}_L(m))$. Whether it remains 0 or becomes positive again depends on whether or not $\hat{Q}_L(m) = \hat{Q}_H(m)$. In any case, employment is increasing in $\underline{w}$ for all $\underline{w} \in (w_1(Q^*; m), W(Q^p))$ and involuntary unemployment is 0 at $\underline{w} = W(Q^p)$. As $\underline{w}$ increases beyond $W(Q^p)$, there will be involuntary unemployment and employment decreases.

The richness and non-monotonicity of the aforementioned effects raises the question how a policymaker could assess whether marginally increasing the minimum wage decreases overall employment and increases or decreases involuntary unemployment. The answer is affirmative and relates to wage dispersion. If at the present minimum wage there is involuntary unemployment and wage dispersion, increasing the minimum wage will increase employment and decrease involuntary unemployment. Likewise, if at the present minimum wage there is no wage dispersion and no involuntary unemployment, increasing the minimum wage will increase employment unless $\underline{w} = W(Q^p)$. In this case, the increase in the minimum wage may induce involuntary unemployment. This increase will be discontinuous. But by the preceding argument, increasing the minimum wage further will reduce it while continuing to increase employment. In sharp contrast, if at the current minimum wage, there is involuntary unemployment and no wage dispersion, then increasing the minimum wage will increase involuntary unemployment and decrease employment.

Putting all of this together yields the following theorem, which specifies the circumstances in which a regulator can expect a local increase in the minimum wage to increase employment. Because the theorem follows from the above, we do not provide a separate proof.

Theorem 1. *Whenever there is involuntary unemployment and wage dispersion at a given* *minimum wage, a sufficiently small increase in the minimum wage increases employment and decreases involuntary unemployment. If there is involuntary unemployment and no wage dispersion at a given minimum wage, increasing the minimum wage decreases employment and increases involuntary unemployment. Moreover, provided $\underline{w} \neq W(Q^p)$, if there is no involuntary unemployment at a given minimum wage, a sufficiently small increase in the minimum wage increases employment.*

Our analysis in this section also points to the possibility of conflicting interests among employed workers concerning the introduction of a minimum wage $\underline{w} \in (w_1(Q^*; m), W(Q_2(m)))$. While those employed at the low wage benefit from the imposition of the minimum wage, workers who earn the high wage in the absence of a minimum wage are harmed by a minimum wage such that $\underline{w} < W(Q_2(m))$. Whenever there is wage dispersion in equilibrium, the high wage decreases in $\underline{w}$. For $\underline{w} \in [W(\hat{Q}_H(m)), W(Q_2(m))]$ all workers earn the minimum wage, and hence the equilibrium wage increases in $\underline{w}$ but is evidently less than $W(Q_2(m))$ for $\underline{w} < W(Q_2(m))$. This effect is also illustrated in Figure 5(b), where the high wage decreases in $\underline{w}$, provided there is wage dispersion and $\underline{w}$ impacts employment and involuntary unemployment. If $Q^p < Q_2(m)$, as is the case in this example, the workers who earn the high wage absent wage regulation are still worse off with a minimum wage equal to $W(Q^p)$ since $W(Q^p) < W(Q_2(m))$.

5 Quantity competition

A natural question that the analysis in Sections 3 and 4 raises is to what extent the effects identified generalize to (imperfectly) competitive environments. To address this question, we now extend the model to allow for quantity competition between firms. We first introduce the setup, derive the equilibrium and discuss its properties. Then we analyze the effects of minimum wages.

5.1 Setup

Suppose now that there are n firms procuring labor. We index these firms by i and for each firm i the marginal value for procuring the x_i-th unit of labor is given by a decreasing function $V(x_i)$, where we use x_i to distinguish individual firms' quantities from the q_1 and q_2 in the previous section. The firms compete in quantities as follows. They simultaneously submit quantities x_i to a Walrasian auctioneer as in standard oligopoly and oligopsony models with quantity competition. However, rather than procuring the $Q := \sum_{i=1}^{n} x_i$ units at the market-clearing wage $W(Q)$, which is the standard assumption in Cournot models and leads to a procurement cost function of C, we assume that the auctioneer can offer an efficiency wage and procure the Q units at minimal total cost $\underline{C}(Q)$. Firm i who employs x_i units has to pay the cost $\frac{x_i}{Q}\underline{C}(Q)$. Modulo replacing the cost function C with $\underline{C}$, this is the same as in standard Cournot models since $\frac{x_i}{Q}\underline{C}(Q) = x_i W(Q)$ for $Q \notin (Q_1(m), Q_2(m))$ for any $m \in \mathcal{M}$. The efficient quantity for a given n is denoted by Q_n^p and such that

$$V\left(\frac{Q_n^p}{n}\right) = W(Q_n^p),$$

which is the quantity that would emerge if the firms were price-takers.

5.2 Equilibrium

The analysis of the previous section then extends to this model, insofar as we will have involuntary unemployment and efficiency wages whenever $Q \in (Q_1(m), Q_2(m))$ for some $m \in \mathcal{M}$.

Denote by Q_n^* the aggregate quantity in a symmetric equilibrium under quantity competition. In models in which market-clearing wages are imposed, this quantity satisfies

$$V\left(\frac{Q_n^*}{n}\right) = W(Q_n^*) + \frac{Q_n^*}{n}W'(Q_n^*), \tag{8}$$

provided a symmetric equilibrium exists. Because $W' > 0$, it follows that $Q_n^* < Q_n^p$, that is, the equilibrium quantity is inefficiently small. As the following proposition shows, in our model of quantity competition the equilibrium is always unique and symmetric. However, for n sufficiently large, $Q_n^p < Q_n^*$ is possible, that is, the equilibrium quantity can be excessively large. To develop an understanding of how such a reversal can occur for n sufficiently large, consider the first-order condition under symmetry,

$$V\left(\frac{Q}{n}\right) = \frac{n-1}{n}\frac{\underline{C}(Q)}{Q} + \frac{1}{n}\underline{C}'(Q),$$

whose right-hand side we denote by $h(Q,n)$. If $Q \in (Q_1(m), Q_2(m))$ for some $m \in \mathcal{M}$, then $h(Q,n)$ is increasing and concave in Q and for any finite n satisfies $h(Q_i(m), n) > W(Q_i(m))$ Moreover, $h(Q,n)$ decreases in n and satisfies $h(Q,1) > W(Q)$ for all $Q \in (Q_1(m), Q_2(m)))$. In contrast, for n sufficiently large, there exists at least one interval $(a_n, b_n) \subset (Q_1(m), Q_2(m))$ such that $h(Q,n) < W(Q)$ for all $Q \in (a_n, b_n)$, where a_n decreases in n and b_n increases in n.[20] Consequently, if $V(Q/n) = h(Q,n)$ for $Q \in (a_n, b_n)$, then

[20] If there are multiple subintervals over which $h(Q,n) < W(Q)$ for some n, index these by k. Then for

$Q^*_n \in (a_n, b_n)$ and $Q^p_n < Q^*_n$. Figure 7 illustrates the relation between $W(Q)$ and $h(Q, n)$ as a function of n for the leading example with a piecewise linear supply function. Intuitively, the first-order condition implies that a firm's perceived marginal cost of procurement $h(Q, n)$ is a convex combination of $\underline{C}'$, which is larger than $W(Q)$, and $\underline{C}(Q)/Q$, which is less than $W(Q)$ for $Q \in (Q_1(m), Q_2(m))$. As n increases, the weight on $\underline{C}(Q)/Q$ increases, eventually leading to $h(Q, n) < W(Q)$ for some Q.

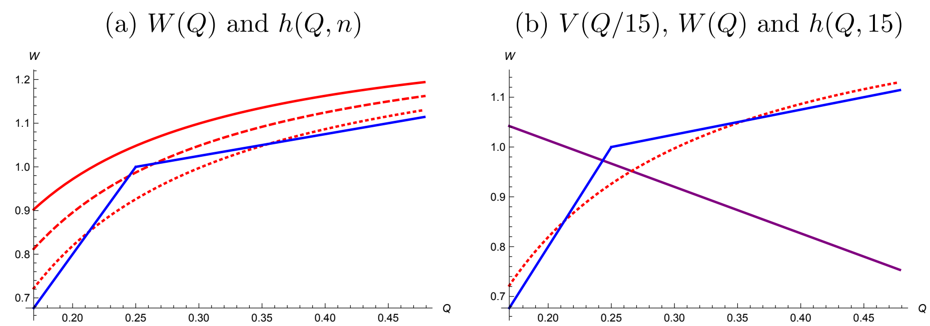


Figure 7: Lefthand panel: $W(Q)$ in blue and $h(Q, n)$ in solid red for $n = 3$, in red dashed for $n = 5$ and red dotted for $n = 15$. Righthand panel: $Q^p_n < Q^*_n$ and $V(Q/n) = 1.2 - 14Q/n$ for $n = 15$.

As n goes to infinity, Q^p_n converges to the efficient (or Walrasian) quantity Q^e, which is such that $V(0) = W(Q^e)$. In the last part of the following proposition, we need to distinguish between the case where there is no $m \in \mathcal{M}$ such that $Q^e \in (C(Q_1(m), Q_2(m)))$ and the case where there exists a $m_e \in \mathcal{M}$ such that $Q^e \in (Q_1(m_e), Q_2(m_e)))$. Observe that in the latter case

$$\underline{C}'(Q^e) = C'(Q_2(m_e)) > W(Q_2(m_e)) > W(Q^e).$$

That is, $\underline{C}'(Q^e) > V(0)$.

Proposition 6. *The quantity setting game has a unique equilibrium, and this equilibrium is symmetric. The aggregate equilibrium quantity Q^*_n is increasing in n. If $Q^p_n \leq Q^*_n$, then $n > 1$ and $Q^*_n \in (Q_1(m), Q_2(m))$ for some $m \in \mathcal{M}$. As $n \to \infty$, we have $Q^*_n \to Q^e$ if $C(Q^e) = \underline{C}(Q^e)$ and otherwise, we have $Q^*_n \to \tilde{Q}$, where $\tilde{Q}$ satisfies $Q^e < \tilde{Q} < Q_2(m_e)$.*

As stated in Proposition 6, key features of the monopsony model —efficiency wages, involuntary unemployment—extend to quantity competition. Moreover, there is no monotone relationship between competition and involuntary unemployment as increasing competition can bring equilibrium quantity into or out of an ironing interval $(Q_1(m), Q_2(m))$.

each k, a^k_n is decreasing in n and b^k_n is increasing in n because h decreases in n. Of course, eventually two or more of these subintervals may collapse into one, that is if $b^k_n < a^{k+1}_n$, we may have $b^k_{n'} \geq a^{k+1}_{n'}$ for some $n' > n$. But this does not invalidate the point that the set of $Q \in (Q_1(m), Q_2(m))$ for which $h(Q, n) < W(Q)$ increases in n in the set inclusion sense.

Moreover, within such an interval, competition decreases wage dispersion and involuntary unemployment and increases $w_1^*(m)$ and employment, while leaving $w_2^*(m)$ fixed, where $w_j^*(m) = Q_j(m)$. Interestingly, when $\underline{C}(Q^e) < C(Q^e)$ we have involuntary unemployment and an efficiency wage even in the limit as $n \to \infty$, yielding what may appear to be a natural unemployment rate of $(Q_2(m_e) - \tilde{Q})/Q_2(m_e)$, since this is the unemployment rate associated with perfect competition. In contrast to the usual notion of a natural unemployment rate, this unemployment is a result of inefficient resource allocation in the form of both random allocation and excessive economic activity (since $\tilde{Q} > Q^e$). In other words, there is the possibility of inefficient perfect competition.

We now illustrate these effects for our leading example in Figures 8 and 9, using $V(x_i) = 1.1 - 8x_i$ for the left panels and $V(x_i) = 1.2 - 8x_i$ for the right panels. This implies that for the left panels, we have $Q^e = 0.45 \in (Q_1, Q_2) = (0.169, 0.478)$ and $\tilde{Q} = 0.4516$, while for the right panels we have $Q^e = 0.65 > Q_2$.

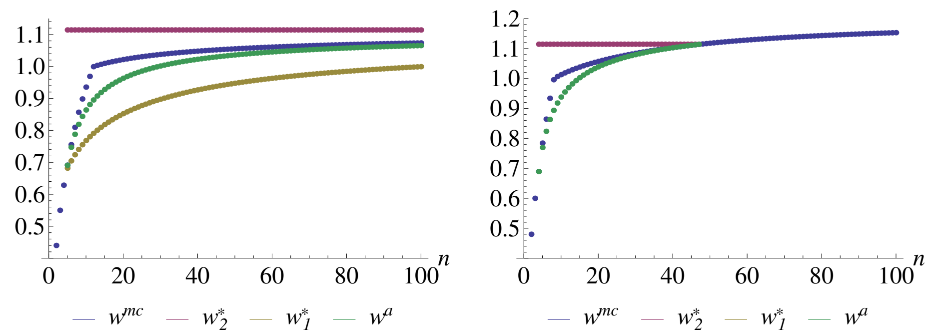


Figure 8: Equilibrium wages as a function on n, where $w^{mc} = W(Q_n^*)$ denotes the market-clearing wage and w^a the average wage $w^a = (w_1^* + w_2^*)/2$. On the left, $W(Q^e) = 1.1 < 1.114 = w_2^*$ and on the right $W(Q^e) = 1.2 > w_2^*$.

5.3 Minimum wage effects and competition

As is reasonably well known, in models with quantity competition and market-clearing wages, minimum wages above the market-clearing wage for the equilibrium quantity Q_n^* absent wage regulation have a positive effect on total employment and, accordingly, workers' pay. To see this, recall that Q_n^p is such that $V\left(\frac{Q_n^p}{n}\right) = W(Q_n^p)$ while the equilibrium quantity satisfies (8), which together with $W' > 0$ implies $Q_n^* < Q_n^p$. Then any minimum wage $\underline{w} \in (W(Q_n^*), W(Q_n^p)]$ has a positive employment effect. Insofar as $\lim_{n\to\infty} Q_n^p = Q^e = \lim_{n\to\infty} Q_n^*$, there is a sense in which the scope for this kind of quantity and social-surplus

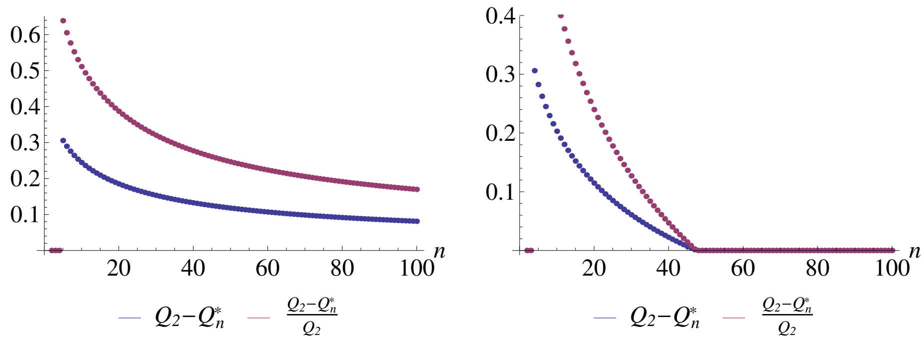


Figure 9: Involuntary unemployment and the unemployment rate as a function on n. On the left, as there is involuntary unemployment of size $Q_2 - \tilde{Q} = 0.0269$ and an unemployment rate of 5.6% as $n \to \infty$.

increasing minimum wage regulation decreases with n since in the limit any such scope has vanished.[21]

Of course, if $Q_n^* \notin (Q_1(m), Q_2(m))$ for any $m \in \mathcal{M}$, then (8) also characterizes the equilibrium quantity in our model in which market-clearing wages are not imposed. Therefore, the same is true with regard to minimum wage effects in our model if, in addition, there is no ironing interval between Q_n^* and Q_n^p.[22] Consequently, the questions of interest concern minimum wage effects when the equilibrium without wage regulation involves involuntary unemployment, that is, if $Q_n^* \in (Q_1(m), Q_2(m))$ for some $m \in \mathcal{M}$. As we show next, similar effects obtain in this model, but there are also subtle yet important differences. A minimum wage slightly above the lower of the two wages absent regulation, that is, $w_1(Q_n^*; m)$, always increases employment and decreases involuntary unemployment.

Because the minimum cost of procurement given a minimum wage $\underline{w}$, $\underline{C}(Q, \underline{w})$ defined in (6), is convex and increasing Q, the results stated in Proposition 6 carry over to the model with a minimum wage, with only one qualification regarding the uniqueness of equilibrium in the case when $V(S(\underline{w})/n)) > \underline{w}$ and all workers that are employed are paid $\underline{w}$. If this occurs, we assume that all firms are rationed equally so that each firm obtains $S(\underline{w})/n$ workers.[23] In particular, the condition

$$V\left(\frac{\underline{Q}_n(\underline{w})}{n}\right) = \frac{n-1}{n}\frac{\underline{C}(\underline{Q}_n(\underline{w}), \underline{w})}{\underline{Q}_n(\underline{w})} + \frac{1}{n}\frac{\partial \underline{C}(\underline{Q}_n(\underline{w}), \underline{w})}{\partial Q} \tag{9}$$

[21] Whether the differences $W(Q_n^p) - W(Q_n^*)$ and $Q_n^p - Q_n^*$ monotonically decrease in n depends on the specifics of the model. If W and V are both linear, then both $W(Q_n^p) - W(Q_n^*)$ and $Q_n^p - Q_n^*$ decrease in n.

[22] That is, if $[Q_n^*, Q_n^p] \cap \bigcup_{m \in \mathcal{M}} (Q_1(m), Q_2(m)) = \emptyset$.

[23] This multiplicity issue is not germane to our problem. It occurs in typical models of quantity competition with uniform pricing when the price is regulated.

still characterizes the symmetric equilibrium. Moreover, the equilibrium is unique if $\underline{Q}_n(\underline{w}) > S(\underline{w})$ because $\underline{C}(Q, \underline{w})$ is convex.

To see when $Q_n^p \leq Q_n^*$ can occur and what it entails, recall from Proposition 6 that $Q_n^p \leq Q_n^*$ implies $Q_n^* \in (Q_1(m), Q_2(m))$ for some $m \in \mathcal{M}$. Observe that $h(Q_i, n) = W(Q_i) + \frac{Q_i}{n}W'(Q_i) > \frac{Q_i}{n}W'(Q_i)$ for $i = 1, 2$, where we write Q_i in lieu of $Q_i(m)$ to ease the notation. Recall also that $h(Q, n)$ is increasing and concave. Therefore, it satisfies, for all $Q \in [Q_1, Q_2]$,

$$h(Q, n) > W(Q_1) + (Q - Q_1)\frac{W(Q_2) - W(Q_1)}{Q_2 - Q_1} = (1 - \alpha)W(Q_1) + \alpha W(Q_2).$$

That is, $h(Q, n)$ is everywhere above the linear function that goes through $(Q_1, W(Q_1))$ and $(Q_2, W(Q_2))$. Expressed in terms of h and V, $Q_n^* \in (Q_1, Q_2)$ satisfies $V(Q_n^*/n) = h(Q_n^*)$. Since Q_n^p is defined by the equality $V(Q_n^p/n) = W(Q_n^p)$, it follows that for $Q_n^* \in (a_n, b_n)$ we have $Q_n^p < Q_n^*$. Observe that $Q_n^p > Q_1$. This holds because $W(Q_n^p) > h(Q_n^*, n) > W(Q_1)$ and because W is increasing. Notice also that we have

$$w_1(Q_n^*; m) = W(Q_1(m)) + (Q_n^* - Q_1(m))\frac{W(Q_2(m)) - W(Q_1(m))}{Q_2(m) - Q_1(m)} < h(Q_n^*, n) < W(Q_n^p).$$

That is, it is always the case that $w_1(Q_n^*; m) < W(Q_n^p)$ (if $Q_n^* < Q_n^p$, this follows trivially from the fact that $w_1(Q_n^*; m) < W(Q_n^*) < W(Q_n^p)$).

Proposition 7. *Consider the model with quantity competition. Suppose $Q_n^* \in (Q_1(m), Q_2(m))$ for some $m \in \mathcal{M}$. Then at $\underline{w} = w_1(Q_n^*; m)$, the marginal effect of increasing the minimum wage is positive, that is, $\frac{d\underline{Q}_n(\underline{w})}{d\underline{w}}|_{\underline{w}=w_1(Q_n^*;m)} > 0$.*

6 Extensions and discussion

We now provide extensions of the model to allow for differentiated jobs and discussion of the effects of policies that prohibit or permit wage discrimination. The section concludes with the analysis of comparative statics and a microfoundation for non-convex costs of procurement under market-clearing wages based on fixed costs of migration. .

6.1 Horizontally differentiated jobs

We first consider a monopsony problem with horizontally differentiated jobs and workers.

Setup Specifically, we now consider a variant of the Hotelling model in which a monopsony with jobs at locations 0 and 1 has a willingness to pay of $V(Q)$ for the Q-th worker employed at a given location $\ell \in \{0, 1\}$. As before, $V(Q)$ is assumed to be decreasing. There is a continuum of workers with linear transportation costs whose locations are uniformly distributed between 0 and 1 and private information of each worker. The total mass of consumers is 1. The value of the outside option of worker is normalized to 0.[24] The payoff of a worker at location x that works at 0 for a wage of w is $w - x$, while this worker's payoff of working at 1 for a wage of w is $w - (1 - x)$. Observe that this implies that the market-clearing wage to hire Q workers at given location is $W(Q) = Q$, which in turn means that the cost of procurement at each location under market-clearing wages is $C(Q) = Q^2$. Of course, the monopsony can hire Q_ℓ workers at $\ell = 1, 2$ if and only if $Q_0 + Q_1 \leq 1$.

Equilibrium We first derive the minimum cost to procure a quantity $Q \in [0, 1/2]$ at given location, given that the same quantity is procured at the other location as well. To this end, notice first that the expected transportation cost of worker at any location $x \in [0, 1]$ who is equally like to work at location 0 and at location 1, conditional on being employed, is 1/2, conditional on being employed. To satisfy the individual rationality constraints of workers employed under these terms and conditions, the wage they receive has to be no less than 1/2. Consequently, by paying a wage of 1/2 and leaving workers employed at this wage in the dark as to where they will work, or having them multi-task by having them spend half of their time at either location, the monopsony can procure any quantity $Q \in [0, 1/2]$ at both locations at a marginal procurement cost of 1/2. Since the marginal cost of procuring Q at a market-clearing wage is $C'(Q) = 2Q$, it follows that the monopsony can procure the quantity $Q \in [0, 1/2]$ at each location at the cost

$$\underline{C}(Q) = \begin{cases} Q^2, & Q \in [0, 1/4] \\ Q/2 - 1/16, & Q \in (1/4, 1/2] \end{cases}$$

by offering a wage of 1/2 to attract "universalists"—workers who are willing to do either job—and a wage of 1/4 to attract "specialists," that is workers with locations no further away from 0 and 1 than 1/4 who will be guaranteed to do the job closest to their location. Notice that individual rationality constraint will bind for all workers who are employed with locations $x \in (1/4, 3/4)$. Consequently, for the marginal worker at 1/4, the incentive compatibility constraint that this worker be indifferent between working as a specialist or as a universalist, coincides with this worker's individual rationality constraint. The preceding arguments establish that this scheme with wage dispersion and random worker-job matchings

[24] This is without loss of generality within the domain of problems in which the value of the outside option and the willingness to pay per worker are independent of the workers' locations since all that matters for these problems is the difference between the latter and the former.

results in smaller procurement costs than market-clearing wages for any $Q \in (1/4, 1/2]$. Arguments along the lines of those in Balestrieri et al. (2021) and Loertscher and Muir (2021b), who study optimal selling mechanisms on the Hotelling line, can be used to establish that $\underline{C}(Q)$ is indeed the minimal cost of procurement, subject to workers' incentive compatibility and individual rationality constraints.[25]

The equilibrium level of employment Q^* at each location is given by the unique number satisfying $V(Q^*) = \underline{C}'(Q^*)$. We say that the equilibrium involves *involuntary unemployment* if at the equilibrium wages there is a positive mass of workers who would be willing to work but are not employed, and we say that it involves *worker-job mismatching* if in equilibrium workers with $x < 1/2$ work at location 1 and workers with $x > 1/2$ at location 0.[26] The following proposition summarizes characteristics of the equilibrium. As it follows directly from the preceding arguments, we omit a proof.

Proposition 8. *If $V(1/4) \leq 1/2$, then $Q^* \leq 1/4$ and the equilibrium involves neither worker-job mismatchings nor involuntary unemployment. If $V(1/4) > 1/2 > V(1/2)$, then $Q^* \in (1/4, 1/2)$ and the equilibrium involves both worker-job mismatchings and involuntary unemployment. If $V(1/2) \geq 1/2$, then $Q^* = 1/2$ and the equilibrium involves worker-job mismatchings but no involuntary unemployment.*

Figure 10 illustrates the case $V(1/4) > 1/2 > V(1/2)$ in Proposition 8 for the linear specification $V(Q) = v - Q$ with $v = 7/8$. For this linear specification, $V(1/4) > 1/2 > V(1/2)$ is equivalent to $v \in (3/4, 1)$.

If a minimum wage of $\underline{w} = 1/2$ is imposed, the strict profitability of worker-jobs mismatching vanishes without any negative effects on the level of employment in equilibrium.

Effects of prohibiting wage discrimination The cost minimizing procurement mechanism involves wage dispersion or wage discrimination whenever the quantity procured at

[25] An outline of the argument, adapted from the monopoly screening problem in Loertscher and Muir (2021b) to the procurement setting and assuming all workers are employed, is as follows. Let $p_\ell(x)$ denote the probability that the worker who reports type $x \in [0, 1]$ works at location $\ell \in \{0, 1\}$. Incentive compatibility implies that $p_1(x) - p_0(x)$ be non-decreasing. Type $\hat{x}$ is worst-off type if $p_1(\hat{x}) = p_0(\hat{x})$. Because all workers are employed, we have $p_0(x) + p_1(x) = 1$, implying $p(x) \equiv p_0(x)$ is sufficient, and incentive compatibility becomes equivalent to $p(x)$ being non-increasing, and $\hat{x}$ is worst-off if $p(\hat{x}) = 1/2$. Given any worst-off type $\hat{x} \in [0, 1]$, incentive compatibility yields the designer's objective in terms of virtual costs and values. Because its pointwise minimizer is not monotone, one needs to iron the virtual types. (Put differently, cost of procurement is not convex in the Q of units procured from location 0.) Given the ironed objective, there is always a pointwise minimizer that also makes $\hat{x}$ worst off. So one is the left to minimize over $\hat{x}$, which in the uniform case yields $\hat{x} = 1/2$.

[26] If worker-job mismatching is optimal, workers who work at the high wage of 1/2 are indifferent between working and not. Thus, those—if any—who are involuntarily unemployed are also indifferent between being unemployed and working.

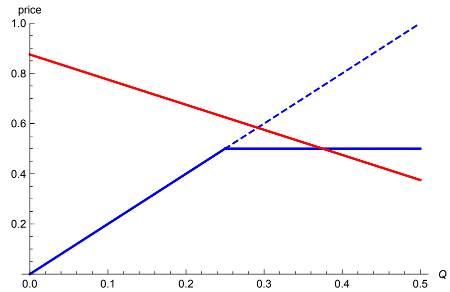


Figure 10: Illustration of Proposition 8 for $V(Q) = v - Q$ with $v = 7/8$ in red and $\underline{C}'$ in blue (and C' dashed).

each location is greater than 1/4. Wage discrimination is often perceived with suspicion in both public and academic debates and has led to pressure for pay transparency in wide range of jurisdictions.[27] We now briefly investigate the effects of prohibiting wage discrimination on the equilibrium quantity, involuntary unemployment, social surplus, workers' total pay and workers' surplus.

Worker-job mismatching being optimal is equivalent to wage discrimination being optimal. From Proposition 8 we know that worker-job mismatching is optimal if and only if $Q^* > 1/4$. Consequently, prohibiting or permitting wage discrimination has no effect on equilibrium outcomes if and only if $Q^* \leq 1/4$, which is equivalent to $V(1/4) \leq 1/2$. In what follows, we therefore focus on the cases with $V(1/4) > 1/2$.

The following effects of prohibiting wage discrimination hold in general, that is, without additional assumptions on $V(Q)$:

Proposition 9. *Assume $V(1/4) > 1/2$. Then prohibiting wage discrimination*

- *weakly decreases the equilibrium quantity and strictly decreases it if $V(1/2) < 1$;*
- *decreases the monopsony's profit;*
- *weakly increases the surplus of all workers and strictly increases the surplus of all but the marginal workers who are employed when wage discrimination is prohibited.*

Proposition 9 implies that for $V(1/4) > 1/2$ and $V(1/2) < 1$, prohibiting wage discrimination decreases both the equilibrium level of employment and eliminates involuntary unemployment. This is similar to the effects observed in Section 5 that employment and involuntary unemployment can move in the same direction. The unambiguous effects of prohibiting wage discrimination on the surplus of individual workers contrast with the effects of minimum wages in Section 4, where, as discussed, high wage earners are typically harmed by minimum wages.

[27]For a comprehensive list of recent references, see, for example, Cullen and Pakzad-Hurson (2021).

We conclude the analysis of prohibiting wage discrimination by studying the effects of wage discrimination on social surplus and total wage payments. Letting Q^* and $\tilde{Q}$ be the equilibrium quantities at each location with and without wage discrimination, the change in social surplus when wage discrimination is permitted compared to when it is not, denoted $\Delta SS(Q^*, \tilde{Q})$, is

$$\Delta SS(Q^*, \tilde{Q}) = \int_{\tilde{Q}}^{Q^*} (V(x) - 1/2)dx - \int_{1/4}^{\tilde{Q}} (1/2 - x)dx$$

while the change in total wage payments, denoted $\Delta C(Q^*, \tilde{Q})$, is

$$\Delta C(Q^*, \tilde{Q}) = \underline{C}(Q^*) - C(\tilde{Q}) = \frac{1}{2}Q^* - \frac{1}{16} - \tilde{Q}^2.$$

The intuition for $\Delta SS(Q^*, \tilde{Q})$ is simple. For all $x \in [\tilde{Q}, Q^*]$, $V(x) - 1/2$ is the social benefit of the additional unit procured with wage discrimination, $V(x)$, minus the cost of production of 1/2, while $\int_{1/4}^{\tilde{Q}}(1/2 - x)dx$ is the additional cost of production on the inframarginal units between 1/4 and $\tilde{Q}$ that are procured with and without wage discrimination.[28]

Recall first that permitting wage discrimination has a positive quantity effect if and only if $\tilde{Q} \in (1/4, 1/2)$.[29] Notice next that $\Delta SS(\tilde{Q}, \tilde{Q}) = \frac{1}{2}\tilde{Q}(\tilde{Q} - 1) + \frac{3}{32} < 0$, which is to say that a positive quantity effect is necessary, that is, $Q^* > \tilde{Q}$, for wage discrimination to increase social surplus, where the inequality follows because for any $\tilde{Q} \in (1/4, 1/2)$ since $\frac{1}{2}\tilde{Q}(\tilde{Q} - 1) < -\frac{3}{32}$.[30] Similarly, $\Delta C(\tilde{Q}, \tilde{Q}) = \frac{1}{2}\tilde{Q}(1 - 2\tilde{Q}) - \frac{1}{16} < 0$, meaning that without a quantity effect, wage payments decrease with wage discrimination.[31] Moreover, we have

$$\frac{\partial \Delta SS(Q^*, \tilde{Q})}{\partial Q^*} = V(Q^*) = \frac{1}{2} = \frac{\partial \Delta C(Q^*, \tilde{Q})}{\partial Q^*}, \tag{10}$$

where the second equality makes use of the first-order condition $V(Q^*) = 1/2$. The proof of the following proposition makes use of these insights and provides the additional steps necessary to establish that:

[28]The expression for $\Delta C(Q^*, \tilde{Q})$ follows straightforwardly by plugging in Q^* and $\tilde{Q}$ into $\underline{C}$ and C, respectively. To see that $\Delta SS(Q^*, \tilde{Q})$ is correct, observe that social surplus with wage discrimination is $\int_0^{Q^*} V(x)dx - \int_0^{1/4} xdx - (Q^* - 1/2)\frac{1}{2}$ while social surplus without wage discrimination is $\int_0^{\tilde{Q}} V(x)dx - \int_0^{\tilde{Q}} xdx$. Subtracting the latter from the former yields $\Delta SS(Q^*, \tilde{Q})$.

[29]For $\tilde{Q} \leq 1/4$, permitting wage discrimination does not affect anything while for $\tilde{Q} = 1/2$, the only effect of permitting wage discrimination is to decrease the procurement cost by mismatching workers to jobs.

[30]The function $\frac{1}{2}\tilde{Q}(\tilde{Q} - 1)$ is convex in $\tilde{Q}$ on $[1/4, 1/2]$, minimized at $\tilde{Q} = 1/2$ and thus maximal at $\tilde{Q} = 1/4$, at which point it is $-3/32$.

[31]To see this, notice that $\frac{1}{2}\tilde{Q}(1 - 2\tilde{Q})$ is maximized at $\tilde{Q} = 1/4$, at which point it equals 1/16. Since $\tilde{Q} > 1/4$, the inequality follows.

Proposition 10. $\Delta C(Q^*, \tilde{Q}) \leq 0$ *implies* $\Delta SS(Q^*, \tilde{Q}) < 0$.

Proposition 10 does not say whether wage discrimination can increase social surplus but merely states that if it does, it will also increase total wage payments. To see that it is indeed possible for permitting wage discrimination to increase social surplus, it is useful to consider the limiting case of a $V(Q)$ decreasing in which case $V(Q) = v$ for all $Q \in [1/4, 1/2]$. For $v \in (1/2, 1)$, this implies $\tilde{Q} = v/2 \in (1/4, 1/2)$ and $\tilde{Q}^* = 1/2$. In this case,

$$\Delta SS(Q^*, \tilde{Q}) = (2\tilde{Q} - 1/2)(Q^* - \tilde{Q}) - \frac{1}{2}\tilde{Q}(1 - \tilde{Q}) + \frac{3}{32} = \frac{2\tilde{Q} - 3\tilde{Q}^2}{2} - \frac{5}{32},$$

where the first equality uses $V(\tilde{Q}) = 2\tilde{Q}$ and the second equality follows from substituting $Q^* = 1/2$ and simplifying. Expressing $\Delta SS(Q^*, \tilde{Q})$ in this way highlights the dual or countervailing role of $\tilde{Q}$: If $\tilde{Q}$ is small, the additional costs due to wage discrimination are small and the benefits $v - 1/2$ are enjoyed over a large domain, namely from $\tilde{Q}$ to $1/2$, but these benefits are themselves small because $\tilde{Q}$ being small means that v is small. Maximizing $\Delta SS(1/2, \tilde{Q})$ over $\tilde{Q}$ yields $\tilde{Q} = 1/3$, which corresponds to $v = 2/3$, and $\Delta SS(1/2, 1/3) = 5/18 - 5/32 > 0$.[32] With constant willingness to pay v, one can show that social surplus increases with wage discrimination if and only if $v \in (1/2, 5/6)$.

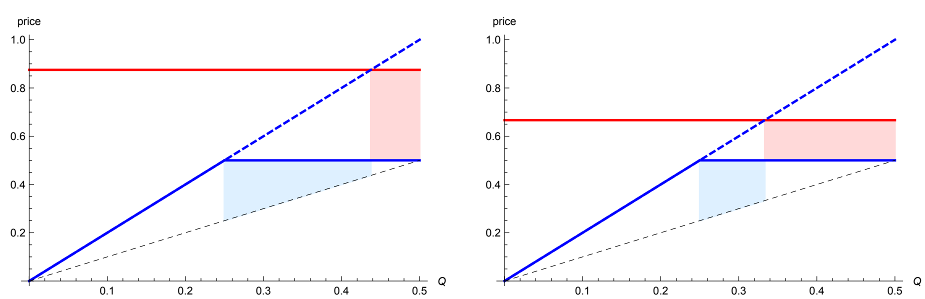


Figure 11: Permitting wage discrimination decreases social surplus in the lefthand panel ($v = 7/8$) and increases it in the righthand panel ($v = 2/3$). Assume $V(Q) = v$, plotted in red. $\underline{C}'$ is plotted in solid blue, C' in blue dashed and $W(Q) = Q$ in black dashed.

Mechanism design and horizontal mergers Suppose now that the plants operated at $\ell \in \{0, 1\}$ are independently owned and operated by firms that each have the same marginal benefit function $V(q)$, so that each maximizes its own profit. Suppose further that, as a stand alone firm, each firm's profit is maximized at a quantity q^* satisfying $q^* \in (1/4, 1/2)$,

[32]If $V(Q) = v - Q$ with $v \in (3/4, 1)$, which implies $Q^* \in (1/4, 1/2)$, then we have $\Delta C(Q^*, \tilde{Q}) > 0 > \Delta SS(Q^*, \tilde{Q})$, that is, prohibiting wage discrimination decreases total wage payments and increases social surplus.

that is, $V(q^*) \in (1/2, 1)$. From the above analysis, we know that a monopsony firm operating both plants jointly would choose a larger employment level at each location. For example, in the case where $V(q) = v$ for all $q \in [1/4, 1/2]$, the multi-jobs monopsony employs a quantity of 1/2 at each location while each independent firm only hires $v/2$ workers. This means that the model exhibits the feature that horizontal mergers are not only profitable but may also increase social surplus since the social surplus effects of the merger are the same as those of allowing the monopsony to wage discriminate. For the case when $V(q) = v$ is constant for $q \in [1/4, 1/2]$, a merger therefore increases social surplus if and only if $v \in (1/2, 5/6)$.[33]

The possibility of social surplus increasing mergers is obtained here without restricting the contracting space of the firms pre merger: Setting a market-clearing wage is part of the optimal mechanism for a single-plant monopsony operating at either end of the Hotelling line with uniformly distributed workers, and when the optimal quantity procured is less than 1/2, the mechanism remains optimal even if the other firm is present. Moreover, without a third party the two independent firms cannot replicate the multi-plants monopsony outcome—firm i cannot offer a wage w_i and say this is how much it will pay a worker if the worker is hired by the other firm.

To see how the implementation of the optimal mechanism for the multi-jobs monopsony can be implemented via a third party, suppose that in addition to the two independent firms, there is a labor market intermediary who offers wages of 1/2 to workers who will then be randomly matched to one of the two firms. Each firm pays a fee of 1/2 to the intermediary for each worker that is referred, possibly up to a quantity constraint. In addition, each independent firm $\ell \in \{0, 1\}$ offers a wage of $w_\ell = 1/4$ to workers it hires directly. These wages are mutually best responses given the intermediary's behaviour. The intermediary makes zero profits and each of the independent firms receives half of the multi-jobs monopsony profit. Of course, because the scheme is strictly profitable, the intermediary could charge the firms a fixed payment for its services. For example, with a constant willingness to pay of v per worker with $v \in (1/2, 1)$, each firm earns $v^2/4$ without the intermediary and $v/2 - 3/16$ with the intermediary. Hence, any fixed fee $\phi \in [0, v/2 - 3/16 - v^2/4)$ will be acceptable for the firms since they are still strictly better off with the intermediary and its fee than without it.

6.2 Discussion

The analysis of and discussion related to the Hotelling model above raises the questions of whether the same or similar effects are also present in the model set up in Section 2

[33]Of course, by the results stated in Proposition 9, the merger harms workers.

and analyzed in Section 3. We now address these questions, beginning with the effects of permitting or prohibiting wage discrimination in that model.

Effects of wage discrimination with homogeneous workers Prohibiting wage discrimination in the baseline model of Section 2 means that the firm will optimally procure a quantity Q satisfying $V(Q) = C'(Q)$. Because there C' is not monotone, there can be multiple local maxima. Of course, the monopsony will choose the quantity that corresponds to the global profit maximum, but this quantity may be larger or smaller than the quantity Q^* that the monopsony procures under the optimal mechanism when $\underline{C}(Q^*) < C(Q^*)$. This means that there is, in general, no monotone quantity effect of prohibiting wage discrimination akin to the one in Proposition 9 for the Hotelling model. Of course, just like there, keeping the quantity fixed, allowing the monopsony to wage discriminate can only decrease total wage payments. Even so, workers who are employed at the high wage are better off with wage discrimination than without it, keeping the employment level fixed. Interestingly, because of the possibility of a positive quantity effect and because, in contrast to the Hotelling model, all but a measure zero of workers who are employed in equilibrium get a strictly positive surplus in the baseline model, it is possible that permitting wage discrimination increases worker surplus.[34] For example, for the piecewise linear specification (4) and $V(Q) = v$ for all $Q \leq 1/4$ and $V(Q) = 0$ otherwise with $v \in (C'(Q_1), C'(1/4))$, the global maximum when wage discrimination is prohibited is always given by a quantity smaller than 1/4. When wage discrimination is permitted, the optimal quantity is 1/4. For v sufficiently small, that is, less than 1.65, worker surplus is larger with wage discrimination than without it.[35]

Heterogeneous tasks and endogenous multi-tasking Thus far, we have assumed that the firm only requires the execution of one homogeneous task, with the willingness to pay of the firm for the Q-th unit being $V(Q)$ with $V' < 0$. We now generalize this by assuming that the firm has demand for n different tasks, indexed by i, with a maximal demand for task i of k_i. The firm's marginal willingness to pay for the Q-th unit when that unit is provided by task i is $V(Q)\theta_i$ while the cost of a worker with opportunity cost $W(Q)$ of executing task i is $W(Q)\theta_i$. We assume $\theta_1 > \ldots \theta_n > 0$ and, for $i \in \{1, \ldots, n\}$, let $K_{(i)} = \sum_{j=1}^{i} k_j$ and $\theta_1 = 1$.

[34]When wage discrimination is prohibited, worker surplus when that employment level is Q is $WS(Q) = C(Q) - \int_0^Q W(x)dx$. When wage discrimination is permitted and $Q \in (Q_1(m), Q_2(m))$, worker surplus is $\underline{WS}(Q) = \underline{C}(Q) - \int_0^{Q_1(m)} W(x)dx - \frac{Q-Q_1(m)}{Q_2(m)-Q_1(m)} \int_{Q_1(m)}^{Q_2(m)} W(x)$.

[35]See Loertscher and Muir (2021a, Proposition 5) for a more elaborate analysis of the related problem of facilitating or prohibiting resale on consumer surplus in a monopoly pricing problem in which the optimal selling mechanism involves rationing. The two problems are related because resale as modelled there induces an efficient allocation, which is what occurs here without wage discrimination.

The cost of procuring Q units of labor for task 1 at the market-clearing wage is still given by $C(Q) = W(Q)Q$.

The problem faced by the profit-maximizing firm in the presence of heterogeneous tasks is to choose the total number of workers it wants to employ, how to allocate the tasks across these workers, and how to match the executed tasks to its inverse demand function $V(Q)$. The answer to the last question is simple. Because V is downward sloping, the optimal matching of executed tasks to its demand is positive assortative, that is, denoting by $q_i \leq k_i$ the units of tasks i that are executed, profit is maximized by matching the best available tasks to the highest-value segment of demand, which generates a benefit of

$$\int_0^{q_1} V(x)dx + \theta_2 \int_{q_1}^{q_1+q_2} V(x)dx + \cdots + \theta_h \int_{\sum_{i=1}^{h-1} q_i}^{\sum_{i=1}^{h} q_i} V(x)dx,$$

where $h \leq n+1$ is the least productive task procured. If C is convex, then C' is increasing, which implies that the least costly way of having any collection of task $(q_1, \ldots, q_h)$ with $q_i \leq k_i$ executed is in a similar positive assortative fashion by having the lowest cost workers executing tasks 1, and so on. It is then not hard to see that the total number of workers employed, Q^*, is given by equating marginal benefit and marginal cost, that is $V(Q^*) = C'(Q^*)$, provided, of course, $K_n \geq Q^*$. If $K_n < Q^*$, then it is optimal to employ K_n workers. In either case, every worker executes exactly one task.

Optimal multi-tasking arises in equilibrium only if C is not convex. The optimal procurement mechanism with heterogeneous tasks can be derived by applying the analysis of Loertscher and Muir (2021a) to the procurement setting. First, without loss of generality, we introduce an arbitrarily large mass of job of intensity $\theta_{n+1} = 0$, and for convenience, we set $K_{(0)} = 0$ and $K_{(n+1)} = \infty$. Then we identify the mass of jobs to be allocated within the interval $[0, \infty)$ by sorting them from most (θ_1) to least (θ_n) intensive, so that for $i \in \{1, \ldots, n\}$ the interval $[K_{(i-1)}, K_{(i)}]$ corresponds to the mass of jobs of intensity θ_i. Similar to the case where C is convex, these tasks are then assigned to the mass of Q workers in a positive assortative fashion so that the highest intensity tasks are allocated to the worker with the lowest cost of supplying labor. However, for each ironing interval $m \in \mathcal{M}$ of the function C, the corresponding mass of tasks that fall within the interval $[Q_1^*(m), Q_2^*(m)]$ are not assigned in a positive assortative fashion and are instead randomly assigned to the corresponding mass of workers. Alternatively, we can think of the firm as repackaging the tasks that fall within the interval $[Q_1^*(m), Q_2^*(m)]$ into a mass $Q_2^*(m) - Q_1^*(m)$ of homogeneous jobs and asking the corresponding mass of workers assigned to these jobs to multi-task. This analysis therefore provides an alternative interpretation of multi-tasking in the sense of Holmström and Milgrom (1991). In our setting, it arises from cost minimization by a monopsony with heterogeneous tasks that faces a non-convex procurement cost function.

Given the optimal mechanism for procuring the Q highest-value units of labor, it only remains to determine the precise mass of workers that are hired under the optimal mechanism. However, following Loertscher and Muir (2021a), this argument proceeds in precisely the same manner as for the case where C is convex, after we replace the cost function C with its concavification $\underline{C}$.

6.3 Migration, efficiency wages, and unemployment

A pervasive feature of the wage increase at the Ford Motor Company in 1914 was that it caused workers to migrate to Detroit (see, for example, Sward, 1948, p.53). As we now show, when workers face a fixed cost of moving or participating in the labor market, this can give rise to a procurement cost function C that is non-convex. Consequently, it may be optimal for the firm to use an efficiency wage and to induce involuntary unemployment. This resonates with a popular view that migration is a cause of unemployment in the region to which workers migrate. The subtle twist, however, is that in this model involuntary unemployment occurs not because of frictions such as costly search or costly wage adjustment, but rather it arises as a consequence of optimal pricing on the part of the firm.

Specifically, we now consider a model with a monopsony firm that operates in a market in which, absent immigration, the inverse labor supply function is W_A. We assume that this function is increasing in Q and satisfies, for the sake of the argument, $W_A'' \geq 0$. This implies that absent migration, the corresponding cost $QW_A(Q)$ of procuring Q units of labor is convex in Q. Thus, absent migration, the firm optimally sets a market-clearing wage. To model migration, we assume that there is another pool of workers whose opportunity costs of working after migrating are described by the inverse supply function W_B, which we also assume to be convex and increasing. Each worker in this pool has the same fixed cost $k > 0$ of moving. For $i \in \{A, B\}$, let $S_i(w) = W_i^{-1}(w)$ and, for $w > W_B(0) + k$, let $S_{AB}(w) = S_A(w) + S_B(w - k)$ denote the supply function that the firm faces. Moreover, for $Q > S_A(W_B(0) + k)$, we take $W_{AB}(Q) = S_{AB}^{-1}(Q)$. Then the inverse labor supply function the firm faces is

$$W(Q) = \begin{cases} W_A(Q), & Q \leq S_A(W_B(0) + k) \\ W_{AB}(Q), & Q > S_A(W_B(0) + k) \end{cases}.$$

Accordingly, accounting for migration, the cost of procuring Q units of labor is $C(Q) = QW(Q)$.

Of course, migration involving a fixed cost is only one possible interpretation of this setup. Equivalently, one can think of workers from the outside pool B as currently working in a different industry. These workers can then switch from industry B to industry A if they bear the adjustment cost k. We can also think of the workers in the outside pool B as a group of workers that are not currently employed and who bear a fixed cost of labor market participation, such as finding and paying for child care.

Letting $\hat{Q} = S_A(W_B(0)+k)$ and $\hat{w} = W_B(0)+k$ (which is the same as $W_A(\hat{Q})$), we have

$$\lim_{Q\uparrow\hat{Q}} C'(Q) = W_A(\hat{Q}) + \hat{Q}S_A^{-1'}(\hat{Q}) > W_{AB}(\hat{Q}) + \hat{Q}S_{AB}^{-1'}(\hat{Q}) = \lim_{Q\downarrow\hat{Q}} C'(Q).$$

Here, the inequality follows because $W_A(\hat{Q}) = W_{AB}(\hat{Q}) = \hat{w}$ and, for $w \geq \hat{w}$, $S_{AB}(w) = S_A(w) + S_B(w-k)$ implies that $S'_{AB}(w) = S'_A(w) + S'_B(w-k) > S'_A(w)$, which in turn implies $S_{AB}^{-1'}(\hat{Q}) = \frac{1}{S'_{AB}(\hat{w})} < \frac{1}{S'_A(\hat{w})} = S_A^{-1'}(\hat{Q})$. Thus, the function C is not convex.

The key implication of all of this is that an increase in the demand for labor by the monopsony in market A—that in equilibrium induces migration to A—can also induce an efficiency wage and involuntary unemployment. As an illustration, for $W_A(Q) = 4Q$, $W_B(Q) = \frac{4}{7}Q + \frac{1}{2}$ and $k = 1/2$, we obtain the specification in (2) and (4).[36] Parameterizing the firm's marginal benefit function as $V(Q) = a - Q$, for a sufficiently small, the intersection of $V(Q)$ with $\underline{C}(Q)$ occurs at a point $Q^*(a)$ such that $Q^*(a) < Q_1$, implying that $\underline{C}(Q^*(a)) = C(Q^*(a))$. Consequently, a market-clearing wage is optimal, and there is no unemployment. As a increases, $Q^*(a)$ increases and eventually $Q^*(a)$ is larger than Q_1 and smaller than Q_2. In this case, an efficiency wage is optimal and involuntary unemployment is part of the optimal procurement mechanism, which now induces migration.[37] Finally, if the firm's labor demand keeps increasing, $Q^*(a)$ will eventually exceed Q_2 and a market-clearing wage will become optimal.

This perspective also offers a novel interpretation of the episode at the Ford Motor Company in the mid 1910s. Contrary to perceived wisdom, Ford did not introduce a uniform wage of \$5 per day in 1914. Until 1916, 30 percent of its workforce earned less than that per day (Sward, 1948). With high enough wages, workers are willing to bear the fixed cost of moving, making C non-convex in the short run and efficiency wages optimal: "the greatest cost saving" (Henry Ford). As the demand for its cars and its demand for labor continued increasing, eventually it became optimal to set market-clearing wages again.

[36] To see this, note that $W_A(Q) = 4Q$ and $W_B(Q) = 4Q/7 + 1/2$ imply $S_A(w) = w/4$ and $S_B(w) = 7(w-1/2)/4$ and hence using $k = 1/2$ for $w \geq \hat{w}$ we have $S_{AB}(w) = S_A(w) + S_B(w-k) = 2w - 7/4$. Inverting S_{AB} yields $W_{AB}(Q) = Q/2 + 7/8$, which is the second line in (2). It remains to verify that $\hat{Q} = 1/4$, which is the case since $S_A(W_B(0)+k) = (1/2+1/2)/4 = 1/4$.

[37] Observe though that while there is unemployment, employment has also increased.

7 Conclusions

Minimum wage legislation is at the forefront of public policy debates. We provide a model in which an appropriately chosen minimum wage increases total employment and decreases involuntary unemployment, possibly to the point of eliminating it. The model merely assumes that a monopsony firm minimizes the cost of procuring labor, subject to respecting workers' incentive compatibility and individual rationality constraints, and that the procurement cost under a market-clearing wage is not convex at the optimal level of employment. Extending the model to allow for quantity competition among firms, we show that there is no monotone relationship between competition and involuntary unemployment. The latter point is perhaps most starkly illustrated by the fact that it is possible to have involuntary unemployment and inefficient allocation under perfect competition.

In the mechanism design approach taken in this paper, randomization in the form of efficiency wages and involuntary unemployment (or mismatching of workers and jobs) occurs because it is optimal for the employer, and not as the result of search or other frictions. An interesting and relevant avenue for further research would therefore be to assess the empirical magnitude of these different causes of randomization. The policy implications differ substantively. If these inefficiencies are caused by frictions, then reducing the frictions will typically improve welfare. If they are by design, then trying to reduce, say, the randomness in workers-job matching may be a cat and mouse game.

References

Akbarpour, M., P. Dworczak, and S. D. Kominers (2020): "Redistributive Allocation Mechanisms," Working paper.

Akerlof, G. and J. Yellen (1986): *Efficiency Wage Models of the Labor Market*, Princeton, New Jersey: Cambridge: Cambridge University Press.

Balestrieri, F., S. Izmalkov, and J. Leao (2021): "The Market for Surprises: Selling Substitute Goods through Lotteries," *Journal of the European Economic Association*, 1, 509–535.

Bhaskar, V., A. Manning, and T. To (2002): "Oligopsony and Monopsonistic Competition in Labor Markets," *Journal of Economic Perspectives*, 16, 155 – 174.

Börgers, T. (2015): *An Introduction to the Theory of Mechanism Design*, New York: Oxford University Press.

Bulow, J. and J. Roberts (1989): "The Simple Economics of Optimal Auctions," *Journal of Political Economy*, 97, 1060–1090.

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

Condorelli, D. (2012): "What Money Can't Buy: Efficient Mechanism Design with Costly Signals," *Games and Economic Behavior*, 75, 613–624.

Cournot, A. (1838): *Rechereches sur les Principles Mathématiques de la Théorie des Richesses*, Paris.

Cullen, Z. B. and B. Pakzad-Hurson (2021): "Equilibrium effects of pay transparency," NBER Working Paper # 28903.

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

Engels, F. (1845): *Die Lage der arbeitenden Klasse in England*, Leipzig: Verlag Otto Wigand.

Holmström, B. and P. Milgrom (1991): "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, & Organization*, 7, 24–52.

Hotelling, H. (1931): "The Economics of Exhaustible Resources," *Journal of Political Economy*, 39, 137–175.

Krueger, A. B. and L. H. Summers (1988): "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica*, 56, 259–293.

Loertscher, S. and E. V. Muir (2021a): "Monopoly pricing, optimal randomization and resale," *Journal of Political Economy*, forthcoming.

——— (2021b): "Optimal Hotelling auctions," Work-in-progress.

Loertscher, S. and C. Wasser (2019): "Optimal Structure and Dissolution of Partnerships," *Theoretical Economics*, 14, 1063–1114.

Marx, K. (1867): *Das Kapital: Kritik der politischen Oekonomie*, Hamburg: Verlag Otto Meissner.

Mussa, M. and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–317.

Myerson, R. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–78.

STARR, G. (1981): *Minimum Wage Fixing: An International Review of Practices and Problems*, Geneva: International Labour Organization.

STIGLER, G. J. (1946): "The Economics of Minimum Wage Legislation," *American Economic Review*, 36, 358 – 365.

SWARD, K. (1948): *The Legend of Henry Ford*, New York: New York: Rinehart.

YELLEN, J. L. (1984): "Efficiency Wage Models of Unemployment," *American Economic Review*, 74, 200–205.

Appendix

A Proofs

A.1 Proof of Lemma 2

Proof. Part I: Proof that the minimal cost is $\underline{C}(Q, \underline{w})$

To prove the first part of the statement, we adopt a mechanism design approach. The designer's problem is to determine the cost-minimizing mechanism for procuring a fixed quantity Q of labor, subject to the constraint that the menu of wages does not include a wage below the minimum wage of $\underline{w}$. We let G denote the cumulative distribution of the opportunity cost of working for the population of workers. We denote the support of this distribution by $[\underline{c}, \overline{c}] \subseteq \mathbb{R}_+$ and its density by g. The assumptions introduced in Section 2 ensure that the function G admits a density g and that g is strictly positive on $[\underline{c}, \overline{c}]$. We let $\Gamma(c) := c + \frac{G(c)}{g(c)}$ the virtual cost function. The value of the worker's outside option of not participating in the mechanism is 0.

We let $\langle x, t \rangle$ represent an arbitrary direct mechanism, with $x(c)$ denoting the probability that the worker has to work when reporting to be of type c and $t(c)$ denoting the expected transfer the worker receives if reporting to be of type c. Note that within loss of generality, we can assume that workers are paid a transfer upon becoming employed. Consequently, the worker's payoff when of type c and reporting to be of type $\hat{c}$ takes the form

$$t(\hat{c}) - x(\hat{c})c.$$

Let $U(c) := t(c) - x(c)c$ denote the worker's payoff when reporting truthfully. Individual rationality requires $U(c) \geq 0$ for all c. Incentive compatibility implies that $x(c)$ is non-increasing and that $U'(c) = -x(c)$ holds almost everywhere. For any $c, \hat{c} \in [\underline{c}, \overline{c}]$ we then have

$$U(c) = U(\hat{c}) + \int_c^{\hat{c}} x(y)dy.$$

Setting this equal to $t(c) - x(c)c$ and solving for $t(c)$ gives

$$t(c) = U(\hat{c}) + x(c)c + \int_c^{\hat{c}} x(y)dy.$$

Observing that for $c < \hat{c}$, $U(c) \geq U(\hat{c})$ holds because $\int_c^{\hat{c}} x(y)dy \geq 0$, the individual rationality constraint is satisfied for all types if and only if $U(\overline{c}) \geq 0$. In an optimal mechanism satisfying incentive compatibility and individual rationality, we must have $U(\bar{c}) = 0$ because otherwise the designer leaves money on the table. Expressing $t(c)$ with $\hat{c} = \bar{c}$ and using $U(\bar{c}) = 0$, we thus obtain

$$t(c) = x(c)c + \int_c^{\bar{c}} x(y)dy.$$

The designer's procurement cost minimization problem, subject to the minimum wage constraint parameterized by $\underline{w}$, is given by

$$\min_x \int_{\underline{c}}^{\bar{c}} t(c)\, dG(c)$$

$$\text{s.t. } x \text{ is non-increasing,} \qquad \int_{\underline{c}}^{\bar{c}} x(c)\, dG(c) = Q, \quad \underline{w}x(c) \le t(c) \text{ for all } c \in [\underline{c}, \bar{c}].$$

Note that $t(c)$ is the expected transfer paid to workers of type c. In line with real-world practice, we assume that the minimum wage $\underline{w}$ represents a constraint on the wage payments made to hired workers. Since workers of type c are hired with probability $x(c)$ and workers are only paid a wage upon being hired, the constraints that the minimum wage $\underline{w}$ imposes on the transfers $t(c)$ are given by $\underline{w}x(c) \le t(c)$.

We have a continuum of constraints given by $\underline{w}x(c) \le t(c)$ for all $c \in [\underline{c}, \bar{c}]$. Under ex post individual rationality (EIR), no worker can ever be paid a wage w that is less than its opportunity cost. This means that for worker types with costs $c > \underline{w}$, the constraint never binds under EIR.

Using the fact that the constraint $\underline{w}x(c) \le t(c)$ is equivalent to $h(c) := \underline{w}x(c) - t(c) \le 0$, we next show that $h(c)$ decreases in c on $[\underline{c}, \underline{w}]$. Specifically, letting $c_0, c_1 \in [\underline{c}, \underline{w}]$ with $v_0 < v_1$, we have

$$\begin{aligned} h(c_1) - h(c_0) &= \underline{w}(x(c_1) - x(c_0)) - (x(c_1)c_1 - x(c_0)c_0) + \int_{c_0}^{c_1} x(y)dy \\ &= (\underline{w} - c_1)(x(c_1) - x(c_0)) + \int_{c_0}^{c_1} x(y)dy - (c_1 - c_0)x(c_0) \le 0, \end{aligned}$$

where the inequality is strict if x is not constant on $[c_0, c_1]$.[38] This shows that it suffices to impose the constraint associated with the minimum wage on the lowest type $c = \underline{c}$.

We let λ denote the Lagrange multiplier corresponding to the lowest type $c = \underline{c}$. Setting aside the quantity constraint for now and using $t(c) = x(c)c + \int_c^{\bar{c}} x(y)\, dy$, the Lagrangian is

[38] Since x is non-increasing, if x is not constant on $[c_0, c_1]$ we have $(c_1 - c_0)x(c_0) > \int_{c_0}^{c_1} x(y)dy$ and $(\underline{w} - c_1)(x(c_1) - x(c_0)) \le 0$ with strict inequality if $c_1 < \underline{w}$.

then given by

$$\begin{aligned}\mathcal{L}(x,\lambda) &= \int_{\underline{c}}^{\overline{c}} t(c)\, dG(c) + \lambda(\underline{w}x(\underline{c}) - t(\underline{c})) \\ &= \int_{\underline{c}}^{\overline{c}} \left(x(c)c + \int_{c}^{\overline{c}} x(y)\, dy \right) dF(v) + \lambda x(\underline{c})(\underline{w} - \underline{c}) - \lambda \int_{\underline{c}}^{\overline{c}} x(c)\, dc.\end{aligned}$$

Using

$$\int_{\underline{c}}^{\overline{c}} \int_{c}^{\overline{c}} g(c)x(y)\, dy\, dc = \int_{\underline{c}}^{\overline{c}} \int_{\underline{c}}^{y} g(c)x(y)\, dc\, dy = \int_{\underline{c}}^{\overline{c}} G(y)x(y)\, dy.$$

we have

$$\begin{aligned}\mathcal{L}(x,\lambda) &= \int_{\underline{c}}^{\overline{c}} \Gamma(c)x(c)\, dG(c) + \lambda x(\underline{c})(\underline{w} - \underline{c}) - \lambda \int_{\underline{c}}^{\overline{c}} x(c)\, dc \\ &= \int_{\underline{c}}^{\overline{c}} \left(\Gamma(c) - \frac{\lambda}{g(c)} \right) x(c)\, dG(c) + \lambda x(\underline{c})(\underline{w} - \underline{c}).\end{aligned}$$

Letting $H(x) = \mathbf{1}(x \geq 0)$ denote the Heaviside step function and using the the probability measure $G_\lambda(c) = \frac{\lambda}{1+\lambda} H(\underline{c} - c) + \frac{1}{1+\lambda} G(c)$, we can rewrite the Lagrangian as

$$\mathcal{L}(x,\lambda) = (1+\lambda) \int_{\underline{c}}^{\overline{c}} \left[\left(\Gamma(c) - \frac{\lambda}{g(c)} \right) \mathbf{1}(c > \underline{c}) + (\underline{w} - \underline{c})\, \mathbf{1}\, (c = \underline{c}) \right] x(v)\, dG_\lambda(c).$$

This is intuitive as $\lambda > 0$ makes the designer favor agents of types $c > \underline{c}$ relative to the lowest types (whose marginal cost is no longer $\underline{c}$ but rather $\underline{w}$). Moreover, increasing $\underline{w}$ harms the designer since it increases $\underline{w} - \underline{c}$ and thereby increases the Lagrangian, which is bad since we aim to minimize expenditure.

We can therefore derive the optimal allocation rule x^* by ironing the function

$$\Psi(c,\lambda) = \begin{cases} \Gamma(c) - \frac{\lambda}{g(c)}, & c \in (\underline{c}, \overline{c}] \\ \underline{w} - \underline{c}, & c = \underline{c} \end{cases}$$

with respect to the probability measure G_λ. Note that if the function Ψ discontinuously decreases at $x = \underline{c}$ this implies that the ironed function $\overline{\Psi}$ contains an ironing interval with an endpoint at $c = \underline{c}$. This ironing interval precisely corresponds to the region identified in Section 4.2 where the optimal mechanism does not involves rationing at the minimum wage $\underline{w}$. Any additional ironing regions correspond to two-price mechanisms with no randomization at the top and rationing at the minimum wage $\underline{w}$. This shows that our restriction to two-price mechanisms when $Q \in (S(\underline{w}), w_1^{-1}(\underline{w}; m))$ is without loss of generality as required.

Part II: Proof of the stated properties of $\underline{C}(Q, \underline{w})$

Since $\underline{w} \in (Q_1(m), Q_2(m)]$, m is fixed, and so for this proof, we omit the dependence of $Q_i(m)$, α_m, $w_1^{-1}(\underline{w}; m)$ and $w_1(Q; m)$ on m and simply write Q_i, α, $\hat{Q}$ and $w_1(Q)$, where $\alpha = \frac{Q-q_1}{q_2-q_1}$.

The Lagrangian for the monopsony's problem

$$\min_{q_1 \in [0,Q], q_2 \geq Q} (1-\alpha)C(q_1) + \alpha C(q_2)$$

subject to the constraint $(1-\alpha)W(q_1) + \alpha W(q_2) \geq \underline{w}$ is

$$\mathcal{L}(q_1, q_2, \lambda) = (1-\alpha)C(q_1) + \alpha C(q_2) - \lambda[(1-\alpha)W(q_1) + \alpha W(q_2) - \underline{w}],$$

where for $Q \in (S(\underline{w}), w_1^{-1}(\underline{w}))$ the constraint will bind, i.e. hold with equality at an optimum, for otherwise the solution would be $q_i = Q_i$ inducing a low wage of $w_1(Q) < \underline{w}$, thereby violating the minimum wage constraint.

Using $C_\lambda(Q) := W(Q)(Q-\lambda)$, the Lagrangian can equivalently be written as

$$\mathcal{L}(q_1, q_2, \lambda) = (1-\alpha)C_\lambda(q_1) + \alpha C_\lambda(q_2) + \lambda \underline{w}.$$

Using the facts that

$$\frac{\partial \alpha}{\partial q_1} = -\frac{1-\alpha}{q_2-q_1} \quad \text{and} \quad \frac{\partial \alpha}{\partial q_2} = -\frac{\alpha}{q_2-q_1},$$

the first-order conditions with respect to q_1 and q_2 are those captured in

$$C'_\lambda(q_1) = \frac{C_\lambda(q_2) - C_\lambda(q_1)}{q_2 - q_1} = C'_\lambda(q_2) \tag{11}$$

while the first-order condition with respect to λ is and

$$(1-\alpha)W(q_1) + \alpha W(q_2) = \underline{w}. \tag{12}$$

Letting

$$H(q_2, q_1, \lambda) = \frac{C_\lambda(q_2) - C_\lambda(q_1)}{q_2 - q_1}.$$

and using subscripts to denote partial derivatives, we have

$$H_1(q_2, q_1, \lambda) = \frac{1}{q_2 - q_1}\left[C'_\lambda(q_2) - H(q_2, q_1, \lambda)\right]$$
$$H_2(q_2, q_1, \lambda) = \frac{1}{q_2 - q_1}\left[H(q_2, q_1, \lambda) - C'_\lambda(q_1)\right]$$
$$H_3(q_2, q_1, \lambda) = \frac{W(q_1) - W(q_2)}{q_2 - q_1}.$$

Note that $H_3 < 0$ because by assumption we have $q_2 > q_1$ and W is an increasing function.

Observe also that (11) is equivalent to

$$C'_\lambda(q_1) = H(q_2, q_1, \lambda) = C'_\lambda(q_2). \tag{13}$$

Denote by $q_1^*(\lambda)$ and $q_2^*(\lambda)$ the values of q_1 and q_2 that satisfy (13). Evaluated at these values, we have

$$H_1(q_2^*(\lambda), q_1^*(\lambda), \lambda) = 0 = H_2(q_2^*(\lambda), q_1^*(\lambda), \lambda).$$

This implies that the second partials of $\mathcal{L}(q_1, q_2, \lambda)$ with respect to q_1 and q_2, evaluated at $q_i = q_i^*$ are

$$\frac{\partial^2 \mathcal{L}(q_1^*, q_2^*, \lambda^*)}{\partial q_1^2} = (1-\alpha)C''_\lambda(q_1^*) \quad \text{and} \quad \frac{\partial^2 \mathcal{L}(q_1^*, q_2^*, \lambda^*)}{\partial q_2^2} = \alpha C''_\lambda(q_2^*)$$

and

$$\frac{\partial^2 \mathcal{L}(q_1^*, q_2^*, \lambda^*)}{\partial q_1 \partial q_2} = 0.$$

The matrix of second partial is thus

$$\begin{pmatrix} (1-\alpha)C''_\lambda(q_1^*) & 0 \\ 0 & \alpha C''_\lambda(q_2^*) \end{pmatrix}.$$

This is positive definite if and only if $(1-\alpha)C''_\lambda(q_1^*) > 0$ and $\alpha C''_\lambda(q_2^*) > 0$. Thus, at the optimum, we have for each $i \in \{1, 2\}$,

$$C''_\lambda(q_i^*) > 0.$$

Totally differentiating $C'_\lambda(q_i^*) = H(q_2^*, q_1^*, \lambda)$ with respect to λ and using $H_1(q_2^*, q_1^*, \lambda) = 0 = H_2(q_2^*, q_1^*, \lambda)$ yields

$$\frac{dq_i^*}{d\lambda} = \frac{H_3(q_2^*, q_1^*, \lambda) + W'(q_i^*)}{C''_\lambda(q_i^*)}.$$

Because $C''_\lambda(q_i^*) > 0$, it follows that $\frac{dq_i^*}{d\lambda}$ has the same sign as

$$H_3(q_2^*, q_1^*, \lambda) + W'(q_i^*) = \frac{W(q_1^*) - W(q_2^*)}{q_2^* - q_1^*} + W'(q_i^*).$$

We next show that this expression is positive for $i = 1$ and negative for $i = 2$.

To this end, we first notice that for$q_1^* < q_2^*$ and $Q \in (q_1^*, q_2^*)$, $C_\lambda(Q)$ is not convex, that is, for all $Q \in (q_1^*, q_2^*)$ we have $\underline{C}_\lambda(Q) < C_\lambda(Q)$, because otherwise there would be no need to convexify $C_\lambda(Q)$. We now show that this implies that $W(Q)$ is not convex on $[q_1^*, q_2^*]$ by showing that convexity of W implies convexity of C_λ.

To see this, for $a \in [0, 1]$ and x_0 and x_1 satisfying $q_1^* \leq x_0 < x_1 \leq q_2^*$, let $x^a := ax_0 + (1-a)x_1$. Convexity of W on $[q_1^*, q_2^*]$ means that

$$W(x^a) \leq aW(x_0) + (1-a)W(x_1).$$

Now by definition of C_λ, we have $C_\lambda(x^a) = C(x^a)(x^a - \lambda)$. By convexity of W this gives us the first inequality in the following:

$$\begin{aligned}
C_\lambda(x^a) &\leq (aW(x_0) + (1-a)W(x_1))\,(ax_0 + (1-a)x_1 - \lambda) \\
&= (aW(x_0) + (1-a)W(x_1))\,(a(x_0 - \lambda) + (1-a)(x_1 - \lambda) \\
&= a\,(aW(x_0) + (1-a)W(x_1))\,(x_0 - \lambda) + (1-a)\,(aW(x_0) + (1-a)W(x_1))\,(x_1 - \lambda)) \\
&= aW(x_0)(x_0 - \lambda) + (1-a)W(x_1)(x_1 - \lambda) + a(1-a)(W(x_1) - W(x_0))(x_0 - x_1) \\
&= aC_\lambda(x_0) + (1-a)C_\lambda(x_1) + a(1-a)(W(x_1) - W(x_0))(x_0 - x_1) \\
&\leq aC_\lambda(x_0) + (1-a)C_\lambda(x_1),
\end{aligned}$$

where the second inequality follows because $W(x_1) - W(x_0) > 0$ and $x_0 - x_1 < 0$ (which also implies that the inequality is strict if $a \in (0, 1)$.) Thus, C_λ is convex if W is convex. Because C_λ is not convex on $[q_1^*, q_2^*]$, this implies that $W(Q)$ is not convex on $[q_1^*, q_2^*]$, that is, for all $Q \in (q_1^*, q_2^*)$,

$$W(Q) > W(q_1^*) + (Q - q_1^*)\frac{W(q_2^*) - W(q_1^*)}{q_2^* - q_1^*}.$$

Finally, because $W(Q)$ intersects with the linear function $W(q_1^*) + (Q - q_1^*)\frac{W(q_2^*) - W(q_1^*)}{q_2^* - q_1^*}$ at $Q = q_2^*$ from above, it follows that the slope of W at that point is smaller than $\frac{W(q_2^*) - W(q_1^*)}{q_2^* - q_1^*}$, that is $W'(q_2^*) < \frac{W(q_2^*)-W(q_1^*)}{q_2^*-q_1^*}$, which is equivalent to

$$\frac{W(q_1^*)-W(q_2^*)}{q_2^*-q_1^*}+W'(q_2^*)<0,$$

which is the same as $H_3(q_2^*, q_1^*, \lambda) + W'(q_2^*) < 0$, which implies

$$\frac{\partial q_2^*(\lambda)}{d\lambda}<0.$$

By the same token, $W(Q)$ intersects with the linear function $W(q_1^*) + (Q - q_1^*)\frac{W(q_2^*)-W(q_1^*)}{q_2^*-q_1^*}$ at $Q = q_1^*$ from below, implying $W(q_1^*) + (q_2^* - q_1^*)W'(q_1^*) > W(q_2^*)$, which is equivalent to

$$\frac{W(q_1^*)-W(q_2^*)}{q_2^*-q_1^*}+W'(q_1^*)>0,$$

which is the same as $H_3(q_2^*, q_1^*, \lambda) + W'(q_1^*) > 0$, implying

$$\frac{\partial q_1^*(\lambda)}{d\lambda}>0.$$

Once we have established the comparative static properties of the solution value $\lambda^*(Q, \underline{w})$ with respect to Q and $\underline{w}$, the comparatives static properties of $q_i^*(Q, \underline{w})$ with respect to these parameters will follow from the definition of $q_i^*(Q, \underline{w})$ via $q_i^*(Q, \underline{w}) = q_i^*(\lambda^*(Q, \underline{w}))$ and the facts $\frac{\partial q_1^*(\lambda)}{d\lambda} > 0 > \frac{\partial q_2^*(\lambda)}{d\lambda}$. Totally differentiating (12) $(1 - \alpha^*)W(q_1^*) + \alpha^* W(q_2^*) = \underline{w}$ with respect to $\underline{w}$, where $\alpha^* = \frac{Q-q_1^*}{q_2^*-q_1^*}$ and where we have dropped dependence on λ^* for notational ease, yields

$$\left\{(1-\alpha^*)\frac{dq_1^*}{d\lambda}\left(W'(q_1^*(\lambda))+H_3\right)+\alpha^*\frac{dq_2^*}{d\lambda}\left(W'(q_2^*(\lambda))+H_3\right)\right\}\frac{d\lambda^*}{d\underline{w}}=1.$$

Thus, $\frac{d\lambda^*}{d\underline{w}}$ is positive if the term in brackets is positive, which is the case if both summands are positive. To see that the second summand is positive, recall $\frac{dq_2^*}{d\lambda} < 0$ and $W'(q_2^*(\lambda)) + \frac{W(q_1^*)-W(q_2^*)}{q_2^*-q_1^*} < 0$, and to see that the first summand is positive, it suffice to recall that $\frac{dq_1^*}{d\lambda} > 0$ and that $W'(q_1^*) + \frac{W(q_1^*)-W(q_2^*)}{q_2^*-q_1^*} > 0$. Because $\frac{dq_i^*(Q,\underline{w})}{d\underline{w}} = \frac{dq_i^*(\lambda)}{d\lambda}\frac{d\lambda^*(Q,\underline{w})}{d\underline{w}}$, it follows that

$$\frac{dq_1^*(Q,\underline{w})}{d\underline{w}}>0>\frac{dq_2^*(Q,\underline{w})}{d\underline{w}}.$$

Similarly, totally differentiating $(1-\alpha^*)W(q_1^*) + \alpha^* W(q_2^*) = \underline{w}$ with respect to Q yields

$$\left\{(1-\alpha^*)\frac{dq_1^*}{d\lambda}\left(W'(q_1^*(\lambda)) + H_3\right) + \alpha^* \frac{dq_2^*}{d\lambda}\left(W'(q_2^*(\lambda)) + H_3\right)\right\}\frac{d\lambda^*}{dQ} = H_3.$$

Since the right-hand side is negative and the term in brackets on the lefthand side is, as just shown, positive, it follows that $\frac{d\lambda^*}{dQ} < 0$, implying

$$\frac{dq_1^*(Q,\underline{w})}{dQ} < 0 < \frac{dq_2^*(Q,\underline{w})}{dQ}.$$

It is useful to note that

$$\frac{d\lambda^*}{dQ} = H_3 \frac{d\lambda^*}{d\underline{w}}.$$

We next show that $\lambda^*(Q,\underline{w}) \downarrow 0$ as $Q \uparrow w_1^{-1}(\underline{w})$. To see this, notice that $q_i^*(0) = Q_i$, in which case (12) is satisfied if $\underline{w} = w_1(Q)$. Hence, $q_i^*(Q,\underline{w}) \to Q_i$ as $Q \to w_1^{-1}(\underline{w})$ follows.

We are left to establish the stated properties of $\mathcal{L}^*(Q,\underline{w})$. By construction, we have

$$\mathcal{L}^*(Q,\underline{w}) = (1-\alpha^*)C_{\lambda^*}(q_1^*) + \alpha^* C_{\lambda^*}(q_2^*) + \lambda^* \underline{w},$$

where $\lambda^* = \lambda^*(Q,\underline{w})$, $q_i^* = q_i^*(Q,\underline{w})$ and $\alpha^* = \frac{Q-q_1^*}{q_2^*-q_1^*}$. Because $\lambda^* = 0$ at $\underline{w} = w_1(Q)$ (and equivalently at $Q = w_1^{-1}(\underline{w})$), $\mathcal{L}^*(Q,\underline{w}) = \underline{C}(Q)$ at $\underline{w} = w_1(Q)$ follows. Likewise, since $\lambda^* = Q$ at $\underline{w} = W(Q)$ implies $C_{\lambda^*}(Q) = 0$ and $\mathcal{L}^*(Q,\underline{w}) = Q\underline{w}$.

By the envelope theorem, we have for $\underline{w} > w_1(Q)$ and $Q < \hat{Q}(\underline{w})$, respectively,

$$\frac{\partial \mathcal{L}^*(Q,\underline{w})}{\partial \underline{w}} = \lambda^* > 0 \quad \text{and} \quad \frac{\partial \mathcal{L}^*(Q,\underline{w})}{\partial Q} = H(q_2^*, q_1^*, \lambda^*) > 0,$$

establishing the required monotonicity properties. Finally, taking the derivative with respect to Q once more yields

$$\frac{\partial^2 \mathcal{L}^*(Q,\underline{w})}{\partial \underline{w} \partial Q} = \frac{\partial \lambda^*}{\partial Q} < 0 \quad \text{and} \quad \frac{\partial^2 \mathcal{L}^*(Q,\underline{w})}{\partial Q^2} = H_3(q_2^*, q_1^*, \lambda^*)\frac{\partial \lambda^*}{\partial Q} > 0,$$

where the first inequality follows because $\frac{\partial \lambda^*}{\partial Q} < 0$ and the second because $\frac{\partial \lambda^*}{\partial Q} < 0$ and because $H_3 < 0$. (To see that $\frac{\partial^2 \mathcal{L}^*(Q,\underline{w})}{\partial Q^2} = H_3(q_2^*, q_1^*, \lambda^*)\frac{\partial \lambda^*}{\partial Q}$ is correct, recall that $H_1 = H_2 = 0$.) Thus, $\mathcal{L}^*(Q,\underline{w})$ is convex in Q, and increases in $\underline{w}$ decrease the marginal cost of procurement $\frac{\partial \mathcal{L}^*(Q,\underline{w})}{\partial Q}$. □

A.2 Proof of Lemma 3

Proof. The fact that $Q^*(\underline{w})$ satisfies $V(Q^*(\underline{w})) = \underline{C}'(Q^*(\underline{w}), \underline{w})$, provided a Q such that $V(Q) = \underline{C}'(Q, \underline{w})$ holds exists, follows from the facts that $\underline{C}(Q, \underline{w})$ is the minimal cost of procuring the quantity Q and that this cost is convex in Q while $V(Q)$ is the marginal benefit of procuring the quantity Q. When no Q exists such that $V(Q) = \underline{C}'(Q, \underline{w})$ holds, the optimal quantity procured is $S(\underline{w})$ because the marginal cost of procuring more, $\lim_{Q \downarrow S(\underline{w})} \underline{C}'(Q, \underline{w})$, is larger than $V(S(\underline{w}))$ in this case.

The optimal procurement mechanism involves wage dispersion if and only if $Q^*(\underline{w}) > S(\underline{w})$ because for $Q \leq S(\underline{w})$, $\underline{C}(Q, \underline{w}) = \underline{w}Q$, which is achieved by procuring the Q workers at the minimum wage $\underline{w}$.

Whenever there is wage dispersion, the optimal mechanism involves involuntary unemployment. Similarly, when $Q^*(\underline{w}) < S(\underline{w})$, there is excess supply at the minimum wage, and hence involuntary unemployment. $\square$

A.3 Proof of Lemma 4

Proof. For $\underline{w} \in (W(Q_1(m)), W(Q_2(m))$ and $Q \in (S(\underline{w}), w_1^{-1}(\underline{w}; m))$, $\partial \underline{C}(Q, \underline{w})/\partial Q$ is continuous in Q and $\underline{w}$. Hence, $\gamma(Q; m)$ is continuous.

Next we show that

$$\gamma(Q_1(m); m) = C'(Q_2(m)) = \gamma(Q_2(m); m).$$

To see this, notice that at $Q = Q_1(m)$, the constraint $\underline{w} = W(Q_1(m))$ does not bind. Thus, $\gamma(Q_1(m); m) = C'(Q_2(m))$. At $Q = Q_2(m)$, we have for $\underline{w} = W(Q_2(m))$. $\underline{C}(Q, \underline{w}) = \underline{w}Q$ for all $Q \leq Q_2(m)$ and $\underline{C}(Q, \underline{w}) = \underline{C}(Q) = C(Q)$ for $Q \in (Q_2(m), Q_2(m) + \delta)$, where $\delta > 0$, implying $\underline{C}'(Q, \underline{w}) = C'(Q)$ for all $Q \in (Q_2(m), Q_2(m) + \delta)$

From Lemma 2, we know that $\underline{C}'(Q, \underline{w})$ is increasing in Q for $Q \in (S(\underline{w}), w_1^{-1}(\underline{w}; m))$. This implies $\gamma(Q; m) < \underline{C}'(Q_2(m))$ for $Q \in (Q_1(m), Q_2(m))$. Because $V(Q)$ is continuous and decreasing and $\gamma(Q; m)$ is continuous and less than $\underline{C}'(Q_2(m))$ for $Q \in (Q_1(m), Q_2(m))$ and equal to $\underline{C}'(Q_2(m))$ for $Q = Q_2(m)$ while $V(Q^*) = \underline{C}'(Q_2(m))$ with $Q^* \in (Q_1(m), Q_2(m))$ by the hypothesis of the lemma, it follows that a smallest and a largest point of intersection of V and γ on $(Q_1(m), Q_2(m))$ exist and that the smallest point if intersection is strictly larger than Q^*. This establishes $Q^* < \hat{Q}_L(m)$.

Because $\underline{C}(Q, \underline{w})$ is convex and $\underline{C}(Q, \underline{w}) = \underline{w}Q$ for $Q \leq S(\underline{w})$, we have $\underline{C}'(Q, \underline{w}) \geq \underline{w}$ for $Q > S(\underline{w})$. This implies $\gamma(Q; m) \geq W(Q)$, which in turn implied $\hat{Q}_H(m) \leq Q^p$, with equality if and only if $\gamma(Q^p; m) = W(Q^p)$. $\square$

A.4 Proof of Proposition 3

Proof. By construction, for $\underline{w} \in (w_1(Q^*; m), W(\hat{Q}_L(m)))$, the point of intersection between $V(Q)$ and $\underline{C}'(Q, \underline{w})$, $Q^*(\underline{w})$, is larger than $S(\underline{w})$. By Lemma 3, this implies that there is wage dispersion and involuntary unemployment, which establishes (i).

We are left to prove (ii).

That the equilibrium quantity increases follows from the fact that $V(Q)$ is downward sloping in Q and that marginal cost of procurement is decreasing in $\underline{w}$ stated in Lemma 2. Formally, $Q^*(\underline{w})$ satisfies

$$V'(Q^*(\underline{w})) = H(q_2^*, q_1^*, \lambda^*),$$

where $H(q_2^*, q_1^*, \lambda^*)$ is the marginal cost of procurement derived in the proof of Lemma 2. Totally differentiating yields

$$\frac{dQ^*(\underline{w})}{d\underline{w}} = \frac{H_3}{V' - H_3 \frac{\partial \lambda^*}{\partial Q}} \frac{\partial \lambda^*}{\partial \underline{w}} > 0,$$

where the inequality holds because $V' < 0$, $H_3 < 0$, $\frac{d\lambda^*}{dQ} < 0 < \frac{d\lambda^*}{d\underline{w}}$.

The lower of the two wages paid in equilibrium is $\underline{w}$, which trivially increases in $\underline{w}$. We are going to show that the higher of the two wages decreases in $\underline{w}$ by showing that $q_2^*(Q^*(\underline{w}), \underline{w})$ decreases in $\underline{w}$.

Using the definition of $q_2^*(Q, \underline{w}) = q_2^*(\lambda^*(Q, \underline{w}))$ and totally differentiation $q_2^*(\lambda^*(Q^*(\underline{w}, \underline{w}))$ with respect to $\underline{w}$ yields the first line in:

$$\begin{aligned} \frac{dq_2^*(\lambda^*(Q^*(\underline{w}), \underline{w}))}{d\underline{w}} &= \frac{\partial q_2^*}{\partial \lambda} \left[\frac{\partial \lambda^*}{\partial Q} \frac{\partial Q^*(\underline{w})}{\partial \underline{w}} + \frac{\partial \lambda^*}{\partial \underline{w}} \right] \\ &= \frac{\partial q_2^*}{\partial \lambda} \frac{\partial \lambda^*}{\partial \underline{w}} \left[H_3 \frac{\partial Q^*(\underline{w})}{\partial \underline{w}} + 1 \right], \end{aligned}$$

where the second line uses $\frac{\partial \lambda^*}{\partial Q} = H_3 \frac{\partial \lambda^*}{\partial \underline{w}}$. Substituting

$$\frac{dQ^*(\underline{w})}{d\underline{w}} = \frac{H_3 \frac{\partial \lambda^*}{\partial \underline{w}}}{V' - H_3 \frac{\partial \lambda^*}{\partial Q}}$$

yields

$$\frac{dq_2^*(\lambda^*(Q^*(\underline{w}),\underline{w}))}{d\underline{w}} = \frac{\partial q_2^*}{\partial \lambda}\frac{\partial \lambda^*}{\partial \underline{w}}\left[\frac{(H_3)^2\frac{\partial \lambda^*}{\partial \underline{w}}}{V' - H_3\frac{\partial \lambda^*}{\partial Q}} + 1\right]$$
$$= \frac{\partial q_2^*}{\partial \lambda}\frac{\partial \lambda^*}{\partial \underline{w}}\left[\frac{(H_3)^2\frac{\partial \lambda^*}{\partial \underline{w}} + V' - H_3\frac{\partial \lambda^*}{\partial Q}}{V' - H_3\frac{\partial \lambda^*}{\partial Q}}\right].$$

Since $\frac{\partial q_2^*}{\partial \lambda} < 0 < \frac{\partial \lambda^*}{\partial \underline{w}}$, $\frac{dq_2^*(\lambda^*(Q^*(\underline{w}),\underline{w}))}{d\underline{w}} < 0$ holds if the term in brackets is positive. To see that this is the case, substitute $\frac{\partial \lambda^*}{\partial Q} = H_3\frac{\partial \lambda^*}{\partial \underline{w}}$ again to obtain

$$\frac{dq_2^*(\lambda^*(Q^*(\underline{w}),\underline{w}))}{d\underline{w}} = \frac{\partial q_2^*}{\partial \lambda}\frac{\partial \lambda^*}{\partial \underline{w}}\left[\frac{V'}{V' - (H_3)^2\frac{\partial \lambda^*}{\partial \underline{w}}}\right].$$

Since $V' < 0$ and $V' - (H_3)^2\frac{\partial \lambda^*}{\partial \underline{w}} < 0$, we have

$$\frac{dq_2^*(\lambda^*(Q^*(\underline{w}),\underline{w}))}{d\underline{w}} < 0$$

as desired and required.

That the minimum wage increase decreases involuntary unemployment is now an implication of the fact that $Q^*(\underline{w})$ increases and q_2^* decreases in $\underline{w}$.

□

A.5 Proof of Lemma 5

Proof. The statement has been established in the proof of Lemma 2. □

A.6 Proof of Proposition 4

Proof. By construction of $\hat{Q}_H(m)$, for $\underline{w} \in W(\hat{Q}_H(m), W(Q_2(m)]$, there will be no wage dispersion. Consequently, all the stated effects follow from standard monopsony pricing with market-clearing wages in the face of a min wage. □

Note A.1. *Assuming that $Q^* \in (Q_1(m), Q_2(m))$, we now briefly discuss the effects of minimum wages above $W(Q_2(m))$. If $Q_2(m) \geq Q^p$ then statement (ii) from Proposition 4 still holds in this case. If $Q_2(m) < Q^p$ and there is no additional ironing range between $Q_2(m)$ and Q^p,[39] then increasing the minimum wage $\underline{w}$ with $\underline{w} \in [W(\hat{Q}_H), W(Q^p))$ increases employment without inducing involuntary unemployment and wage dispersion. Increasing the minimum*

[39]Formally, that is, if $(Q_2(m), Q^p] \cap \bigcup_{m' \in \mathcal{M}}(Q_1(m'), Q_2(m')) = \emptyset$.

wage beyond $W(Q^p)$ will induce involuntary unemployment but no wage dispersion. To see this, notice that if $Q^p \in (Q_1(m'), Q_2(m'))$ for some $m' \in \mathcal{M}$, then we have $\hat{Q}_H(m') \leq Q^p$ by the same arguments as those underlying Lemma 4. Consequently, for any $Q \geq \hat{Q}_H(m')$, which corresponds, to $\underline{w} \geq W(\hat{Q}_H(m'))$, there will be no wage dispersion and $Q^*(\underline{w})$ is such that $V(Q^*(\underline{w}))) = \underline{w}$. If $Q^p < Q_1(m')$, then no minimum wage $\underline{w} \in [W(Q_1(m')), W(Q_2(m'))]$ will induce wage dispersion because $\gamma(Q; m') > V(Q)$ for all $Q \in [Q_1(m'), Q_2(m')]$. This follows from the facts that (i) $V(Q^p) = W(Q^p)$, (ii) V is decreasing and W is increasing, and (iii) $\gamma(Q; m') \geq W(Q)$.

A.7 Proof of Proposition 5

Proof. Over intervals with an even index, we have $V(Q) < \gamma(Q; m)$, implying that there is no wage dispersion and no involuntary unemployment, while over intervals with an odd index, we have $V(Q) > \gamma(Q; m)$, implying that there is wage dispersion and involuntary unemployment. □

A.8 Proof of Proposition 6

Proof. Firm i's first-order condition is

$$V(x_i) = \frac{Q - x_i}{Q^2}\underline{C}(Q) + \frac{x_i}{Q}\underline{C}'(Q).$$

The left-hand side is decreasing in x_i. The partial derivative of the right-hand side with respect to x_i is $-\frac{1}{Q^2}(\underline{C}(Q) - Q\underline{C}'(Q))$, which is positive because $\underline{C}$ is convex. This implies that for any aggregate quantity Q there is a unique x_i that satisfies the first-order condition. This x_i must thus be the same for all i. Hence, any equilibrium is symmetric. Given this, we can write the first-order condition as

$$V\left(\frac{Q}{n}\right) = \frac{n-1}{n}\frac{\underline{C}(Q)}{Q} + \frac{1}{n}\underline{C}'(Q). \tag{14}$$

The left-hand side is decreasing in Q. The derivative of the right-hand side with respect to Q is

$$-\frac{n-1}{nQ^2}\left(\underline{C}(Q) - Q\underline{C}'(Q)\right) + \frac{1}{n}\underline{C}''(Q) \geq 0, \tag{15}$$

where the inequality follows because $\underline{C}$ is convex, implying $\underline{C}'' \geq 0$ and $Q\underline{C}'(Q) \geq \underline{C}(Q)$. Because at $Q = 0$, the left-hand side is larger than the right-hand side, there is a unique Q that satisfies (14). This proves that the equilibrium is unique and symmetric.

To see that Q_n^* is increasing in n, suppose to the contrary that it is not and we have $Q_n^* \geq Q_{n+1}^*$ for some n. This implies $\frac{Q_n^*}{n} > \frac{Q_{n+1}^*}{n+1}$ and therefore

$$\begin{aligned} V\left(\frac{Q_{n+1}^*}{n+1}\right) > V\left(\frac{Q_n^*}{n}\right) &= \frac{n-1}{n}\frac{\underline{C}(Q_n^*)}{Q_n^*} + \frac{1}{n}\underline{C}'(Q_n^*) \\ &\geq \frac{n-1}{n}\frac{\underline{C}(Q_{n+1}^*)}{Q_{n+1}^*} + \frac{1}{n}\underline{C}'(Q_{n+1}^*) \\ &\geq \frac{n}{n+1}\frac{\underline{C}(Q_{n+1}^*)}{Q_{n+1}^*} + \frac{1}{n+1}\underline{C}'(Q_{n+1}^*). \end{aligned}$$

Here, the first weak inequality is due to (15) and the second follows from the fact that the derivative of $\frac{n-1}{n}\frac{\underline{C}(Q)}{Q} + \frac{1}{n}\underline{C}'(Q)$ with respect to n is

$$\frac{1}{n^2 Q}\left[\underline{C}(Q) - Q\underline{C}'(Q)\right] \leq 0,$$

where the inequality holds because $\underline{C}(Q)$ is convex. Since in equilibrium

$$V\left(\frac{Q_{n+1}^*}{n+1}\right) = \frac{n}{n+1}\frac{\underline{C}(Q_{n+1}^*)}{Q_{n+1}^*} + \frac{1}{n+1}\underline{C}'(Q_{n+1}^*),$$

we have the desired contradiction.

That $Q_n^p < Q_n^*$ for n sufficiently small follows from the discussion preceding the proposition by choosing $n = 1$ since $h(Q,1) > W(Q)$ for all $Q \in (Q_1(m), Q_2(m))$. Moreover, $Q_n^p \leq Q_n^*$ requires $Q_n^* \in (Q_1(m), Q_2(m))$ for some $m \in \mathcal{M}$ since otherwise $h(Q,n) = W(Q) + \frac{Q}{n}W'(Q)$, which implies $Q_n^* < Q_n^p$. The arguments preceding the proposition imply that $h(Q,n) < W(Q)$ for some $Q \in (Q_1(m), Q_2(m))$ can only occur if n is large enough.

Assume now that $\underline{C}(Q^e) = C(Q^e)$ and take $Q_\infty := \lim_{n\to\infty} Q_n^*$. Taking limits of both sides of (14) yields

$$V(0) = \frac{\underline{C}(Q_\infty)}{Q_\infty}. \tag{16}$$

The definition of Q^e then implies that $V(0) = \frac{\underline{C}(Q_\infty)}{Q_\infty} = W(Q^e) = \frac{\underline{C}(Q^e)}{Q^e}$. Using

$$\frac{d}{dQ}\left(\frac{\underline{C}(Q)}{Q}\right) = \frac{Q\underline{C}'(Q) - \underline{C}(Q)}{Q^2} \geq 0,$$

where the inequality holds because $\underline{C}$ is convex, we have that the solution to the equation $V(0) = \frac{\underline{C}(Q_\infty)}{Q_\infty}$ is unique. Since Q^e satisfies this equation we thus have $Q_\infty = Q^e$. Hence, if $Q^e \notin \cup_{m\in\mathcal{M}}(Q_1(m)), Q_2(m))$ then Q^e is also the aggregate quantity in the limit as claimed.

Assume now that $Q^e \in (Q_1(m_e)), Q_2(m_e))$ for some $m_e \in \mathcal{M}$. For $Q \in (Q_1(m_e)), Q_2(m_e)),$

$\underline{C}(Q)$ increases linearly from $C(Q_1(m_e))$ to $C(Q_2(m_e))$ with a slope that is greater than $V(0)$. The latter follows from our observation that $\underline{C}'(Q^e) > V(0)$, which appeared immediately prior to the proposition statement. Using that W is increasing we have

$$\frac{C(Q_1(m_e))}{Q_1(m_e)} = W(Q_1(m_e)) < W(Q^e) = V(0) < W(Q_2(m_e)) = \frac{C(Q_2(m_e))}{Q_2(m_e)}.$$

This implies that there exists a unique number $\tilde{Q} \in (Q_1(m_e)), Q_2(m_e))$ such that $\frac{\underline{C}(\tilde{Q})}{\tilde{Q}} = V(0)$. This is thus the aggregate quantity in the limit if $Q^e \in (Q_1(m_e)), Q_2(m_e))$ as claimed.

We are left to show that in case $Q^e \in (Q_1(m_e), Q_2(m_e))$, $\tilde{Q} > Q^e$. To see that this holds, rearrange (16) to

$$Q_\infty V(0) = \underline{C}(Q_\infty)$$

and recall that $Q^e V(0) = C(Q^e)$. Since $C(Q^e) > \underline{C}(Q^e)$, $\tilde{Q} = Q_\infty > Q^e$ follows. □

A.9 Proof of Proposition 7

Proof. We show that $\frac{d\underline{Q}_n(\underline{w})}{d\underline{w}}|_{\underline{w}=w_1(Q_n^*;m)} > 0$. Totally differentiating (9) with respect to $\underline{w}$, dropping arguments and writing $\underline{C}'$ and $\underline{C}''$ in lieu of $\frac{\partial \underline{C}}{\partial Q}$ and $\frac{\partial^2 \underline{C}}{\partial Q^2}$ yields

$$\left[V' - (n-1)\left[\frac{\underline{Q}_n \underline{C}' - \underline{C}}{\underline{Q}_n^2}\right] - \underline{C}''\right]\frac{d\underline{Q}_n}{d\underline{w}} = (n-1)\frac{\partial \underline{C}}{\partial \underline{w}}\frac{1}{\underline{Q}_n} + \frac{\partial \underline{C}'}{\partial \underline{w}}.$$

Since the term in brackets on the left-hand-side is negative, $\frac{d\underline{Q}_n}{d\underline{w}}$ has the opposite sign of $(n-1)\frac{\partial \underline{C}}{\partial \underline{w}}\frac{1}{\underline{Q}_n} + \frac{\partial \underline{C}'}{\partial \underline{w}}$. From the proof of Lemma 2 , we know that $\frac{\partial \underline{C}}{\partial \underline{w}} = \lambda^* \geq 0$ and $\frac{\partial \underline{C}'}{\partial \underline{w}} = \frac{\partial \lambda^*}{\partial Q} \leq 0$, where λ^* is the solution value of the Lagrange multiplier associated with the minimum wage constraint. At $\underline{w} = w_1(Q;m)$, we have $\lambda^* = 0$ and $\frac{\partial \lambda^*}{\partial Q} < 0$, whence $\frac{d\underline{Q}_n}{d\underline{w}}|_{\underline{w}=w_1(Q_n^*;m)} > 0$ follows.

□

A.10 Proof of Proposition 9

Proof. We prove the proposition statement by statement.

When wage discrimination is prohibited, the monopsony optimally procures the quantity $\tilde{Q}$ at each location satisfying $V(\tilde{Q}) = C'(\tilde{Q})$, provided $\tilde{Q} \leq 1/2$. Otherwise, we have $\tilde{Q} = 1/2$. Since $V(1/2) < 1$ is equivalent to $\tilde{Q} < 1/2$ and $Q^* > \tilde{Q}$ holds whenever $V(1/4) > 1/4$ and $V(1/2) < 1$, the statement follows.

That the monopsony's profit decreases when wage discrimination is prohibited follows simply because for $V(1/4) > 1/2$, wage discrimination is strictly optimal.

With wage discrimination, only workers with $x \in [0, 1/4)$ and $x \in (3/4, 1]$ enjoy a positive surplus. All other workers are indifferent between working and not working and hence have a surplus of 0. When wage discrimination is prohibited, all workers with $x \in [0, \tilde{Q})$ and $x \in (1 - \tilde{Q}, 1]$ enjoy a positive surplus and are paid a wage of of $\tilde{Q}$, which is larger than 1/4 since $V(1/4) > 1/2$ implies $\tilde{Q} > 1/4$. □

A.11 Proof of Proposition 10

Proof. Note first that $\Delta SS(\tilde{Q}, \tilde{Q}) < \Delta C(\tilde{Q}, \tilde{Q})$ is equivalent to $\tilde{Q}(2 - \tilde{Q}) > 5/16$, which is equivalent to $\tilde{Q} \in (1/4, 5/12)$. Together with the slope condition (10), $\Delta SS(\tilde{Q}, \tilde{Q}) < \Delta C(\tilde{Q}, \tilde{Q})$ implies the statement in the proposition. Consequently, the proof is complete if we can show that for $\tilde{Q} > 5/12$, $\Delta SS(Q^*, \tilde{Q}) < 0$. To see that this is the case, notice that because $V(Q)$ is decreasing, for all $x > \tilde{Q}$, we have $V(x) \le V(\tilde{Q}) = 2\tilde{Q}$, where the equality uses the first-order condition for $\tilde{Q}$. This implies the first inequality in the following display

$$\begin{aligned} \Delta SS(Q^*, \tilde{Q}) &\le (2\tilde{Q} - 1/2)(Q^* - \tilde{Q}) - \frac{1}{2}\tilde{Q}(1 - \tilde{Q}) + \frac{3}{32} \\ &\le (2\tilde{Q} - 1/2)(1/2 - \tilde{Q}) - \frac{1}{2}\tilde{Q}(1 - \tilde{Q}) + \frac{3}{32}, \end{aligned}$$

where the second inequality follows because the righthand side in the first line increases in Q^* since $\tilde{Q} > 1/4$. Because

$$(2\tilde{Q} - 1/2)(1/2 - \tilde{Q}) - \frac{1}{2}\tilde{Q}(1 - \tilde{Q}) + \frac{3}{32} = -\frac{5}{32} + \frac{2\tilde{Q} - 3\tilde{Q}^2}{2},$$

the righthand side of which is decreasing for $\tilde{Q} > 1/3$ and negative at $\tilde{Q} = 5/12$.

Thus, for all $\tilde{Q} \in (1/4, 5/12)$, we have $\Delta SS(\tilde{Q}, \tilde{Q}) < \Delta C(\tilde{Q}, \tilde{Q})$, which jointly with (10) implies $\Delta SS(Q^*, \tilde{Q}) < \Delta C(Q^*, \tilde{Q})$, and for all $\tilde{Q} \ge 5/12$, we have $\Delta SS(Q^*, \tilde{Q})$. □